ELSEVIER Journal of Computational and Applied Mathematics 77 (1997) 5–34

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Recursiveness in matrix rational interpolation problems[1]

Bernhard Beckermann[a,*], George Labahn[b]

[a] *Laboratoire d'Analyse Numérique et d'Optimisation, UFR IEEA – M3, USTL Flandres Artois, F-59655 Villeneuve d'Ascq CEDEX, France*

[b] *Department of Computer Science, University of Waterloo, Waterloo, Ont., Canada*

Abstract

We consider the problem of computing solutions to a variety of matrix rational interpolation problems. These include the partial realization problem for matrix power series and Newton–Padé, Hermite–Padé, simultaneous Padé, M-Padé and multipoint Padé approximation problems along with their matrix generalizations. A general recurrence relation is given for solving these problems. Unlike other previous recursive methods, our recurrence works along arbitrary computational paths. When restricted to specific paths, the recurrence relation generalizes previous work of Antoulas, Cabay and Labahn, Beckermann, Van Barel and Bultheel and Gutknecht along with others.

Our results rely on the concept of *extended M-Padé approximation* introduced in this paper. This is a natural generalization of the two-point Padé approximation problem extended to multiple interpolation points (including infinity) and matrix Laurent and Newton series. By using module-theoretic techniques we determine complete parameterizations of all solutions to this problem. Our recurrence relation then efficiently computes these parameterizations. This recursion requires no conditions on the input data.

We also discuss the concept of duality which was shown to be of particular interest for a stable computation of those approximants. Finally, we show the invariance of our approximation problem under linear transformations of the extended complex plane.

Keywords: Partial realization; Hermite Padé approximant; Simultaneous Padé approximant; Matrix Padé approximant; Newton–Padé approximant; Multipoint Padé approximant

AMS classification: 65D05; 41A21; CR: G.1.2

1. Introduction

Let m be an integer with $m \geqslant 2$, $\mathbb{F}$ a field and $\mathbb{F}_0$ a (finite or infinite) subset of $\mathbb{F}$ (the set of "knots" or interpolation points). Throughout this paper, we will assume that we have $m \times m$ matrices $\boldsymbol{G}$ and $\boldsymbol{H}$, where each entry of $\boldsymbol{G}$ has an expansion as a formal Newton series in z while each entry of $\boldsymbol{H}$ is a right-truncated Laurent series in z, and $\det \boldsymbol{G} \neq 0 \neq \det \boldsymbol{H}$.

* Corresponding author. E-mail: bbecker@ano.univ.lille1.fr.

[1] This work was partially supported by the HCM project ROLLS, under contract CHRX-CT93-0416.

PII S0377-0427(96)00120-3

We study the *extended M-Padé approximation problem* for $\boldsymbol{G}$ and $\boldsymbol{H}$. Roughly speaking, this is the problem of finding an m-vector $\boldsymbol{Q}$ of polynomials satisfying

$$\boldsymbol{G} \cdot \boldsymbol{Q} = \mathrm{O}_+(\boldsymbol{\omega}),$$

$$\boldsymbol{H} \cdot \boldsymbol{Q} = \mathrm{O}_-(z^u),$$

where u is an integer and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_m)$ with polynomial components $\omega_j \in \mathbb{F}[z]$ having only zeros from $\mathbb{F}_0$.

Using various choices of $\boldsymbol{\omega}, \boldsymbol{G}$ and $\boldsymbol{H}$, extended M-Padé approximation generalizes a wide variety of approximation and interpolation problems. These include two-point Padé approximation [30], Hermite–Padé and simultaneous Padé approximation [19, 29, 39], Newton–Padé and multipoint Padé approximation [21] and also the partial realization problem for matrix sequences [1]. These rational approximation and interpolation problems are used in a wide variety of applications (cf. [2]).

By using module-theoretic techniques we are able to give a complete description of the space of solutions of the extended M-Padé approximation problem. The results generalize previous work by one of the authors [3] in the case of scalar M-Padé approximation at finite points. Special cases of our results have also been given in [5, 8, 9, 20, 36], all also making use of module-theoretic tools. In the case of two-point Padé, multipoint Padé or Newton–Padé approximation, these results provide simple proofs that the corresponding Padé-like tables have a block structure with unique rational forms inside each block (cf. [17, 21, 30]).

We may distinguish between two different kinds of algorithms for solving the above mentioned approximation problems. First there are single-step methods (see, e.g., [5, 9, 20, 33–36, 39]) where one is interested to compute a sequence of neighboring entries of the respective solution table (or, in case of singularities, a maximal subsequence). Here an elementary step consists in solving an interpolation problem obtained by adding one interpolation condition and/or by changing the degree constraints by units. For example, the algorithm given in [9] computes power Hermite–Padé approximants of type $(n_1, \ldots, n_m)$ by recursively solving subproblems on some diagonal path, namely of type $(n_1(\delta), \ldots, n_m(\delta))$ where $n_i(\delta) = \max(0, n_i - \delta + 1)$. Alternate computational paths for single-step methods are considered in [5, 36]. However, single-step methods have the drawback that one may get poor numerical results if a singular block in the respective solution table is not correctly detected, or if one encounters ill-conditioned subproblems.

A second class of methods is given by the hybrid methods described in [1, 7, 9, 12–16, 21–23, 25, 27, 37, 38]. Here one solves the original interpolation problem by recursively dividing a single interpolation basis problem into two smaller interpolation problems. One of these two problems will be of the same type as the original one, and the other is usually solved by building up a 'small' system of linear equations, which then is solved by some stable classical method such as Gaussian elimination. This method has the advantage that by some "look-ahead" techniques one may also avoid ill-conditioned subproblems, leading to weakly stable algorithms [7, 13, 16, 38].

On the other hand, the above hybrid methods are only based on a proper "divide and conquer" approach, namely, breaking the original problem into two problems of the same type, if one follows diagonal paths. However, in many applications alternate computational paths are desired, motivated for example by convergence results such as the theorem of Montessus de Ballore. As another example, diagonal paths for Padé computation result in Hankel matrix solvers while Padé computation along straight line row paths result in Toeplitz matrix solvers. The latter are often more useful especially in the cases where the Toeplitz matrices have added structure (for example, positive-definite so all leading principal minors are nonsingular) not inherited by their Hankel counterparts.

Our main result is the use of extended M-Padé approximation to solve the problem of recursively computing matrix rational interpolation problems on arbitrary paths by applying the "divide and conquer" principle, i.e., we give an algorithm that divides a single interpolation problem into two smaller interpolation problems of the same type. This generalizes previous work of Gutknecht [23] for the problem of (scalar) multipoint Padé approximation (a generalization of two-point Padé approximation).

The primary tool in our approach is a so-called interpolation basis for an extended M-Padé problem. These are bases of the module of all solutions of our interpolation problems. We also introduce the notion of *normal data* for our interpolation problem. These are cases which allow for unique solutions of an interpolation problem, at least up to normalization. We also discuss dual interpolation problems along with their interpolation bases. These are one of the fundamental requirements in using the Cabay–Meleshko approach for creating weakly stable arbitrary path algorithms for interpolation problems. Such an algorithm will be presented in a later publication. Finally, we show that our formalism for an interpolation problem in the extended complex plane is fully invariant under Moebius transformations.

The remainder of this paper is divided as follows. In the next section we use concepts from module theory to study solutions of only the order condition. We introduce the concept of an order basis and transfer matrices from one basis to another one. Section 3 discusses interpolation at infinity using the notion of $\boldsymbol{H}$-degree and $\boldsymbol{H}$-reduced, introduced in this paper. Section 4 provides a characterization of our bases in terms of these degree concepts and shows how they are related to the usual degree bounds found in most rational interpolation problems. Section 5 gives our recurrence for computing such bases along arbitrary paths while Section 6 discusses duality and the invariance under Moebius transforms. The closing section gives some topics for future research.

Notation. For a space $\mathscr{M}$ with scalars from the field $\mathbb{F}$ (for instance $\mathscr{M} = \mathbb{F}^{p\times q}$, the space of $p \times q$ matrices over $\mathbb{F}$), $\mathscr{M}[z]$ will denote the set of polynomials in z with coefficients from $\mathscr{M}$ while $\mathscr{M}[[z]]_{\mathbb{F}_0}$ represents the set of formal Newton series in z with coefficients from $\mathscr{M}$. The latter is specified (with respect to $\mathbb{F}_0$) as follows: $\boldsymbol{G} \in \mathscr{M}[[z]]_{\mathbb{F}_0}$ iff for all $z_0 \in \mathbb{F}_0$ and all $k \in \mathbb{N}_0$ the kth derivative of $\boldsymbol{G}$ at z_0 is known and is an element of $\mathscr{M}$. Note that $\mathscr{M}[z] \subset \mathscr{M}[[z]]_{\mathbb{F}_0}$ and that if $\mathscr{M}$ is an algebra, then $\mathscr{M}[z]$ and $\mathscr{M}[[z]]_{\mathbb{F}_0}$ are also algebras (multiplication being the classical product rule). Because much of our work involves square matrices we also set $\mathscr{F} := \mathbb{F}^{m\times m}$. In addition we define

$$\mathscr{M}[[z]]_- := \left\{ A(z) = \sum_{k=-\infty}^{\infty} a_k z^k : a_k \in \mathscr{M}, \exists K \text{ with } A_k = 0 \; \forall k > K \right\},$$

the set of right-truncated matrix Laurent series and its subset $\mathscr{M}[[z]]_\infty$ containing formal power series in the variable z^{-1}.

2. Order constraints

2.1. The definition of order

The type of interpolation problems that will be considered are specified by requirements that solutions satisfy both certain order constraints and also bounds on their degrees. However, in this section we look only at the order conditions and disregard any other constraints until later sections. For the case $\mathbb{F}_0 = \{0\}$ order constraints are classically specified as equations of the form

$$\boldsymbol{G} \cdot \boldsymbol{P} = \mathrm{O}_+(z^N) = c_N z^N + c_{N+1} z^{N+1} + \cdots$$

for an $m \times 1$ vector $\boldsymbol{P}$ and $1 \times m$ matrix of power series $\boldsymbol{G}$. Order constraints for the vector problem are similar to the above, except that $\boldsymbol{G}$ will be an $s \times m$ matrix with $s \leqslant m$. Note that the order constraints are specified on individual rows, hence by padding the order conditions if necessary (cf. Example 2.3) one can make the assumption that $\boldsymbol{G}$ is a square matrix. In addition, the $\boldsymbol{G}$ of interest in applications usually also has certain invertibility properties.

Definition 2.1. A formal Newton series $\boldsymbol{G} \in \mathscr{F}[[z]]_{\mathbb{F}_0}$ is called *regular* if it has an inverse in $\mathscr{F}[[z]]_{\mathbb{F}_0}$ or, equivalently, if $\det \boldsymbol{G}$ does not vanish at any element of $\mathbb{F}_0$. Similarly, a formal power series around infinity $\boldsymbol{H} \in \mathscr{F}[[z]]_\infty$ is regular if $\boldsymbol{H}(\infty)$ is regular.

A vector $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_m)$ of polynomials is called an order vector (with respect to $\mathbb{F}_0$) if each component ω_j is a monic polynomial with all zeros being elements of $\mathbb{F}_0$.

Definition 2.2. The matrix polynomial $\boldsymbol{P} \in \mathbb{F}^{m \times s}[z]$ is said to have order $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_m)$ (with respect to $\boldsymbol{G} \in \mathscr{F}[[z]]_{\mathbb{F}_0}$) if

$$\boldsymbol{G} \cdot \boldsymbol{P} = \mathrm{diag}\,(\omega_1, \ldots, \omega_m) \cdot \boldsymbol{R} \quad \text{with } \boldsymbol{R} \in \mathbb{F}^{m \times s}[[z]]_{\mathbb{F}_0}. \tag{1}$$

To be more precise, we also will use the notation that $\boldsymbol{P}$ has order $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{R})$, and $\boldsymbol{R}$ will be called the *order residual* of $\boldsymbol{P}$. In addition, we define

$$\mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G}) := \{\boldsymbol{P} \in \mathbb{F}^{m \times 1}[z] : \boldsymbol{P} \text{ has order } \boldsymbol{\omega}\},$$

the set of all elements in $\mathbb{F}^{m \times 1}[z]$ of order $\boldsymbol{\omega}$ with respect to $\boldsymbol{G}$.

Example 2.3 (*M-Padé approximants, type*-I *Hermite–Padé approximants* [29]). Given $z_0, z_1, \ldots \in \mathbb{F}_0$, formal Newton series $f_1, \ldots, f_m \in \mathbb{F}[[z]]_{\mathbb{F}_0}$, and a vector $\boldsymbol{n} = (n_1, \ldots, n_m)$ of non-negative integers, an *M-Padé approximant* of type $\boldsymbol{n}$ is a column vector $\boldsymbol{P} = (P_1, \ldots, P_m)^{\mathrm{T}}$ of polynomials with degree of P_j being bounded by n_j, $j = 1, \ldots, m$, such that the polynomial linear combination $f_1 \cdot P_1 + \cdots + f_m \cdot P_m$ has the zeros $z_0, \ldots, z_N$ counting multiplicities, where $N = n_1 + \cdots + n_m + m - 2$. In the particular case $\mathbb{F}_0 = \{0\}$ we obtain type-I Hermite–Padé approximants. These were studied by Hermite in 1873. Note also that when $m = 2$ and $f_2 = -1$ then we have classical Newton–Padé approximants.

Setting $\boldsymbol{F} := (f_1, \ldots, f_m)$, $\omega(z) := (z - z_0) \cdot \cdots \cdot (z - z_N)$, the above is equivalent to specifying that $\boldsymbol{F} \cdot \boldsymbol{P}$ contains the factor ω, that is, $\boldsymbol{F} \cdot \boldsymbol{P} = \omega \cdot R$ with $R \in \mathbb{F}[[z]]_{\mathbb{F}_0}$. In accordance with

Definition 2.2 we may extend this order condition to matrix form by taking $\boldsymbol{\omega} = (\omega, 1, \ldots, 1)$ and

$$\boldsymbol{F} = (f_1, \ldots, f_m) \in \mathbb{F}^{1\times m}[[z]]_{\mathbb{F}_0}, \qquad \boldsymbol{F}_1 = (\boldsymbol{0}\ \boldsymbol{I}) \in \mathbb{F}^{(m-1)\times m}[z], \qquad \boldsymbol{G} = \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{F}_1 \end{pmatrix}, \tag{2}$$

such that $\boldsymbol{G} \in \mathscr{F}[[z]]_{\mathbb{F}_0}$, where $\boldsymbol{I}$ and $\boldsymbol{0}$ denote identity and zero matrices of suitable size. Note that in order to obtain a regular $\boldsymbol{G}$ we have to assume that at least one f_i has no zero from $\mathbb{F}_0$ and, by renumbering if necessary, we may assume that f_1 has this property.

Example 2.4 (*Matrix Padé approximants* [2, 27]). Let $\mathbb{F}_0 = \{0\}$ and $\boldsymbol{A}$ be an $s \times s$ matrix power series. Then a right-hand matrix Padé form for $\boldsymbol{A}$ of type (p, q) is a pair $(\boldsymbol{U}, \boldsymbol{V})$ of matrix polynomials of size $s \times s$ having degree bounds $\deg \boldsymbol{U} \leqslant p$, and $\deg \boldsymbol{V} \leqslant q$, respectively, and satisfying

$$\boldsymbol{A} \cdot \boldsymbol{V} - \boldsymbol{U} = \mathrm{O}(z^{p+q+1}).$$

Let $m = 2s$ and set $\boldsymbol{\omega} = (z^{p+q+1}, \ldots, z^{p+q+1}, 1, \ldots, 1)$, a vector of length m having s ones,

$$\boldsymbol{G} = \begin{pmatrix} -\boldsymbol{I} & \boldsymbol{A} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \in \mathscr{F}[[z]]_{\mathbb{F}_0}, \tag{3}$$

where $\boldsymbol{I}$ and $\boldsymbol{0}$ denote identity and zero matrices of appropriate size so that $\boldsymbol{G}$ is $m \times m$. Then $\boldsymbol{G}$ is regular, and the set $\mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G})$ describes the possible columns of the combined matrix $(\boldsymbol{U}^{\mathrm{T}}, \boldsymbol{V}^{\mathrm{T}})^{\mathrm{T}}$.

Generalizations of Example 2.3 to vector-valued data such as the vector Hermite–Padé or vector M-Padé problems have been studied [8, 9, 20, 36]. Also, notations similar to those of Example 2.4 have been introduced for the case where $\boldsymbol{A}$ is an $(r \times s)$ rectangular matrix power series.

2.2. Order modules and their bases

Our interest is not so much in determining a single solution to a given rational interpolation problem but rather in characterizing all such solutions. This is particularly useful if one wants to classify singular cases. Such a characterization has been given before, for example for the particular case of Newton–Padé approximation [18, 21], for the M-Padé approximation problem [3–5, 34], and for vector-valued generalizations [8, 9, 36]. Here we refine ideas proposed in [3, 34] by using module-theoretic properties. For details of module theory we refer the reader to [28]. Roughly speaking, many results for linear spaces still are valid if one considers as the set of scalars a ring instead of a field.

In order to be self-contained, we summarize and prove in the following lemma the required assertions. We assume $\mathbb{D}$ to be a principal ideal domain (e.g., $\mathbb{D} = \mathbb{F}[z]$).

Lemma 2.5. (a) *Each submodule $\mathscr{S}$ of $\mathbb{D}^m$ has a basis $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_\mu$ of $\mu \leqslant m$ elements, i.e., for each $\boldsymbol{P}_0 \in \mathscr{S}$ there exist unique $\alpha_1, \ldots, \alpha_\mu \in \mathbb{D}$ such that $\boldsymbol{P}_0 = \boldsymbol{P}_1 \cdot \alpha_1 + \cdots + \boldsymbol{P}_\mu \cdot \alpha_\mu$.*

(b) *Let the submodule $\mathscr{S}$ of $\mathbb{D}^m$ have a basis of m elements, arranged as columns in a matrix from $\mathbb{D}^{m\times m}$. Then there exists a $\chi(\mathscr{S}) \in \mathbb{D}$ such that*

$$\{\alpha \cdot \chi(\mathscr{S}) : \alpha \in \mathbb{D}\} = \{\det \boldsymbol{P} : \boldsymbol{P} \in \mathbb{D}^{m\times m} \text{ and columns of } \boldsymbol{P} \text{ are in } \mathscr{S}\}. \tag{4}$$

Moreover, $\boldsymbol{P}$ is a basis of $\mathscr{S}$ if and only if $\alpha \in \mathbb{D}$, being defined by $\det \boldsymbol{P} = \alpha \cdot \chi(\mathscr{S})$, has an inverse in $\mathbb{D}$. Equivalently, a basis of $\mathscr{S}$ is unique up to multiplication on the right by a unimodular matrix. [2]

Proof. (a) The assertion is obvious for $m = 1$. If $m \geqslant 2$, let

$$\mathscr{S}' = \{\boldsymbol{P}_m : (\boldsymbol{P}_1, \dots, \boldsymbol{P}_m) \in \mathscr{S}\} \subset \mathbb{D}.$$

$\mathscr{S}$ is a submodule of $\mathbb{D}^m$, hence $\mathscr{S}'$ is an ideal. By assumption on $\mathbb{D}$, there exists a $\boldsymbol{P}_1^* = (\boldsymbol{P}_{1,1}^*, \dots, \boldsymbol{P}_{1,m}^*) \in \mathscr{S}$ such that $\mathscr{S}'$ is generated by $\boldsymbol{P}_{1,m}^*$. Since the submodule $\mathscr{S}^*$ of $\mathscr{S}$,

$$\mathscr{S}^* = \{\boldsymbol{P} = (\boldsymbol{P}_1, \dots, \boldsymbol{P}_m) \in \mathscr{S} : \boldsymbol{P}_m = 0\}$$

can also be understood as a submodule of $\mathbb{D}^{m-1}$, assertion (a) follows by induction.

(b) Let $\boldsymbol{P}, \boldsymbol{P}^* \in \mathbb{D}^{m \times m}$ have columns from $\mathscr{S}$, with $\boldsymbol{P}^*$ forming a basis of $\mathscr{S}$. Then by definition of a basis there exists a matrix $\boldsymbol{U} \in \mathbb{D}^{m \times m}$ such that $\boldsymbol{P} = \boldsymbol{P}^* \cdot \boldsymbol{U}$. Taking determinants yields (4) with $\chi(\mathscr{S}) := \det \boldsymbol{P}^*$. If, in addition, $\boldsymbol{P}$ is a basis, then there also exists a matrix $\boldsymbol{V} \in \mathbb{D}^{m \times m}$ such that $\boldsymbol{P}^* = \boldsymbol{P} \cdot \boldsymbol{V}$ and therefore $\boldsymbol{P} \cdot (\boldsymbol{V} \cdot \boldsymbol{U} - \boldsymbol{I}) = \boldsymbol{P}^* \cdot (\boldsymbol{U} \cdot \boldsymbol{V} - \boldsymbol{I}) = 0$. Consequently, $\boldsymbol{U}$ must be unimodular. □

Theorem 2.6. *The set $\mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G})$ is a submodule of the module $\mathbb{F}^m[z]$ of polynomial column vectors with respect to the ring $\mathbb{F}[z]$ of univariate polynomials, and $\mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G})$ has a basis of m elements. A matrix $\boldsymbol{P} \in \mathbb{F}^{m \times m}[z]$ with columns $\boldsymbol{P}_1, \dots, \boldsymbol{P}_m$ forming a basis of $\mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G})$ will be called an* $(\boldsymbol{\omega}, \boldsymbol{G})$-basis. *Finally, for any $\boldsymbol{Q} \in \mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G})$ the polynomial coefficients α_j in the representation*

$$\boldsymbol{Q} = \boldsymbol{P}_1 \cdot \alpha_1 + \cdots + \boldsymbol{P}_m \cdot \alpha_m \tag{5}$$

may be calculated by $(j = 1, \dots, m)$

$$\alpha_j = \frac{\det(\boldsymbol{P}_1, \dots, \boldsymbol{P}_{j-1}, \boldsymbol{Q}, \boldsymbol{P}_{j+1}, \dots, \boldsymbol{P}_m)}{\det(\boldsymbol{P}_1, \dots, \boldsymbol{P}_m)}. \tag{6}$$

Proof. That $\mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G})$ is a submodule of $\mathbb{F}^m[z]$ is clear. Since $\mathbb{D} = \mathbb{F}[z]$ is a principal ideal domain, from Lemma 2.5(a) we may conclude that there exists a basis of $\mu \leqslant m$ elements. If we set $\Omega := \omega_1 \cdot \cdots \cdot \omega_m$, then the column vectors $\boldsymbol{P}_j := (0, \dots, \Omega, 0, \dots, 0)^{\mathrm{T}}$ trivially are linearly independent elements of $\mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G})$, considered as a vector space over the quotient field of rational functions. Therefore we have $\mu = m$. The representation (6) follows from Cramer's rule in the quotient field of rational functions. Both numerator and denominator of the representation (6) are polynomial multiples of $\chi(\mathscr{S})$ with the multiple being a constant in the case of a basis. Therefore each $\alpha_j \in \mathbb{F}[z]$, rather than from the quotient field. □

In the sequel, we will use for the generator of $\mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G})$ as defined in (4) the shorthand notation $\chi(\boldsymbol{\omega}, \boldsymbol{G})$, and we will suppose without loss of generality that $\chi(\boldsymbol{\omega}, \boldsymbol{G})$ is monic, which yields its uniqueness. As seen in Lemma 2.5(b), an $(\boldsymbol{\omega}, \boldsymbol{G})$-basis $\boldsymbol{P}$ must satisfy $\det \boldsymbol{P} = c \cdot \chi(\boldsymbol{\omega}, \boldsymbol{G})$ with $c \in \mathbb{F}$ being different from zero. Therefore let us have a closer look at this quantity $\chi(\boldsymbol{\omega}, \boldsymbol{G})$.

[2] By definition, a matrix $\boldsymbol{U} \in \mathbb{D}^{m \times m}$ is called unimodular if there exists an inverse $\boldsymbol{V} \in \mathbb{D}^{m \times m}$ with $\boldsymbol{U} \cdot \boldsymbol{V} = \boldsymbol{V} \cdot \boldsymbol{U} = \boldsymbol{I}$, or, equivalently, if $\det \boldsymbol{U}$ has an inverse in $\mathbb{D}$.

For M-Padé approximation, and F as in Example 2.3, it was proved in [5] that $\chi(\omega, G) = \omega_1 = \omega_1 \cdot \cdots \cdot \omega_m$, provided that we are able to find an extension G of F being regular with respect to a subset of $\mathbb{F}_0$ containing the zeros of ω_1 (or, equivalently, supposing that F does not vanish at any zero of ω_1). Similar to the ideas in [5], characterizations of $\chi(\omega, G)$ have already been given for arbitrary G (see, e.g., [9, 36]). We give here a simpler constructive proof for completeness.

Lemma 2.7. *Let ω be an order vector and $\Omega := \det \operatorname{diag} \omega$. Then there holds:*
(a) *provided that G is regular, $\chi(\omega, G)$ is a (polynomial) multiple of Ω;*
(b) *for all $G \in \mathcal{F}[[z]]_{\mathbb{F}_0}$, the generator $\chi(\omega, G)$ is a divisor of Ω.*

Proof. Assertion (a) follows immediately from applying the determinant function on (1) and using the fact that $\det G$ has no zeros from $\mathbb{F}_0$. We prove (b) by induction on $\deg \Omega$. In case $\deg \Omega = 0$ we have $\omega = (1, \ldots, 1)$ and $\chi(\omega, G) = 1$ since $\mathcal{M}(\omega, G) = \mathbb{F}^m[z]$. Hence suppose that $\deg \Omega > 0$. Then there exists $\omega' = (\omega'_1, \ldots, \omega'_m)$, and $\omega'_\ell(z) \cdot (z-a) = \omega_\ell$, whereas $\omega'_j = \omega_j$ for $j \neq \ell$, such that $\chi(\omega', G)$ is a divisor of $\Omega(z)/(z-a)$. If $\chi(\omega, G) = \chi(\omega', G)$ then the assertion is trivial. Otherwise there exists a $P' \in \mathcal{F}[z]$ with $\det P' = \chi(\omega', G)$ having order (ω', G, R'), but not order ω. Let $(R'_{\ell,1}, \ldots, R'_{\ell,m})$ constitute the ℓth row of R', then at least one component of the vector $(R'_{\ell,1}(a), \ldots, R'_{\ell,m}(a))$ is different from zero, say the πth. We define $U \in \mathcal{F}[z]$ by

$$U(z) := \begin{pmatrix} 1 & 0 \cdots & \cdots & \cdots & \cdots & \cdots\cdots & 0 \\ 0 & 1 \quad 0 & \cdots & \cdots & \cdots & \cdots\cdots & 0 \\ \vdots & \ddots \quad \ddots & \ddots & \cdots & \cdots & \cdots\cdots & \vdots \\ 0 & \cdots \quad 0 & 1 & 0 & \cdots & \cdots\cdots & 0 \\ -\dfrac{R'_{\ell,1}(a)}{R'_{\ell,\pi}(a)} & \cdots\cdots & -\dfrac{R'_{\ell,\pi-1}(a)}{R'_{\ell,\pi}(a)} & z-a & -\dfrac{R'_{\ell,\pi+1}(a)}{R'_{\ell,\pi}(a)} & \cdots\cdots & -\dfrac{R'_{\ell,m}(a)}{R'_{\ell,\pi}(a)} \\ 0 & \cdots\cdots & \cdots & 0 & 1 & 0 \cdots & 0 \\ \vdots & \cdots\cdots & \cdots & \cdots & \ddots & \ddots \quad \ddots & \vdots \\ 0 & \cdots\cdots & \cdots & \cdots & \cdots & 0 \quad 1 & 0 \\ 0 & \cdots\cdots & \cdots & \cdots & \cdots & \cdots \quad 0 & 1 \end{pmatrix}, \tag{7}$$

then it is straightforward to show that the matrix $P := P' \cdot U$ has order ω and its determinant takes the value $(z-a) \cdot \det P' = (z-a) \cdot \chi(\omega', G)$. Consequently, $\det P = \chi(\omega, G) = (z-a) \cdot \chi(\omega', G)$ which is a divisor of Ω. $\square$

Note that, when G is regular, then Lemma 2.7(a),(b) implies that

$$\chi(\omega, G) = \det \operatorname{diag} \omega. \tag{8}$$

For our recursive approach in later sections it will be useful that the order residual inherits properties of the original matrix power series G.

Lemma 2.8. *Suppose that G is regular and let P have order (ω, G, R). Then P is an (ω, G)-basis iff R is regular. Furthermore, if τ is a divisor of all components of ω, then τ is also a divisor of P and $(1/\tau) \cdot P$ has order $(\omega/\tau, G, R)$.*

Proof. There exists a regular $G' \in \mathscr{F}[[z]]_{\mathbb{F}_0}$ with $G' \cdot G = I$, where I is the identity matrix of suitable size. Hence

$$P = G' \operatorname{diag} \omega \cdot R \in \mathscr{F}[z]$$

which together with (8) implies both assertions. □

Example 2.9. Let

$$G(z) = \begin{bmatrix} \frac{1}{2} + z^2 - z^4 & 1 + \sin(z^2)^4 & \dfrac{1}{\sqrt{1+z^2}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and $\omega = (z^8, 1, 1)$. Then one basis for $\mathscr{M}(\omega, G)$ is given by

$$P(z) = \begin{bmatrix} z^4 + \frac{11}{2}z^2 & -\frac{10}{19}z^2 + \frac{2}{19} & \frac{9}{19}z^2 - \frac{11}{76} \\ -\frac{59}{4}z^2 & z^2 - \frac{33}{19} & -\frac{5}{4}z^2 + \frac{59}{152} \\ 12z^2 & \frac{32}{19} & z^2 - \frac{6}{19} \end{bmatrix},$$

with $\det P = z^8$. In this case the first 4 terms of the order residual R of P are given by

$$R(z) = \begin{bmatrix} -\frac{19}{4} - \frac{367}{32}z^2 - \frac{189}{64}z^4 + O(z^6) & -\frac{97}{76} + \frac{89}{152}z^2 + O(z^4) & -\frac{13}{1216} - \frac{1093}{1216}z^2 + O(z^4) \\ -\frac{59}{4}z^2 & -\frac{33}{19} + z^2 & \frac{59}{152} - \frac{5}{4}z^2 \\ 12z^2 & \frac{32}{19} & -\frac{6}{19} + z^2 \end{bmatrix}.$$

Example 2.10. Let (m,n) be a normal point in the Padé table of a power series $A(z)$, that is, a point where the coefficient Hankel matrix of the corresponding linear system is nonsingular. A normal point can also be described as a nonzero entry (m,n) in the C-table for $A(z)$ [2, p.23]. Normal points for interpolation problems are discussed further in a later section. It is well known that this is equivalent to finding polynomials p, q, u, v satisfying

$$A(z)q(z) - p(z) = z^{m+n-1}r(z) \quad \text{with } r(0) \neq 0,$$

$$A(z)v(z) - u(z) = z^{m+n+1}w(z) \quad \text{with } v(0) \neq 0,$$

and with degrees bounded by $m-1, n-1, m, n$, respectively. Let

$$P = \begin{bmatrix} zp & u \\ zq & v \end{bmatrix}.$$

Then again it is well known that $\det P = z \cdot (p \cdot v - q \cdot u) = z^{m+n} \cdot r(0) \cdot v(0)$. Thus P is an order basis for order $((z^{m+n}, 1), G)$ where G is as in Example 2.4 (over the scalars).

Notice also that, if $w_0 = \cdots = w_{k-1} = 0$, then for any $-2 \leqslant r \leqslant k$ the matrix polynomial

$$P_{(r)} = \begin{bmatrix} z^{2+r}p & u \\ z^{2+r}q & v \end{bmatrix}$$

satisfies $\det \boldsymbol{P}_{(r)} = z^{2+r}(pv - qu) = z^{m+n+r+1}r(0)v(0)$ so that $\boldsymbol{P}_{(r)}$ is an order basis for order $((z^{m+n+r+1}, 1), \boldsymbol{G})$.

When $r = -1$ or 0 we get the so-called Padé system from [17]. Similar Padé-like systems also exist in the context of matrix Padé approximation [27], Hermite–Padé and matrix Hermite–Padé approximation [25] and matrix simultaneous Padé approximation [14, 26].

2.3. *Transfer matrices*

We are interested in the recursive computation of bases for $\mathcal{M}(\boldsymbol{\omega}, \boldsymbol{G})$ in terms of bases of "lower" order. As such it is important to determine the possible choices for transferring a basis to one of higher order. This was initially considered by Mahler [29] in his study of M-Padé approximants (types I and II).

Definition 2.11 (*Transfer matrices*). Let $\boldsymbol{\omega}^{(1)}, \boldsymbol{\omega}^{(2)}$ be order vectors. We say that $\boldsymbol{\omega}^{(2)} = (\omega_1^{(2)}, \ldots, \omega_m^{(2)})$ contains $\boldsymbol{\omega}^{(1)} = (\omega_1^{(1)}, \ldots, \omega_m^{(1)})$ if $\omega_j^{(1)}$ is a divisor of $\omega_j^{(2)}$, $j = 1, \ldots, m$, or, equivalently, if $\boldsymbol{\omega}^{(1,2)}$ being defined by $\operatorname{diag} \boldsymbol{\omega}_j^{(1,2)} = (\operatorname{diag} \boldsymbol{\omega}_j^{(1)})^{-1} \cdot \operatorname{diag} \boldsymbol{\omega}_j^{(2)}$ is an order vector. Let $\boldsymbol{P}^{(1)} \in \mathscr{F}[z]$ be an $(\boldsymbol{\omega}^{(1)}, \boldsymbol{G})$-basis. Then the matrix $\boldsymbol{P}^{(1,2)}$ is called an $(\boldsymbol{\omega}^{(1)}, \boldsymbol{\omega}^{(2)})$-*transfer matrix* if $\boldsymbol{P}^{(2)} := \boldsymbol{P}^{(1)} \cdot \boldsymbol{P}^{(1,2)} \in \mathscr{F}[z]$ is an $(\boldsymbol{\omega}^{(2)}, \boldsymbol{G})$-basis.

Notice that since $\mathcal{M}(\boldsymbol{\omega}^{(2)}, \boldsymbol{G}) \subset \mathcal{M}(\boldsymbol{\omega}^{(1)}, \boldsymbol{G})$, Lemma 2.5(b) implies that each column of an $(\boldsymbol{\omega}^{(2)}, \boldsymbol{G})$-basis can be expressed as a polynomial combination of the columns of $\boldsymbol{P}^{(1)}$. Thus any transfer matrix is a polynomial matrix.

Theorem 2.12. *With the notations of Definition 2.11, let $\boldsymbol{R}^{(1)}$ denote the $(\boldsymbol{\omega}^{(1)}, \boldsymbol{G})$ order residual of $\boldsymbol{P}^{(1)}$. Then $\boldsymbol{P}^{(1,2)} = (\boldsymbol{P}^{(1)})^{-1} \cdot \boldsymbol{P}^{(2)}$ is an $(\boldsymbol{\omega}^{(1,2)}, \boldsymbol{R}^{(1)})$-basis if and only if $\boldsymbol{P}^{(2)} = \boldsymbol{P}^{(1)} \cdot \boldsymbol{P}^{(1,2)}$ is an $(\boldsymbol{\omega}^{(2)}, \boldsymbol{G})$-basis, both having the same order residual.*

Proof. Notice that the columns of $\boldsymbol{P}^{(2)} = \boldsymbol{P}^{(1)} \cdot \boldsymbol{P}^{(1,2)}$ are elements of $\mathcal{M}(\boldsymbol{\omega}^{(2)}, \boldsymbol{G})$ if and only if the columns of $\boldsymbol{P}^{(1,2)} = (\boldsymbol{P}^{(1)})^{-1} \cdot \boldsymbol{P}^{(2)}$ are elements of $\mathcal{M}(\boldsymbol{\omega}^{(1,2)}, \boldsymbol{R}^{(1)})$. Therefore from (4) we obtain

$$\chi(\boldsymbol{\omega}^{(2)}, \boldsymbol{G}) = \chi(\boldsymbol{\omega}^{(1)}, \boldsymbol{G}) \cdot \chi(\boldsymbol{\omega}^{(1,2)}, \boldsymbol{R}^{(1)}) \tag{9}$$

implying Theorem 2.12. $\square$

Example 2.13. Let $\boldsymbol{G}$ and $\boldsymbol{\omega}^{(1)}$ be as in Example 2.9. Then with $\boldsymbol{\omega}^{(1,2)} = (z^4, 1, 1)$ one basis for $\mathcal{M}(\boldsymbol{\omega}^{(1,2)}, \boldsymbol{R}^{(1)})$ is given by

$$\boldsymbol{P}^{(1,2)}(z) = \begin{bmatrix} z^2 & \frac{1208}{13\,775} & -\frac{9409}{34\,800} \\ -\frac{1405}{388}z^2 & z^2 - \frac{44}{145} & \frac{19}{32}z^2 + \frac{27\,257}{27\,840} \\ -12z^2 & -\frac{1984}{725} & \frac{9409}{2900} \end{bmatrix}.$$

Multiplying $\boldsymbol{P}^{(1)} = \boldsymbol{P}$ from Example 2.9 with $\boldsymbol{P}^{(1,2)}$ gives

$$\boldsymbol{P}^{(2)}(z) = \begin{bmatrix} z^6 + \frac{167}{97}z^4 + \frac{263}{194}z^2 & -\frac{318}{725}z^4 - \frac{398}{725}z^2 + \frac{264}{725} & -\frac{5071}{8700}z^4 - \frac{1753}{4350}z^2 - \frac{25\,511}{69\,600} \\ -\frac{327}{97}z^4 + \frac{633}{388}z^2 & z^4 + \frac{63}{725}z^2 - \frac{388}{725} & \frac{19}{32}z^4 - \frac{2779}{23\,200}z^2 - \frac{20\,467}{46\,400} \\ -\frac{224}{97}z^2 & \frac{256}{725} & z^2 + \frac{1358}{2175} \end{bmatrix}.$$

This gives a basis for $\mathcal{M}(\boldsymbol{\omega}^{(2)}, \boldsymbol{G})$ (where $\boldsymbol{\omega}^{(2)} = (z^{12}, 1, 1)$) with $\det \boldsymbol{P}^{(2)} = z^{12}$. By Theorem 2.12 all bases of $\mathcal{M}(\boldsymbol{\omega}^{(2)}, \boldsymbol{G})$ are determined in such a way. □

3. Interpolation at infinity

As of now we have only considered the order conditions of rational interpolation problems. The degree constraints that are common to such problems (e.g., $\boldsymbol{n}$-reduced bases [36]) have been ignored. In our case we deal with this problem by considering a more general concept using the notion of interpolation at infinity.

As an example illustrating our approach, let $\boldsymbol{n} = (n_1, \ldots, n_m)$ and consider the Hermite–Padé approximation problem of type $\boldsymbol{n}$ as discussed initially in Example 2.3. By multiplying $\boldsymbol{Q}$ on the left by $\boldsymbol{H} = \mathrm{diag}(z^{-n_1}, \ldots, z^{-n_m})$ the condition that the ith component of $\boldsymbol{Q}$ has degree at most n_i for all i is equivalent to

$$\boldsymbol{H} \cdot \boldsymbol{Q} = \mathcal{O}(z^0)_{z \to \infty}, \tag{10}$$

i.e., an interpolation condition at infinity. Similarly, in the case of right-handed matrix Padé forms of type (p, q) given in Example 2.4, specifying that the rows of $\boldsymbol{U}$ and $\boldsymbol{V}$ have degrees at most p and q, respectively, is the same as looking for those $\boldsymbol{Q}$ in $\mathcal{M}(\boldsymbol{\omega}, \boldsymbol{G})$ satisfying Eq. (10) for $\boldsymbol{H} = \mathrm{diag}(z^{-p}, \ldots, z^{-p}, z^{-q}, \ldots, z^{-q})$. From previous work on matrix-like Padé problems [8, 9] we know that it is useful to also have information on the differences between the degree constraints and the degrees of each column. This motivates the concept of an $\boldsymbol{H}$-degree of a matrix of polynomials.

For the remainder of this paper we will use the following notations: for an integer vector $\boldsymbol{d} = (d_1, \ldots, d_s)$, and a scalar $\alpha \in \mathbb{F}$, let $(z - \alpha)^{\boldsymbol{d}}$ denote the diagonal matrix-valued function $\mathrm{diag}((z - \alpha)^{d_1}, \ldots, (z - \alpha)^{d_s})$. Also, define the "norm" $|\boldsymbol{d}| = d_1 + \cdots + d_s$ and set $\boldsymbol{e} := (1, \ldots, 1)$. We will also assume that we have $\boldsymbol{H} \in \mathscr{F}[[z]]_-$, a right-truncated matrix Laurent power series satisfying in addition $\det \boldsymbol{H} \neq 0$.

Definition 3.1. $\boldsymbol{P} \in \mathbb{F}^{m \times s}[z]$ is said to have $\boldsymbol{H}$-degree $\boldsymbol{d}$, if

$$\boldsymbol{H} \cdot \boldsymbol{P} = \boldsymbol{S} \cdot z^{\boldsymbol{d}}, \tag{11}$$

with $\boldsymbol{S} \in \mathbb{F}^{m \times s}[[z]]_\infty$ (called the *degree residual*), and $\boldsymbol{S}(\infty)$ containing no zero columns (if $\boldsymbol{P}$ contains a zero column, then the corresponding quantity d_j is defined to be $+\infty$). It is called $\boldsymbol{H}$-reduced if the matrix $\boldsymbol{S}(\infty)$ in (11) has maximal rank.

Definition 3.2. $\boldsymbol{P} \in \mathscr{F}[z]$ will be called an $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis if it is both $\boldsymbol{H}$-reduced and an $(\boldsymbol{\omega}, \boldsymbol{G})$-basis.

For example, if $H = z^{(-2,-2,-2)}$, then the H-degree of the matrix polynomial P of Example 2.9 is $(2,0,0)$. In addition, P is H-reduced and hence is an (ω, G, H)-basis (where ω and G also come from Example 2.9). Similarly, if $H = z^{(-m,-n)}$ in Example 2.10 then, using well-known properties of Padé approximants at normal points, both P and $P_{(r)}$ are H-reduced bases with H-degrees $(0,0)$ and $(r+1,0)$, respectively.

Note that, similar to the order residual, the components of the H-degree vector of P are the H-degrees of the columns of P, that is, if $\{P_i\}$ represent the columns of P then

$$H\text{-deg}\, P = (H\text{-deg}\, P_1, \ldots, H\text{-deg}\, P_m).$$

The concept of the H-degree is known for many special cases.

Example 3.3 (*τ-degree, n-defect*). Let $H = z^{\tau}$ for an integer vector τ and let $P = (P_1, \ldots, P_m)$ be a column vector. Then the H-degree coincides with the τ-degree ([20, 34–36]), i.e.,

$$H\text{-deg}\, P = \tau\text{-deg}\, P = \max_j\{\deg P_j + \tau_j\}.$$

Furthermore, a polynomial matrix P is H-reduced iff it is τ-reduced.

Similarly, if $H = z^{-(n+e)}$ for an integer vector n then

$$H\text{-deg}\, P = -\text{dct}_n(P) =: -\min_j\{n_j + 1 - \deg P_j\}$$

gives the H-degree in terms of the n-defect ([3–5, 8, 9]) of P.

From Example 3.3 we may also conclude that the I-degree (or z^0-degree) coincides with the classical column-degree of a polynomial matrix.

Example 3.4 (*Vector biorthogonal polynomials*). Denote by $I \subset \mathbb{R}$ a compact set and let $W: I \to \mathbb{R}^{r\times s}$ be continuous. For given vectors m, n of nonnegative integers, the polynomial $Q^R \in \mathbb{R}^{s\times 1}[z]$ is called (m,n)-right orthogonal if $\deg Q^R \leqslant n$ (rowwise), and

$$\int Q^L(x) \cdot W(x) \cdot Q^R(x)\, dx = 0$$

for all $Q^L \in \mathbb{R}^{1\times r}[z]$ satisfying $\deg Q^L \leqslant m$ (columnwise). Similarly, $Q^L \in \mathbb{R}^{1\times r}[z]$ is called (m,n)-left orthogonal if $\deg Q^L \leqslant m$ and the above orthogonality relation holds for all $Q^R \in \mathbb{R}^{s\times 1}[z]$ with $\deg Q^R \leqslant n$. For particular so-called *perfect multiindices* m, n, vector biorthogonal polynomials have been successfully applied in the spectral theory of difference operators of order $(r + s + 1)$ (see [6,32]). It is not difficult to check that there is at least one (m,n)-right orthogonal polynomial if $|m| - r = |n| - s + 1$. Denote the corresponding matrix-valued symbol by

$$F(z) = \int \frac{W(x)}{z-x}\, dx \in R^{r\times s}[z].$$

Then the above conditions on an (m,n)-right orthogonal polynomial may be rewritten in terms of H-degrees as z^{-n}-deg $Q^R \leqslant 0$ and z^{m+2e}-deg$(F(z) \cdot Q^R(z) - P^R(z)) \leqslant 0$ with some $P^R \in \mathbb{R}^{r\times 1}[z]$.

Equivalently, we have

$$\boldsymbol{H}\text{-deg}\begin{pmatrix}\boldsymbol{P}^{\mathrm{R}}\\ \boldsymbol{Q}^{\mathrm{R}}\end{pmatrix} \leqslant 0 \quad \text{with } \boldsymbol{H} := z^{(m+2e_r-\boldsymbol{n})} \cdot \begin{pmatrix}-\boldsymbol{I}_r & \boldsymbol{F}\\ \boldsymbol{0}_{s\times r} & \boldsymbol{I}_s\end{pmatrix}.$$

Thus, vector right orthogonal polynomials are solutions of an extended M-Padé problem with interpolation only at infinity. Also, we have an integral representation for the upper part of the residual at infinity,

$$(-\boldsymbol{I}, \boldsymbol{F}(z)) \cdot \begin{pmatrix}\boldsymbol{P}^{\mathrm{R}}(z)\\ \boldsymbol{Q}^{\mathrm{R}}(z)\end{pmatrix} = \int \frac{W(x)\cdot \boldsymbol{Q}^{\mathrm{R}}(x)}{z-x}\,\mathrm{d}x,$$

as a function of the second kind.

For $\boldsymbol{P} \in \mathscr{F}[z]$ we can relate the $\boldsymbol{H}$-degree of $\boldsymbol{P}$ to the degree of its determinant. For order bases we can be even more precise.

Lemma 3.5. *Let* $\eta(\boldsymbol{H}) = \deg\det \boldsymbol{H}$, *i.e*, $\eta(\boldsymbol{H}) := \min\{k : \lim_{z\to\infty} z^{-k}\cdot\det \boldsymbol{H}(z)$ *is finite*$\}$. *Then for* $\boldsymbol{H}\text{-deg}\,\boldsymbol{P} = \boldsymbol{d}$ *we have*

(a) $|\boldsymbol{d}| \geqslant \deg\det \boldsymbol{P} + \eta(\boldsymbol{H})$,

(b) $|\boldsymbol{d}| = \deg\det \boldsymbol{P} + \eta(\boldsymbol{H})$ *iff* $\boldsymbol{P}$ *is* $\boldsymbol{H}$*-reduced,*

(c) *if* $\boldsymbol{P}$ *is an* $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$*-basis, then* $|\boldsymbol{d}| = \deg\chi(\boldsymbol{\omega}, \boldsymbol{G}) + \eta(\boldsymbol{H})$.

Proof. Lemma 3.5 follows directly by taking determinants in Eq. (11). □

As a direct consequence of Lemma 3.5 and (4) we see that there is no $\boldsymbol{Q} \in \mathscr{F}[z]$ with $\det \boldsymbol{Q} \neq 0$ having order $\boldsymbol{\omega}$ such that the norm of its $\boldsymbol{H}$-degree is smaller than $\deg\chi(\boldsymbol{\omega}, \boldsymbol{G}) + \eta(\boldsymbol{H})$. A similar minimality property has been observed in [20, Theorem 2].

It is a straightforward process to transform an $(\boldsymbol{\omega}, \boldsymbol{G})$-basis to one that is $\boldsymbol{H}$-reduced. To see this, let $\boldsymbol{P}_0 \in \mathscr{F}[z]$ with $\det \boldsymbol{P}_0 \neq 0$, and furthermore let $\boldsymbol{d}_0 = (d_{0,1}, \ldots, d_{0,m})$ and $B_0 := \boldsymbol{S}(\infty)$ be defined as in (11). If B_0 is regular, then $\boldsymbol{P}_0$ is already $\boldsymbol{H}$-reduced. Otherwise there exists a nontrivial column vector $b_0 = (b_{0,1}, \ldots, b_{0,m})^{\mathrm{T}}$ with $B_0 \cdot b_0 = 0$. Select k with $b_{0,k} \neq 0$ and $d_{0,k}$ as large as possible and define the unimodular polynomial matrix U_0 by

$$z^{\boldsymbol{d}_0}\cdot U_0 \cdot z^{-\boldsymbol{d}_0} := \begin{pmatrix}
1 & 0 & \cdots & 0 & b_{0,1} & 0 & \cdots & \cdots & 0\\
0 & 1 & \ddots & \vdots & \vdots & \vdots & & & \vdots\\
\vdots & \ddots & \ddots & 0 & \vdots & \vdots & & & \vdots\\
\vdots & & \ddots & 1 & b_{0,k-1} & \vdots & & & \vdots\\
\vdots & & & 0 & b_{0,k} & 0 & & & \vdots\\
\vdots & & & \vdots & b_{0,k+1} & 1 & \ddots & & \vdots\\
\vdots & & & \vdots & \vdots & 0 & \ddots & \ddots & \vdots\\
\vdots & & & \vdots & \vdots & \vdots & \ddots & 1 & 0\\
0 & \cdots & \cdots & 0 & b_{0,m} & 0 & \cdots & 0 & 1
\end{pmatrix}. \tag{12}$$

By multiplying Eq. (11) on the right, the resulting matrix $\boldsymbol{P}_1 := \boldsymbol{P}_0 \cdot \boldsymbol{U}_0$ satisfies $\deg\det \boldsymbol{P}_1 = \deg\det \boldsymbol{P}_0$ having an $\boldsymbol{H}$-degree $\boldsymbol{d}_1$ with $|\boldsymbol{d}_1| \leqslant |\boldsymbol{d}_0| - 1$ (since $d_{1,j} = d_{0,j}$ for $j \neq k$ and $d_{1,k} \leqslant d_{0,k} - 1$). Iterating this process yields the desired $\boldsymbol{H}$-reduced counterpart. Note that the process terminates since in each step we decrease $|\boldsymbol{d}_j|$, a number which is bounded from below because of Lemma 3.5.

Corollary 3.6. *Given any $\boldsymbol{P} \in \mathscr{F}[z]$ with $\det \boldsymbol{P} \neq 0$ and $\boldsymbol{H}$-deg $\boldsymbol{P} = \boldsymbol{d}$, we can successively construct a unimodular matrix $\boldsymbol{U} \in \mathscr{F}[z]$ such that $\boldsymbol{P} \cdot \boldsymbol{U}$ is $\boldsymbol{H}$-reduced and $\boldsymbol{H}$-deg $\{\boldsymbol{P} \cdot \boldsymbol{U}\} =: \boldsymbol{d}'$ with $\boldsymbol{d}' \leqslant \boldsymbol{d}$ (componentwise). In particular, if $\boldsymbol{P}$ is an $(\boldsymbol{\omega}, \boldsymbol{G})$-basis, then $\boldsymbol{P} \cdot \boldsymbol{U}$ is an $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis.*

Example 3.7. Let

$$\boldsymbol{P} = \begin{bmatrix} 2+z & 2 \\ 1-z & 1 \end{bmatrix}, \qquad \boldsymbol{H} = \begin{bmatrix} z^{-3} & 0 \\ 0 & z^{-2}+z^{-3} \end{bmatrix},$$

$$\boldsymbol{H} \cdot \boldsymbol{P} = \left(\begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} z^{-1} + \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix} z^{-2} \right) \cdot \begin{bmatrix} z^{-1} & 0 \\ 0 & z^{-2} \end{bmatrix}$$

and note that $\det \boldsymbol{H} = z^{-5} + z^{-6}$, $\det \boldsymbol{P} = 3z$ so that $|\boldsymbol{d}| = -3 > \eta(\boldsymbol{H}) + \deg\det \boldsymbol{P} = -4$. Set

$$\boldsymbol{U} = \begin{bmatrix} z^1 & 0 \\ 0 & z^2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} z^{-1} & 0 \\ 0 & z^{-2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ z & 1 \end{bmatrix}.$$

Then

$$\boldsymbol{P}^* := \boldsymbol{P} \cdot \boldsymbol{U} = \begin{bmatrix} 2+3z & 2 \\ 1 & 1 \end{bmatrix},$$

$$\boldsymbol{H} \cdot \boldsymbol{P}^* = \left(\begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} z^{-1} \right) \cdot \begin{bmatrix} z^{-2} & 0 \\ 0 & z^{-2} \end{bmatrix},$$

showing that $\boldsymbol{P}^*$ is $\boldsymbol{H}$-reduced.

Using a similar construction gives

Corollary 3.8. *Given any $\boldsymbol{A} \in \mathscr{F}[[z]]_\infty$ with $\det \boldsymbol{A} \neq 0$, there exist a unimodular matrix $\boldsymbol{V} \in \mathscr{F}[z]$ and a vector $\boldsymbol{\delta}$ containing only nonpositive integers such that $\boldsymbol{A}(z) = \boldsymbol{A}_1(z) \cdot z^{\boldsymbol{\delta}} \cdot \boldsymbol{V}(1/z)$ with $\boldsymbol{A}_1 \in \mathscr{F}[[z]]_\infty$ being regular.*

Proof. The assertion of this corollary is similar to the *Smith–McMillan normal form* of $\boldsymbol{A}$. We will show it directly. With $\boldsymbol{P} = \boldsymbol{I}$ we have $\boldsymbol{A}$-deg $\boldsymbol{P} \leqslant \boldsymbol{0}$, and Corollary 3.6 leads to a unimodular matrix $\boldsymbol{U} \in \mathscr{F}[z]$ such that $\boldsymbol{A}(z) \cdot \boldsymbol{U}(z) \cdot z^{-\boldsymbol{d}} \in \mathscr{F}[[z]]_\infty$ is regular, where $\boldsymbol{d} \leqslant \boldsymbol{0}$ (componentwise). However, for the assertion we do not want $\boldsymbol{U}(z)$ but $\boldsymbol{U}(1/z)$ to be unimodular. Therefore we have to generalize the concept of $\boldsymbol{H}$-degree as well as the above construction. Introducing the set of *Laurent polynomials*

$$\mathscr{F}[z]_- := \{z^\kappa \cdot \boldsymbol{P} : \kappa \text{ an integer}, \ \boldsymbol{P} \in \mathscr{F}[z]\},$$

we see that Definition 3.2 naturally extends to these quantities. Also Lemma 3.5(a),(b) remains valid if we replace $\deg\det \boldsymbol{P}$ by $\eta(\boldsymbol{P})$ (these quantities coincide for $\boldsymbol{P} \in \mathscr{F}[z]$). Therefore we may also perform the (terminating) process described before Corollary 3.6 for any $\boldsymbol{P} \in \mathscr{F}[z]_-$ with $\det \boldsymbol{P} \neq 0$. Moreover, if in each step we select k with $d_{0,k}$ as small as possible (instead of being as large as possible), the resulting factor U_0 and therefore the factor $\boldsymbol{U}$ of Corollary 3.6 has the required properties that $\boldsymbol{U} \in \mathscr{F}[z]_-$, and $\boldsymbol{U}'$ being defined by $\boldsymbol{U}'(z) = \boldsymbol{U}(1/z)$ is unimodular. Taking $\boldsymbol{V} = \boldsymbol{U}'^{-1}$ yields the assertion. □

The following lemma gives some simple observations that help in understanding a little more the concept of the $\boldsymbol{H}$-degree. Of particular interest are properties for an $\boldsymbol{H}$ of the form $z^{\boldsymbol{q}} \cdot \boldsymbol{A} \cdot z^{\boldsymbol{p}}$ with $\boldsymbol{A} \in \mathscr{F}[[z]]_\infty$ being regular and integer vectors $\boldsymbol{p}$ and $\boldsymbol{q}$. This is a common form for $\boldsymbol{H}$ in applications.

Lemma 3.9. (a) *Let $\boldsymbol{A} \in \mathscr{F}[[z]]_\infty$ be regular. Furthermore, let k be an integer, and $\boldsymbol{P} \in \mathscr{F}[z]$. Then*

$$\boldsymbol{H}\text{-}\deg \boldsymbol{P} = (\boldsymbol{A} \cdot \boldsymbol{H})\text{-}\deg \boldsymbol{P} = (z^k \cdot \boldsymbol{H})\text{-}\deg \boldsymbol{P} - k \cdot \boldsymbol{e}.$$

(b) *Suppose $\boldsymbol{H} = z^{\boldsymbol{q}} \cdot \boldsymbol{A} \cdot z^{\boldsymbol{p}}$ with $\boldsymbol{A} \in \mathscr{F}[[z]]_\infty$ being regular and $q_{\min} := \min \boldsymbol{q}$. Furthermore, let $\boldsymbol{P} \in \mathscr{F}[z]$ be $\boldsymbol{H}$-reduced with $\boldsymbol{H}$-$\deg \boldsymbol{P} = \boldsymbol{d}$ and $z^{\boldsymbol{p}}$-$\deg \boldsymbol{P} \leqslant N \cdot \boldsymbol{e} + \boldsymbol{d}$ with minimal N. Then necessarily $N + q_{\min} = 0$.*

Proof. Assertion (a) is clear using the definition of the $\boldsymbol{H}$-degree. In order to show (b), notice that by (11) and by definition of N

$$z^{-N \cdot \boldsymbol{e} - \boldsymbol{q}} \cdot \boldsymbol{S} = \boldsymbol{A} \cdot z^{\boldsymbol{p}} \cdot \boldsymbol{P} \cdot z^{-N \cdot \boldsymbol{e} - \boldsymbol{d}} \in \mathscr{F}[[z]]_\infty.$$

Since $\boldsymbol{S}(\infty)$ is regular, $N \cdot \boldsymbol{e} + \boldsymbol{q}$ may only contain nonnegative components. With the additional assumption of regularity of $\boldsymbol{A}(\infty)$, we may even conclude that the above expression, evaluated at infinity, is not identical zero giving $N + q_{\min} = 0$. □

Note that, at least theoretically, Lemma 3.9 covers all $\boldsymbol{H} \in \mathscr{F}[[z]]_-$ since with help of Corollary 3.8 we may rewrite $\boldsymbol{H}$ as $\boldsymbol{H} = \boldsymbol{H}_0 \cdot z^k = \boldsymbol{A}_0 \cdot z^{\boldsymbol{\delta} + k \cdot \boldsymbol{e}} \cdot \boldsymbol{A}$ with $\boldsymbol{H}_0, \boldsymbol{A}, \boldsymbol{A}_0 \in \mathscr{F}[[z]]_\infty$ and $\boldsymbol{A}, \boldsymbol{A}_0$ being regular, and thus the $\boldsymbol{H}$-degree and the $z^{\boldsymbol{\delta} + k \cdot \boldsymbol{e}} \cdot \boldsymbol{A}$-degree coincide. Also, as a consequence of Lemma 3.9(b), an upper bound for the $\boldsymbol{H}$-degree implies an upper bound for the degrees of the entries of $\boldsymbol{P}$.

4. Properties of interpolation bases

4.1. Bounds for the degrees of the coefficients

Theorem 2.6 gives representations of interpolation problems as polynomial combinations of basis elements. Specifying degree constraints in such problems serves to limit the possibilities for the components α_i appearing in the representation (5). For example, a Hermite–Padé approximant $\boldsymbol{Q}$ of type $\boldsymbol{n} = (2,2,2)$ for $(\frac{1}{2} + z^2 - z^4, 1 + \sin(z^2)^4, \frac{1}{\sqrt{1+z^2}})$ has order $((z^8, 1, 1), \boldsymbol{G})$ where $\boldsymbol{G}$ is as in Example 2.3. Hence if $\boldsymbol{P}_1, \boldsymbol{P}_2$ and $\boldsymbol{P}_3$ denote the columns of $\boldsymbol{P}$ in Example 2.3, then there is a representation of the form

$$\boldsymbol{Q} = \alpha_1 \boldsymbol{P}_1 + \alpha_2 \boldsymbol{P}_2 + \alpha_3 \boldsymbol{P}_3$$

for polynomials $\alpha_1, \alpha_2, \alpha_3$. Because of the degree constraints on the components of $\boldsymbol{Q}$ it is easy to see that $\alpha_1 = 0$ while both α_2 and α_3 must be constants since here the degree constraints impose a bound on the difference of the degree of each α_i and the $\boldsymbol{n}$-defect of the corresponding $\boldsymbol{P}_i$ (see, e.g., [5, 8, 9]). This is also the case for rational interpolation problems where the degree constraints are replaced by interpolation conditions at infinity.

Theorem 4.1. *Let $\boldsymbol{P}$ be an $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis with $\boldsymbol{H}$-deg $\boldsymbol{P} = \boldsymbol{d} = (d_1, \ldots, d_m)$. Moreover, let $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ denote the columns of $\boldsymbol{P}$. Then given any $\boldsymbol{Q}$ being element of the submodule $\mathcal{M}(\boldsymbol{\omega}, \boldsymbol{G})$ with $\delta = \boldsymbol{H}$-deg $\boldsymbol{Q}$ there exist unique polynomials α_j with*

$$\boldsymbol{Q} = \alpha_1 \boldsymbol{P}_1 + \cdots + \alpha_m \boldsymbol{P}_m, \quad \deg \alpha_j \leqslant \delta - d_j. \tag{13}$$

Proof. Taking the notations of Theorem 2.6, we know the determinantal representation (6). Further by assumption we have

$$|\boldsymbol{H}\text{-deg}\,(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_{j-1}, \boldsymbol{Q}, \boldsymbol{P}_{j+1}, \ldots, \boldsymbol{P}_m)| = |(d_1, \ldots, d_{j-1}, \delta, d_{j+1}, \ldots, d_m)| = |\boldsymbol{d}| - d_j + \delta.$$

From Lemma 3.5 we can conclude that

$$\begin{aligned} \deg \alpha_j &= \deg \det (\boldsymbol{P}_1, \ldots, \boldsymbol{P}_{j-1}, \boldsymbol{Q}, \boldsymbol{P}_{j+1}, \ldots, \boldsymbol{P}_m) - \deg \det (\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m) \\ &\leqslant (|\boldsymbol{d}| - d_j + \delta - \eta(\boldsymbol{H})) - (|\boldsymbol{d}| - \eta(\boldsymbol{H})) = \delta - d_j. \quad \square \end{aligned}$$

Note that property (13) together with the requirement $\boldsymbol{P}_j \in \mathcal{M}(\boldsymbol{\omega}, \boldsymbol{G})$, $j = 1, \ldots, m$, has been used in earlier papers (see, e.g., [9, Definition 3.2]) in order to define the so-called σ-bases $\{\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m\}$. In fact, it is not difficult to show that both approaches are equivalent.

Example 4.2. Theorem 4.1 provides a useful tool for characterizing the singular structure of various Padé-like tables. For example, suppose (m, n) is a normal point of the Padé table and that the order condition of the approximant of type (m, n) overshoots its order condition by k. Let (q_1, q_2) be a Padé approximant of type $(m + r, n + s)$ with $1 \leqslant r, s \leqslant k$.

From Example 2.10 we know that $\boldsymbol{P}_{(r)}$ is an $\boldsymbol{H}$-reduced basis for $((z^{m+n+r+1}, 1), \boldsymbol{G})$ where $\boldsymbol{H} = z^{(-m,-n)}$ and that $\boldsymbol{H}$-deg $\boldsymbol{P}_{(r)} = (r + 1, 1)$. Let $\boldsymbol{Q} = (q_1, q_2)^{\mathrm{T}}$. Then $\boldsymbol{Q}$ also has order $((z^{m+n+r+1}, 1), \boldsymbol{G})$ and hence can be written as

$$\boldsymbol{Q} = \alpha_1 \boldsymbol{P}_1 + \alpha_2 \boldsymbol{P}_2 \quad \text{with } \deg \alpha_1 \leqslant c - (r + 1) \text{ and } \deg \alpha_2 \leqslant c,$$

where $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are the columns of $\boldsymbol{P}_{(r)}$ and where $c = \boldsymbol{H}$-deg $\boldsymbol{Q}$. But with at least one of r or s at least one we have that $0 \leqslant c \leqslant r$.

Therefore $\alpha_1 = 0$ and $\boldsymbol{Q} = \alpha_2 \boldsymbol{P}_2$, hence $q_1/q_2 = u/v$ is unique. This is just the classical block structure of the singular Padé table.

As in [3, pp. 212–214], we can also show the following invariance property:

Theorem 4.3. *The $\boldsymbol{H}$-degree of two $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-bases coincides up to permutation of columns.*

Proof. Let $\boldsymbol{P}^{(i)}$ be an $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis with $\boldsymbol{H}$-deg $\boldsymbol{P}^{(i)} = \boldsymbol{d}^{(i)} = (d_1^{(i)}, \ldots, d_m^{(i)})$, for $i = 1, 2$. Moreover, let the columns of $\boldsymbol{P}^{(i)}$ be permuted such that $d_1^{(i)} \leqslant \cdots \leqslant d_m^{(i)}$. Let us show that $\boldsymbol{d}^{(1)} = \boldsymbol{d}^{(2)}$. Suppose in contrast that $d_j^{(1)} = d_j^{(2)}$, $j < s$, but $d_s^{(1)} \neq d_s^{(2)}$. Without any loss of generality we may assume that $d_s^{(1)} < d_s^{(2)}$ and therefore by the ordering we have

$$d_j^{(1)} - d_i^{(2)} \leqslant d_s^{(1)} - d_s^{(2)} < 0 \quad \text{for } 1 \leqslant j \leqslant s,\ s \leqslant i \leqslant m.$$

Since both $\boldsymbol{P}^{(1)}$ and $\boldsymbol{P}^{(2)}$ are $(\boldsymbol{\omega}, \boldsymbol{G})$-bases, Theorem 2.12 implies that there exists a nonsingular matrix $\boldsymbol{U} = (\alpha_{i,j}) \in \mathscr{F}[z]$ with $\boldsymbol{P}^{(1)} = \boldsymbol{P}^{(2)} \cdot \boldsymbol{U}$. Moreover, from Theorem 4.1 we can conclude that $\deg \alpha_{i,j} \leqslant d_j^{(1)} - d_i^{(2)}$ for all i, j. Hence by construction we get $\alpha_{i,j} = 0$ for $1 \leqslant j \leqslant s$, $s \leqslant i \leqslant m$, a contradiction to the fact that $\boldsymbol{U}$ is nonsingular. □

From the proof of Theorem 4.3 we can conclude that the $(\boldsymbol{\omega}, \boldsymbol{\omega})$-transfer matrix transforming a reduced basis into another one has to be block triangular (up to a permutation of rows and columns), where the blocks on the diagonal only contain elements from $\mathbb{F}$. Also, we are now prepared to introduce the concept of normality, something that is well established for particular cases.

Definition 4.4. The data $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$ will be called *normal* if all components of the $\boldsymbol{H}$-degree of a $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis coincide. More generally, we will speak of *weakly normal* data if the components of the $\boldsymbol{H}$-degree of a $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis differ at most by one.

Notice that in the case of normal data we implicitly require that $\deg \chi(\boldsymbol{\omega}, \boldsymbol{G}) + \eta(\boldsymbol{H})$ has to be a multiple of m. However, for most of the classical examples such as Padé systems the latter quantity equals zero. In fact, here one usually chooses a fixed regular $\boldsymbol{G}$, and $\boldsymbol{H} = z^{-\boldsymbol{n}}$ with some varying multiindex $\boldsymbol{n}$, and finally a particular order vector $\boldsymbol{\omega}$ corresponding to $\boldsymbol{n}$. Thus, there is a correspondence between normal points in some m-dimensional solution table and normal data $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$ (see e.g., Example 2.10).

As a consequence of Theorem 4.3 we have the following uniqueness result:

Corollary 4.5. *An $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis is unique up to multiplication on the right by a (arbitrary) nonsingular scalar matrix if and only if $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$ is normal.*

There are different well-established normalizations of bases corresponding to normal data: we may choose the residual at infinity satisfying $\boldsymbol{S}(\infty) = \boldsymbol{I}$, which includes the "monic" systems of polynomials introduced by Mahler [29] for the case of scalar Hermite–Padé and simultaneous Padé approximants and [26] for the matrix case. A "comonic" normalization, namely $\boldsymbol{R}(0) = \boldsymbol{I}$, is chosen for the Padé-type systems of [12, 14, 15, 26] (provided that $\det \boldsymbol{G}(0) \neq 0$).

4.2. Connection to classical interpolation problems

The rational interpolation problems that we study are often described in terms of order and interpolation at infinity as the linear set

$$\mathscr{M}(\boldsymbol{\omega}, \delta, \boldsymbol{G}, \boldsymbol{H}) := \{\boldsymbol{Q} \in \mathbb{F}^{m \times 1}[z] : \boldsymbol{Q} \text{ has order } \boldsymbol{\omega} \text{ and } \boldsymbol{H}\text{-deg}\, \boldsymbol{Q} \leqslant \delta\}. \tag{14}$$

Newton–Padé, simultaneous Padé, Hermite–Padé and multipoint Padé approximation problems can all be presented in this way. For example, with $\boldsymbol{G}$ as in Example 2.3 and $\boldsymbol{H}=z^{-\boldsymbol{n}}$ the set $\mathcal{M}((z^{|\boldsymbol{n}|+m-1}, 1,\ldots,1),0,\boldsymbol{G},\boldsymbol{H})$ describes the set of Hermite–Padé (M-Padé) approximants of type $\boldsymbol{n}$. Similarly, with $\boldsymbol{G}$ as in Example 2.4 and $\boldsymbol{H}=z^{-(p,\ldots,p,q,\ldots,q)}$, the set $\mathcal{M}((z^{p+q+1},\ldots,z^{p+q+1},1,\ldots,1),0,\boldsymbol{G},\boldsymbol{H})$ can be used to describe the columns of all matrix Padé forms of type (p,q).

Theorem 4.1 enables us to characterize all solutions in the space $\mathcal{M}(\omega,\delta,\boldsymbol{G},\boldsymbol{H})$. For example, we obtain as the dimension of the linear space (over $\mathbb{F}$) $\mathcal{M}(\omega,\delta,\boldsymbol{G},\boldsymbol{H})$ the quantity

$$\dim_{\mathbb{F}}\mathcal{M}(\omega,\delta,\boldsymbol{G},\boldsymbol{H})=\sum_{j=1}^{m}\max\{0,\delta+1-d_j\}. \tag{15}$$

Some further properties are summarized in the next

Corollary 4.6. *Let $s:=(\delta+1)\cdot m-\deg\chi(\omega,\boldsymbol{G})-\eta(\boldsymbol{H})$ be an integer between 1 and m. Then:*

(a) *$\mathcal{M}(\omega,\delta,\boldsymbol{G},\boldsymbol{H})$ has at least s solutions linearly independent over $\mathbb{F}$;*

(b) *$\mathcal{M}(\omega,\delta,\boldsymbol{G},\boldsymbol{H})$ has s solutions linearly independent over $\mathbb{F}[z]$ if and only if $\delta\cdot\boldsymbol{e}-\boldsymbol{d}$ contains s nonnegative components;*

(c) *a matrix of size $(m\times s)$ built up with s solutions from $\mathcal{M}(\omega,\delta,\boldsymbol{G},\boldsymbol{H})$, linearly independent over $\mathbb{F}$, is unique up to multiplication on the right by a nonsingular scalar matrix of size $s\times s$ if and only if the data $(\omega,\boldsymbol{G},\boldsymbol{H})$ is weakly normal. In this case, these s solutions are also linearly independent over $\mathbb{F}[z]$.*

Proof. In Section 3 we have already shown the existence of an $(\omega,\boldsymbol{G},\boldsymbol{H})$-basis $\boldsymbol{P}$ with $\boldsymbol{H}$-degree $\boldsymbol{d}$ satisfying $|\boldsymbol{d}|=\deg\chi(\omega,\boldsymbol{G})+\eta(\boldsymbol{H})$ and therefore $|\boldsymbol{d}|=(\delta+1)\cdot m-s$. Thus, part (a) follows from (15). Assertion (b) is an immediate consequence of Theorem 4.1. In order to show part (c), note that the data $(\omega,\boldsymbol{G},\boldsymbol{H})$ is weakly normal if and only if the vector $\delta\cdot\boldsymbol{e}-\boldsymbol{d}$ contains only the entries -1 and 0, namely exactly s components equal to 0. By construction, this is equivalent to saying that $\delta\cdot\boldsymbol{e}-\boldsymbol{d}$ contains exactly s nonnegative components, all being equal to 0. Taking into account Theorem 4.1 we get the equivalent characterization that all solutions of (14) are obtained by taking scalar linear combinations of s basis elements, showing the first part of (c). The second part now follows using assertion (b). □

Let us also mention that if $\boldsymbol{H}$ satisfies the conditions of Lemma 3.9(b) with $\boldsymbol{p}=(p_1,\ldots,p_m)$, $\boldsymbol{q}=(q_1,\ldots,q_m)$, then all $\boldsymbol{Q}=(Q_1,\ldots,Q_m)^{\mathrm{T}}$ in $\mathcal{M}(\omega,\delta,\boldsymbol{G},\boldsymbol{H})$ must satisfy the degree constraints

$$\max_j\{\deg Q_j+p_j\}\leqslant\delta-\min_j\{q_j\}. \tag{16}$$

4.3. Scalar multipoint Padé approximants

Suppose that $\mathbb{F}$ is the field of real numbers (the extension to the complex numbers is immediate). Let $\mathbb{F}_0$ contain a sequence of *knots* $z_0,z_1,\ldots,$ and define

$$\omega_{i,j}(z):=(z-z_i)\cdot(z-z_{i+1})\cdot\cdots\cdot(z-z_{j-1})$$

if $i < j$ while $\omega_{i,j}(z) := 1$ otherwise. Let $f_1, g_1 \in \mathbb{F}[[z]]_{\mathbb{F}_0}$, such that for each j one of the quantities $f_1(z_j)$, $g_1(z_j)$, is different from zero. Also, let $\hat{f}_1, \hat{g}_1 \in \mathbb{F}[[z]]_\infty$, not both vanishing at infinity. Then we may find $f_2, g_2 \in \mathbb{F}[[z]]_{\mathbb{F}_0}$ and $\hat{f}_2, \hat{g}_2 \in \mathbb{F}[[z]]_\infty$ such that with $m = 2$

$$G(z) := \begin{pmatrix} g_1(z) & f_1(z) \\ g_2(z) & f_2(z) \end{pmatrix} \in \mathscr{F}[[z]]_{\mathbb{F}_0} \quad \text{is regular,} \tag{17}$$

$$A(z) := \begin{pmatrix} \hat{g}_1(z) & \hat{f}_1(z) \\ \hat{g}_2(z) & \hat{f}_2(z) \end{pmatrix} \in \mathscr{F}[[z]]_\infty \quad \text{is regular.} \tag{18}$$

We define furthermore for integers i, μ, ν, with ν nonnegative and $i \geqslant -1$,

$$\begin{aligned} H(z) &:= z^{(\nu, \mu - i)} \cdot A(z) \cdot z^{(0,i)} \\ &= \begin{pmatrix} z^\nu \cdot \hat{g}_1(z) & z^{\nu+i} \cdot \hat{f}_1(z) \\ z^{\mu-i} \cdot \hat{g}_2(z) & z^\mu \cdot \hat{f}_2(z) \end{pmatrix} \in \mathscr{F}[[z]]_-. \end{aligned} \tag{19}$$

Depending on varying $[\mu, \nu]$ in the range

$$-\nu - 1 \leqslant \mu \leqslant \nu + i, \tag{20}$$

we look for a solution $(u, v)^T$ of (14) with parameters $(\boldsymbol{\omega}, \mu + \nu, \boldsymbol{G}, \boldsymbol{H})$ with $\boldsymbol{\omega} = (\omega_{0,\mu+\nu+1}, 1)$. This problem is related to the well-known multipoint Padé approximation problem (see for example [21–23, 30]).

Note that $\eta(\boldsymbol{H}) = \mu + \nu$, hence there is a $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis $\boldsymbol{P}$ such that $|\boldsymbol{H}\text{-deg}\,\boldsymbol{P}| = 2\mu + 2\nu + 1$. Denote the columns of $\boldsymbol{P}$ by $\boldsymbol{P}_1, \boldsymbol{P}_2$, and suppose without loss of generality that $\boldsymbol{H}\text{-deg}\,P_1 \leqslant \boldsymbol{H}\text{-deg}\,P_2$. Since $\boldsymbol{H}\text{-deg}\,P_1 + \boldsymbol{H}\text{-deg}\,P_2 = 2\mu + 2\nu + 1$, we may conclude that $\boldsymbol{H}\text{-deg}\,P_2 > \mu + \nu \geqslant \boldsymbol{H}\text{-deg}\,P_1$. Thus, by Theorem 4.1, any solution in $\mathscr{M}(\boldsymbol{\omega}, \mu + \nu, \boldsymbol{G}, \boldsymbol{H})$ differs from $\boldsymbol{P}_1$ by multiplication with a polynomial. This implies that the fraction u/v is unique giving a classical "block" structure to the corresponding table of rational approximants formed by this multipoint problem. Also, by Theorem 4.1, $\boldsymbol{P}_1$ is unique up to multiplication with a scalar. Thus, following [17], the components of $\boldsymbol{P}_1$, denoted in the sequel by $p_{[\mu,\nu]}, q_{[\mu,\nu]}$, form a "minimal solution" of the (linearized) multipoint Padé approximation problem.

Let us show that our problem coincides with the proper multipoint Padé problem as introduced in [21, 23]. Writing down more explicitly the three essential conditions given implicitly by (17) and (19) for an element $(u, v)^T$ of $\mathscr{M}(\boldsymbol{\omega}, \mu + \nu, \boldsymbol{G}, \boldsymbol{H})$ gives

$$g_1(z)u(z) + f_1(z)v(z) = r(z) \cdot \omega_{0,\mu+\nu+1} \quad \text{with a Newton series } r, \tag{21}$$

$$\hat{g}_1(z)u(z) + z^i \hat{f}_1(z)v(z) = z^\mu \cdot (c + o(1)_{z \to \infty}) \quad \text{with } c \in \mathbb{F}, \tag{22}$$

$$z^{-i}\hat{g}_2(z)u(z) + \hat{f}_2(z)v(z) = z^\nu \cdot (c' + o(1)_{z \to \infty}) \quad \text{with } c' \in \mathbb{F}. \tag{23}$$

The matrix $\boldsymbol{H}$ satisfies the conditions of Lemma 3.9(b) with $\boldsymbol{p} = (0, i)$ and $\boldsymbol{q} = (\nu, \mu - i)$. Since $\max\{\nu, \mu - i\} = \nu$ by (20), from (16) we obtain the degree constraints

$$\deg u \leqslant \nu + i, \qquad \deg v \leqslant \nu \tag{24}$$

and we may drop condition (23) since it will be always true. The remaining interpolation problem (21), (22), and (24) coincides with Gutknecht's proper multipoint Padé problem.

Finally, note that working with two interpolation conditions at infinity (instead of one as in [21, 23]) enables us also to include ordinary Newton–Padé approximation. In fact, in the special case, $\hat{g}_1(z)=\hat{f}_2(z)=1$ and $\hat{g}_2(z)=\hat{f}_1(z)=0$ conditions (22), (23) become the ordinary degree constraints $\deg u \leqslant \mu$, $\deg v \leqslant \nu$ (independent of i), and the pair $(p_{[\mu,\nu]}, q_{[\mu,\nu]})$ coincides with the minimal solution for Newton–Padé approximation as described in [17].

5. The general recurrence

5.1. Dividing a problem into two subproblems

A type of recursive algorithm for an arbitrary path algorithm for computing Newton–Padé approximants was given in [21]. This algorithm divides a Newton–Padé problem into two smaller problems — one a Newton–Padé computation and the other a multipoint Padé problem. In [23], Gutknecht extends this recurrence formula to multipoint Padé problems, here one breaks apart a large problem into two smaller problems of the *same* type. Note that these problems are both special cases of extended M-Padé approximation. In fact, as in [23] we obtain a simple recurrence relation being an immediate consequence of Theorem 2.12.

Theorem 5.1. *Let $\omega^{(1)}, \omega^{(2)}$ be order vectors, $\omega^{(2)}$ containing $\omega^{(1)}$, and let* $\operatorname{diag}\omega^{(2)} = \operatorname{diag}\omega^{(1)} \cdot \operatorname{diag}\omega^{(1,2)}$. *In addition, let $\mathbf{P}^{(1)} \in \mathscr{F}[z]$ be an $(\omega^{(1)}, \mathbf{G})$-basis with order residual $\mathbf{R}^{(1)}$ and $\mathbf{H}^{(2)} \in \mathscr{F}[[z]]_-$. Set $\mathbf{H}^{(1,2)} := \mathbf{H}^{(2)} \cdot \mathbf{P}^{(1)}$.*

If $\mathbf{P}^{(1,2)}$ is an $(\omega^{(1,2)}, \mathbf{R}^{(1)}, \mathbf{H}^{(1,2)})$-basis with order residual $\mathbf{R}^{(1,2)}$, degree residual $\mathbf{S}^{(1,2)}$, and $\mathbf{H}^{(1,2)}$-degree $\mathbf{d}^{(1,2)}$, then $\mathbf{P}^{(2)}=\mathbf{P}^{(1)} \cdot \mathbf{P}^{(1,2)}$ is an $(\omega^{(2)}, \mathbf{G}, \mathbf{H}^{(2)})$-basis with the same order residual $\mathbf{R}^{(2)}=\mathbf{R}^{(1,2)}$, degree residual $\mathbf{S}^{(2)} = \mathbf{S}^{(1,2)}$, and $\mathbf{H}^{(2)}$-degree $\mathbf{d}^{(2)} = \mathbf{d}^{(1,2)}$.

Conversely, if $\mathbf{P}^{(2)}$ is an $(\omega^{(2)}, \mathbf{G}, \mathbf{H}^{(2)})$-basis with order residual $\mathbf{R}^{(2)}$, degree residual $\mathbf{S}^{(2)}$, and $\mathbf{H}^{(2)}$-degree $\mathbf{d}^{(2)}$, then $\mathbf{P}^{(1,2)} = (\mathbf{P}^{(1)})^{-1} \cdot \mathbf{P}^{(2)}$ is an $(\omega^{(1,2)}, \mathbf{R}^{(1)}, \mathbf{H}^{(1,2)})$-basis with the same residual $\mathbf{R}^{(1,2)} = \mathbf{R}^{(2)}$, degree residual $\mathbf{S}^{(1,2)} = \mathbf{S}^{(2)}$, and $\mathbf{H}^{(1,2)}$-degree $\mathbf{d}^{(1,2)} = \mathbf{d}^{(2)}$.

Note that normalization (both monic and comonic) is preserved under our recurrence. Also, we see that $(\omega^{(2)}, \mathbf{G}, \mathbf{H}^{(2)})$ are (weakly) normal data if and only if $(\omega^{(1,2)}, \mathbf{R}^{(1)}, \mathbf{H}^{(1,2)})$ are (weakly) normal data.

Example 5.2. Let $\mathbf{G}$ and $\mathbf{P}^{(1)} = \mathbf{P}$ be as in Example 2.9. Then to compute a $\mathscr{M}((z^{16}, 1, 1), \mathbf{G}, \mathbf{H}^{(2)})$-base $\mathbf{P}^{(2)}$ where $\mathbf{H}^{(2)}=\operatorname{diag}(z^{-3}, z^{-4}, z^{-5})$ one can determine a basis for $\mathscr{M}((z^6, 1, 1), \mathbf{R}^{(1)}, \mathbf{H}^{(1,2)})$ where $\mathbf{R}^{(1)}$ represents the first 8 terms of the order residual series and

$$\mathbf{H}^{(1,2)}(z) = \begin{bmatrix} z + \frac{11}{2}z^{-1} & -\frac{10}{19}z^{-1} + \frac{2}{19}z^{-3} & \frac{9}{19}z^{-1} - \frac{11}{76}z^{-3} \\ -\frac{59}{4}z^{-2} & z^{-2} - \frac{33}{19}z^{-4} & -\frac{5}{4}z^{-2} + \frac{59}{152}z^{-4} \\ 12z^{-3} & \frac{32}{19}z^{-5} & z^{-3} - \frac{6}{19}z^{-5} \end{bmatrix}.$$

In this case a basis for $\mathcal{M}((z^6,1,1),\boldsymbol{R}^{(1)},\boldsymbol{H}^{(1,2)})$ is given by

$$\boldsymbol{P}^{(1,2)}(z)=\begin{bmatrix} \frac{10\,278\,436}{27\,153\,755} & \frac{2\,361\,878}{10\,759\,035} & \frac{7}{38}z^2 \\ \frac{10\,208\,327}{4\,001\,606}z^2-\frac{1\,901\,158}{1\,429\,145} & -\frac{857\,615}{792\,771}z^2-\frac{908\,503}{1\,132\,530} & \frac{5}{4}z^4-\frac{593}{916}z^2 \\ \frac{41\,481\,712}{10\,004\,015}z^2-\frac{13\,388\,256}{1\,429\,145} & -\frac{6\,601\,544}{3\,963\,855}z^2-\frac{333\,656}{188\,755} & z^4-\frac{1044}{229}z^2 \end{bmatrix}$$

with $\boldsymbol{H}^{(1,2)}$-degree $(1,0,1)$. A basis for $\mathcal{M}((z^{16},1,1),\boldsymbol{G},\boldsymbol{H}^{(2)})$ having $\boldsymbol{H}^{(2)}$-degree $(1,0,1)$ is then given as $\boldsymbol{P}^{(1)}\cdot\boldsymbol{P}^{(1,2)}$.

Note that in this case one computes a basis for the Hermite–Padé forms of type $(2,2,2)$ and recursively computes a basis for all Hermite–Padé forms of type $(3,4,5)$. In general, the recursion of Theorem 5.1 provides an arbitrary path algorithm for computing reduced bases for the Hermite–Padé approximation problem.

Example 5.3. Let $\boldsymbol{n}^{(1)}=(n_1,\dots,n_m)$ be a multiindex and suppose that we have a $\boldsymbol{H}^{(1)}=z^{-\boldsymbol{n}^{(1)}}$-reduced basis $\boldsymbol{P}^{(1)}$ for the Hermite–Padé problem of order $(\boldsymbol{\omega}^{(1)},\boldsymbol{G})$ where $\boldsymbol{\omega}^{(1)}=(z^{|\boldsymbol{n}^{(1)}|},1,\dots,1)$ and $\boldsymbol{G}$ is as in Example 2.3. Further we assume that $\boldsymbol{H}^{(1)}\text{-deg}\,\boldsymbol{P}^{(1)}=0$ so that we are at a normal point of the Hermite–Padé table. Let $\boldsymbol{n}^{(2)}=(n_1+s,\dots,n_m+s)=\boldsymbol{n}^{(1)}+s\cdot\boldsymbol{e}$. We are interested in computing a $\boldsymbol{H}^{(2)}=z^{-\boldsymbol{n}^{(2)}}$-reduced basis $\boldsymbol{P}^{(2)}$ for the Hermite–Padé problem again having $\boldsymbol{H}^{(2)}$-deg $\boldsymbol{P}^{(2)}=0$. Thus we are interested in computing bases along normal points of an offdiagonal path.

Note that $\boldsymbol{P}^{(1)}$ is also $\boldsymbol{H}^{(2)}$-reduced with $\boldsymbol{H}^{(2)}$-deg $\boldsymbol{P}^{(1)}=-s\cdot\boldsymbol{e}$. Theorem 5.1 together with Lemma 3.9(a) imply that $\boldsymbol{P}^{(2)}$ can be determined by computing an $z^{-s\cdot\boldsymbol{e}}$-reduced basis $\boldsymbol{P}^{(1,2)}$ of order $((z^{s\cdot m},1,\dots,1),\boldsymbol{R}^{(1)})$ for the residual $\boldsymbol{R}^{(1)}$ of $\boldsymbol{P}^{(1)}$ with $z^{-s\cdot\boldsymbol{e}}\text{-deg}\,\boldsymbol{P}^{(1,2)}=0$. Thus the data for computing the intermediate basis $\boldsymbol{P}^{(1,2)}$ will be also normal. Such a recursion is used in the hybrid algorithms in [15, 26]. Here, the intermediate problems are solved using Gaussian elimination of the associated linear systems. Similar hybrid algorithms can also be found in the matrix Padé problem [27] and matrix simultaneous Padé problem [16, 26].

Example 5.4. The recursion for the Hermite–Padé problem is a bit more involved in the case where we are taking an arbitrary, rather than offdiagonal path. For example, having $\boldsymbol{n}^{(2)}=(n_1+s,n_2,\dots,n_m)$ gives a recursion along a row path in the Hermite–Padé table (such a path corresponds to a Toeplitz-like solver rather than the Hankel-like solver of the previous paragraph).

In the general case of moving from one normal point to the next, let $\boldsymbol{n}^{(1)}=(n_1,\dots,n_m)$ be an integer multiindex and again assume that we have a $\boldsymbol{H}^{(1)}=z^{-\boldsymbol{n}^{(1)}}$-reduced basis $\boldsymbol{P}^{(1)}$ as in the previous example. We are now interested in computing a $\boldsymbol{H}^{(2)}=z^{-\boldsymbol{n}^{(2)}}$-reduced basis $\boldsymbol{P}^{(2)}$ of order $(\boldsymbol{\omega}^{(2)},\boldsymbol{G})$ with $\boldsymbol{\omega}^{(2)}=(z^{|\boldsymbol{n}^{(2)}|},1,\dots,1)$ where $\boldsymbol{n}^{(2)}=\boldsymbol{n}^{(1)}+\boldsymbol{s}=(n_1+s_1,\dots,n_m+s_m)$ with varying s_i. In this case our initial basis $\boldsymbol{P}^{(1)}$ is not necessarily $\boldsymbol{H}^{(2)}$-reduced. Therefore we may first determine a unimodular matrix $\boldsymbol{P}^{(1,3)}$ such that $\boldsymbol{P}^{(3)}=\boldsymbol{P}^{(1)}\cdot\boldsymbol{P}^{(1,3)}$ is $\boldsymbol{H}^{(2)}$ reduced and has order $(\boldsymbol{\omega}^{(1)},\boldsymbol{G})$. Suppose $\boldsymbol{H}^{(2)}\text{-deg}\,\boldsymbol{P}^{(3)}=-\boldsymbol{d}$. Then, since $\boldsymbol{H}^{(2)}\cdot\boldsymbol{P}^{(3)}=\boldsymbol{S}\cdot z^{-\boldsymbol{d}}$ with $\boldsymbol{S}(\infty)$ nonsingular, the recursion requires us to compute a $z^{-\boldsymbol{d}}$-reduced basis $\boldsymbol{P}^{(3,2)}$ of order $((z^{|\boldsymbol{s}|},1,\dots,1),\boldsymbol{R}^{(3)})$ with $\boldsymbol{R}^{(3)}=\boldsymbol{R}^{(1)}\cdot\boldsymbol{P}^{(1,3)}$ and having $z^{-\boldsymbol{d}}$-degree 0. Thus again we solve our recursion by solving a Hermite–Padé problem for $\boldsymbol{R}^{(3)}$ of type $\boldsymbol{d}$ at a normal point. Note that $|\boldsymbol{d}|=|\boldsymbol{s}|$ so that the intermediate problem can be solved via a linear system of equations. This allows for hybrid algorithms along arbitrary paths.

Example 5.5. The Hermite–Padé algorithm of [9] follows a different approach than that used by the previous two examples. Let $\boldsymbol{n}=(n_1,\ldots,n_m)$ and $\boldsymbol{H}=z^{-\boldsymbol{n}}$. Then this algorithm iteratively computes a sequence of bases $\boldsymbol{I}=\boldsymbol{P}^{(0)},\boldsymbol{P}^{(1)},\ldots,\boldsymbol{P}^{(N)}$ along an offdiagonal of the Hermite–Padé table, each of order $(\boldsymbol{\omega}^{(i)},\boldsymbol{G})$, $\boldsymbol{\omega}^{(i)}=(z^i,1,\ldots,1)$. Here i varies from 0 to $N=|\boldsymbol{n}|+m-1$. In this case the $\boldsymbol{H}$ remains fixed throughout with each $\boldsymbol{P}^{(i)}$ being $\boldsymbol{H}$-reduced. Unlike the previous examples, the $\boldsymbol{H}$-degree potentially changes at each step.

At the ith step suppose that $\boldsymbol{H}\text{-deg}\,\boldsymbol{P}^{(i)}=-\boldsymbol{d}^{(i)}$ and let $\boldsymbol{R}^{(i)}$ be the $m\times m$ residual. Note that since $\boldsymbol{P}^{(i)}$ is $\boldsymbol{H}$-reduced we have that $|\boldsymbol{n}|=|\boldsymbol{d}^{(i)}|+i$. We increase the order condition using a construction similar to that encountered in Corollaries 3.6 and 3.8. Let $S=\{\mu \mid \boldsymbol{R}^{(i)}_{1,\mu}(0)\neq 0\}$. If S is empty then we set $\boldsymbol{P}^{(i+1)}=\boldsymbol{P}^{(i)}$. Otherwise let k be an element of S having a maximum $d^{(i)}_k$ and let $\boldsymbol{U}$ be the matrix formed by replacing the kth row of the identity by $(\boldsymbol{R}^{(i)}_{1,1}(0),\ldots,\boldsymbol{R}^{(i)}_{1,m}(0))/\boldsymbol{R}^{(i)}_{1,k}(0)$. Set $\boldsymbol{d}^{(i+1)}=\boldsymbol{d}^{(i)}-\boldsymbol{e}_k$ where $\boldsymbol{e}_k$ denotes the kth row of the identity matrix. Then the matrix $\boldsymbol{V}=z^{\boldsymbol{d}^{(i)}}\cdot\boldsymbol{U}\cdot z^{-\boldsymbol{d}^{(i+1)}}$ is a polynomial matrix (by construction). Setting $\boldsymbol{P}^{(i+1)}=\boldsymbol{P}^{(i)}\cdot\boldsymbol{V}$ gives a basis of order $(\boldsymbol{\omega}^{(i+1)},\boldsymbol{G})$, which is $\boldsymbol{H}$-reduced with $\boldsymbol{H}$-degree $-\boldsymbol{d}^{(i+1)}$. Therefore, also this recurrence is a special case of Theorem 5.1.

Also the main recurrences of the algorithms presented in [3, 5, 20, 33, 34, 36–38] may be viewed as special cases of our Theorem 5.1.

When $\boldsymbol{\omega}^{(1)}=\boldsymbol{\omega}^{(2)}=(1,\ldots,1)$ then the only interpolation conditions occur at infinity. In this case, the definition of a basis are all those matrix polynomials $\boldsymbol{P}\in\mathscr{F}[z]$ having a constant determinant, that is, which are unimodular matrices. In this case the recursion in Theorem 5.1 generalizes the work of Antoulas [1] in his study of recursiveness in linear systems theory and takes the simple form

Corollary 5.6. *Let $\boldsymbol{P}^{(1)}\in\mathscr{F}[z]$ be unimodular, $\boldsymbol{H}\in\mathscr{F}[[z]]_-$ and define $\boldsymbol{H}':=\boldsymbol{H}\cdot\boldsymbol{P}^{(1)}$. Then $\boldsymbol{P}^{(2)}=\boldsymbol{P}^{(1)}\cdot\boldsymbol{P}^{(1,2)}$ is $\boldsymbol{H}$-reduced with $\boldsymbol{H}$-degree $\boldsymbol{d}$ iff $\boldsymbol{P}^{(1,2)}=(\boldsymbol{P}^{(1)})^{-1}\cdot\boldsymbol{P}^{(2)}$ is $\boldsymbol{H}'$-reduced with $\boldsymbol{H}'$-degree $\boldsymbol{d}'$. In both cases, $\boldsymbol{P}^{(1,2)}$ and $\boldsymbol{P}^{(2)}$ are unimodular, and $\boldsymbol{d}=\boldsymbol{d}'$.*

5.2. Connections to known recurrences for scalar rational interpolation

Theorem 5.1 also generalizes exactly the recursions in [21; 23, Theorems 9 and 9′] in the context of rational interpolation. Let $(p_{[a,b]},q_{[a,b]})$ denote the minimal Newton–Padé or proper multipoint Padé form of type $[a,b]$ as described in Section 4.3. The basic tool of the Gutknecht recursion are matrix polynomials consisting of neighboring interpolants

$$\boldsymbol{P}^{(1)}=\begin{pmatrix} p_{[\mu-1,\nu-1]} & p_{[\mu,\nu]} \\ q_{[\mu-1,\nu-1]} & q_{[\mu,\nu]} \end{pmatrix}. \tag{25}$$

The point $[\mu,\nu]$ is called weakly regular in the terminology of [21, 23] iff $\det\boldsymbol{P}^{(1)}\neq 0$. With $\boldsymbol{G},\boldsymbol{H}$ as in (17), (19), and $\boldsymbol{H}^{(1)}:=z^{-\mu-\nu}\cdot\boldsymbol{H}=z^{(-\mu,-\nu-i)}\cdot\boldsymbol{A}\cdot z^{(0,i)}$, it is not difficult to establish the equivalent characterization that $\boldsymbol{P}^{(1)}$ is an $\boldsymbol{H}^{(1)}$-reduced $(\boldsymbol{\omega}^{(1)},\boldsymbol{G})$-basis with $\boldsymbol{H}^{(1)}$-degree $(-1,0)$ (and increased order condition for the second column).

In the recursions [21, 23] one takes as a starting point the matrix $\boldsymbol{P}^{(1)}$ with $[\mu,\nu]$ being weakly regular, and wants to compute

$$\boldsymbol{P}^{(2)}=\begin{pmatrix} p_{[\mu+\kappa-1,\nu+\lambda-1]} & p_{[\mu+\kappa,\nu+\lambda]} \\ q_{[\mu+\kappa-1,\nu+\lambda-1]} & q_{[\mu+\kappa,\nu+\lambda]} \end{pmatrix} \tag{26}$$

for some $-\lambda \leqslant \kappa \leqslant \lambda$, i.e., an $H^{(2)}$-reduced $(\omega^{(2)}, G)$-basis with $\omega^{(2)} = (\omega_{0,\mu+\nu+\kappa+\lambda-1}, 1)$ and $H^{(2)} = z^{(-\kappa,-\lambda)} \cdot H^{(1)}$. Let us suppose for simplicity that also $[\mu+\kappa, \nu+\lambda]$ is a weakly regular point (otherwise, a link to [23, Theorems 9 and 9′] may be obtained by exploiting the relation $d^{(1,2)} = d^{(2)}$ of Theorem 5.1). By assumption,

$$G \cdot P^{(1)} = \omega^{(1)} \cdot R^{(1)}, \quad H^{(1)} \cdot P^{(1)} = S^{(1)} \cdot z^{(-1,0)},$$

with $R^{(1)} \in \mathscr{F}[[z]]_{\mathbb{F}_0}$, $S^{(1)} \in \mathscr{F}[[z]]_\infty$ being regular. Consequently, the quantity $H^{(1,2)}$ of Theorem 5.1 equals

$$H^{(1,2)} := H^{(2)} \cdot P^{(1)} = z^{(-\kappa-1,-\lambda-1)} \cdot S^{(1)} \cdot z^{(0,1)},$$

and the data $(G, A, H, \mu, \nu, i) := (R^{(1)}, S^{(1)}, z^{\kappa+\lambda+1} \cdot H^{(1,2)}, \kappa+1, \lambda, 1)$ fullfill the requirements for multipoint Padé problems as described in Section 4.3. We see that the transfer matrix $P^{(1,2)}$ of Theorem 5.1 contains as columns the two multipoint approximants of type $[\kappa, \lambda-1]$ and $[\kappa+1, \lambda]$ (with respect to these new data) as stated in [21, 23].

Note that our generalization has the added advantage that the starting point is allowed to be singular. In addition the starting point does not necessarily satisfy any degree constraints, although we may of course add some if convenient.

5.3. Remarks on complexity

With the notations of Theorem 5.1, and $H^{(1)} \in \mathscr{F}[[z]]_-$, it is quite easy to determine the complexity of calculating a $(\omega^{(2)}, G, H^{(2)})$-basis $P^{(2)}$ from a $(\omega^{(1)}, G, H^{(1)})$-basis $P^{(1)}$. Here we count as an essential operation the multiplication of a matrix polynomial on the right with matrices obtained from the $(m \times m)$ identity matrix by replacing a suitable column/row by a simple other column/row containing only scalars or simple monomials or $z - z_\nu$. We construct $P^{(2)}$ as a product $P^{(2)} = P^{(1)} \cdot P^{(1,3)} \cdot P^{(3,2)}$. Firstly, the unimodular $P^{(1,3)}$ is determined so that $P^{(1)} \cdot P^{(1,3)}$ is $H^{(2)}$-reduced and has the same order $(\omega^{(1)}, G)$. Proceeding as explained after Lemma 3.5, the number of required essential operations is at most

$$|H^{(2)}\text{-deg}\, P^{(1)}| - \chi(\omega^{(1)}, G) - \eta(H^{(2)}) = |H^{(2)}\text{-deg}\, P^{(1)}| - |H^{(1)}\text{-deg}\, P^{(1)}| + \eta(H^{(1)}) - \eta(H^{(2)}).$$

Secondly, $P^{(3,2)}$ is constructed so that $\{P^{(1)} \cdot P^{(1,3)}\} \cdot P^{(3,2)}$ remains $H^{(2)}$-reduced and has the order $(\omega^{(2)}, G)$. Proceeding as in Example 5.5 (see [5, 9, 36]), the number of essential operations in this part is at most

$$\deg \chi(\omega^{(1,2)}, (\operatorname{diag} \omega^{(1)})^{-1} \cdot G \cdot \{P^{(1)} \cdot P^{(1,3)}\}) = \deg \chi(\omega^{(2)}, G) - \deg \chi(\omega^{(1)}, G).$$

Let us also notice that for a numerically more stable procedure it may be preferable to solve the extended M-Padé approximation problem corresponding to the intermediate “small” problem $P^{(1,2)}$ by some classical method such as Gaussian elimination with partial pivoting.

6. Duality and invariance

6.1. Dual systems

So far we have considered only order conditions and interpolation at infinity that involves matrix multiplication on the left. There are similar concepts that can be defined for matrix multiplication on the right. This concept is best developed by the use of *dual systems*. Here we will make use of the cofactor A^* of a square matrix A of size m, being defined by $A^* := (\operatorname{adj} A)^{\mathrm{T}}$. Recall that $\det A^* = (\det A)^{m-1}$, and that in the case $\det A \neq 0$ we have

$$A^* = (\operatorname{adj} A)^{\mathrm{T}} = (A^{-1})^{\mathrm{T}} \cdot \det A, \qquad (A^*)^* = \operatorname{adj}(\operatorname{adj} A) = A \cdot (\det A)^{m-2}. \tag{27}$$

Definition 6.1. Given $\boldsymbol{G} \in \mathscr{F}[[z]]_{\mathbb{F}_0}$, $\boldsymbol{H} \in \mathscr{F}[[z]]_-$, a multi-index $\boldsymbol{d}$ and an order vector $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_m)$, respectively, we refer to the quantities $\boldsymbol{G}^* = (\operatorname{adj} \boldsymbol{G})^T \in \mathscr{F}[[z]]_{\mathbb{F}_0}$, $\boldsymbol{H}^* = (\operatorname{adj} \boldsymbol{H})^T \in \mathscr{F}[[z]]_-$,

$$\boldsymbol{d}^* := |\boldsymbol{d}| \cdot \boldsymbol{e} - \boldsymbol{d}, \quad \boldsymbol{\omega}^* := \left(\frac{\Omega}{\omega_1}, \ldots, \frac{\Omega}{\omega_m} \right), \quad \Omega = \det \operatorname{diag} \boldsymbol{\omega},$$

as the corresponding *dual parameters*.

The definition of dual order vectors and dual multiindices is motivated by the properties $\operatorname{diag}(\boldsymbol{\omega}^*) = (\operatorname{diag} \boldsymbol{\omega})^*$, and $z^{(\boldsymbol{d}^*)} = (z^{\boldsymbol{d}})^*$. Notice also that $\boldsymbol{G} \in \mathscr{F}[[z]]_{\mathbb{F}_0}$ is regular if and only if its dual counterpart $\boldsymbol{G}^*$ is regular.

In [29], Mahler established a close relationship between Hermite–Padé approximation problems of type I and II, which may be rewritten as a duality relation between particular basis elements of $\mathscr{M}(\boldsymbol{\omega}, \boldsymbol{G})$ and $\mathscr{M}(\boldsymbol{\omega}^*, \boldsymbol{G}^*)$. For the case of vector Hermite–Padé approximation, such a duality relation at normal points has been found by one of the authors [26, 25] considering particular degree constraints, and by De Samblanx et al. [20] for arbitrary degree constraints. Notice that duality relations are basic for deriving inversion formulas of block Hankel matrices [26, 25]. In addition, they are also one of the basic tools for proving weak stability of the Cabay–Meleshko algorithm for Padé approximation [17] and its matrix type generalizations [7, 13, 38].

The aim of this section is to establish similar duality results for our general framework. Let us first show that the adjoint operation induces a correspondence between order bases and also their behavior with regard to interpolation at infinity of these two order modules.

Lemma 6.2. *Let* $\boldsymbol{P} \in \mathscr{F}[z]$ *with* $\det \boldsymbol{P} \neq 0$, *and consider* $\boldsymbol{P}^* = (\operatorname{adj} \boldsymbol{P})^{\mathrm{T}}$. *Then:*
(a) *if* $\boldsymbol{P}$ *has order* $(\boldsymbol{\omega}, \boldsymbol{G})$ *then* $\boldsymbol{P}^*$ *has order* $(\boldsymbol{\omega}^*, \boldsymbol{G}^*)$;
(b) *if* $\boldsymbol{P}$ *has* $\boldsymbol{H}$*-degree of at least* $\boldsymbol{d}$ *then* $\boldsymbol{P}^*$ *has* $\boldsymbol{H}^*$*-degree of at least* $\boldsymbol{d}^*$;
(c) $\boldsymbol{P}$ *is* $\boldsymbol{H}$*-reduced iff* $\boldsymbol{P}^*$ *is* $\boldsymbol{H}^*$*-reduced, and in this case* $\boldsymbol{H}^*\text{-}\deg \boldsymbol{P}^* = (\boldsymbol{H}\text{-}\deg \boldsymbol{P})^*$.

Proof. Note that the matrices $\boldsymbol{P}, \boldsymbol{G}, \boldsymbol{H}, \boldsymbol{R}$ involved in (1) and (11) together with their dual counterpart all have a nontrivial determinant by assumption. Suppose that with suitable $\boldsymbol{R} \in \mathscr{F}[[z]]_{\mathbb{F}_0}$,

$S \in \mathscr{F}[[z]]_\infty$,

$$G \cdot P = \operatorname{diag} \omega \cdot R, \quad \text{and/or} \quad H \cdot P = S \cdot z^d.$$

Then

$$G^* \cdot P^* = \operatorname{diag}(\omega^*) \cdot R', \quad \text{and/or} \quad H^* \cdot P^* = S' \cdot z^{d^*},$$

where $R' = R^* \in \mathscr{F}[[z]]_{\mathbb{F}_0}$, and $S' = S^* \in \mathscr{F}[[z]]_\infty$, leading to (a), (b). Moreover, also part (c) follows, since $S(\infty)$ is regular if and only if $S'(\infty) = S^*(\infty) = S(\infty)^*$ is regular. □

From (27) we see that complete symmetry may only be expected for regular G.

Corollary 6.3. *Let in addition G be regular. Then:*
(a) *P is an (ω, G)-basis iff $P^* = (\operatorname{adj} P)^{\mathrm{T}}$ is an (ω^*, G^*)-basis;*
(b) *P is an (ω, G, H)-basis iff $P^* = (\operatorname{adj} P)^{\mathrm{T}}$ is an (ω^*, G^*, H^*)-basis;*
(c) *(ω, G, H) are (weakly) normal data if and only if (ω^*, G^*, H^*) are (weakly) normal data.*

Proof. First we know from Lemma 6.2 that P has the correct order iff P^* has. Also, the matrix G is regular iff its dual G^* is regular. Thus part (a) follows by applying the criterion of Lemma 2.8 since, with the notations of the proof of Lemma 6.2, $\det R(\alpha) \neq 0$ iff $\det R'(\alpha) \neq 0$ for any interpolation knot $\alpha \in \mathscr{F}[[z]]_{\mathbb{F}_0}$. Statement (b) is a trivial consequence of part (a) and Lemma 6.2(c). For part (c) we again apply Lemma 6.2(c). □

For the particular case of ordinary degree constraints, Corollary 6.3(a),(b) reduces to [20, Theorems 1 and 3]. Notice also that monic or comonic normalization are preserved under duality transformations.

From Corollary 6.3 we see that with $\operatorname{adj} P$ we have found the solution of some dual problem. However, the reciprocal of this statement has received much more attention in the last years since it is one of the main steps of the matrix generalizations of the weakly stable Cabay–Meleshko algorithm. In fact, a main criterion for "admissible" subproblems is that both the coefficients of a basis P and of its adjoint should be small. However, calculating explicitly $\operatorname{adj} P$ from P by taking determinants leads to very poor numerical results. Therefore it is preferable to solve simultaneously the dual problem, since the adjoint of P at normal locations may be obtained by suitably normalizing a solution of the dual problem (see Corollaries 4.5 and 6.3(c)).

Let us notice that there is also a close correspondence between transfer matrices of a system and transfer matrices of the dual system, as discussed, e.g., in [29, p. 125] for the particular case of M-Padé approximation. This follows at once by the fact that transfer matrices themselves are basis matrices, see Theorem 5.1.

Example 6.4 (*Type-*II *Hermite–Padé approximants* [29]). Let

$$F = (f_1, \ldots, f_m) \in \mathbb{F}^{1 \times m}[[z]]_{\mathbb{F}_0} \tag{28}$$

be a vector of power series and construct $\boldsymbol{G}$ as in Example 2.3. Then with $\boldsymbol{\omega} = (\omega, 1, \ldots, 1)$, and $\boldsymbol{H} = z^{-\boldsymbol{n}}$, we obtain M-Padé (type I) approximants of type $\boldsymbol{n}$. In this case, the dual system is

$$\boldsymbol{G}^* = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ -f_2 & f_1 & 0 & \ldots & 0 \\ -f_3 & 0 & f_1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -f_m & 0 & \ldots & 0 & f_1 \end{bmatrix} \in \mathscr{F}[[z]]_{\mathbb{F}_0}, \tag{29}$$

and $\boldsymbol{\omega}^* = (1, \omega, \ldots, \omega)$, $\boldsymbol{H}^* = \operatorname{adj} z^{-\boldsymbol{n}} = z^{-|\boldsymbol{n}|\cdot\boldsymbol{e}+\boldsymbol{n}}$. Therefore, the set

$$\mathscr{M}(\boldsymbol{\omega}^*, 0, \boldsymbol{G}^*, \boldsymbol{H}^*)$$

coincides with the set of simultaneous Padé approximants of type $\boldsymbol{n}$ [8, 9, 29]. □

Example 6.5 (*Left-hand and right-hand matrix Newton–Padé approximants* [8, 9, 25–27]). Let

$$\boldsymbol{G} = \begin{pmatrix} -\boldsymbol{I} & \boldsymbol{A} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \in \mathscr{F}[[z]]_{\mathbb{F}_0}, \qquad \boldsymbol{G}^* = (-1)^s \cdot \boldsymbol{G}^{\mathrm{T}} = (-1)^s \cdot \begin{pmatrix} -\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{A}^{\mathrm{T}} & \boldsymbol{I} \end{pmatrix} \in \mathscr{F}[[z]]_{\mathbb{F}_0}, \tag{30}$$

where $\boldsymbol{I}$ and $\boldsymbol{0}$ denote identity and zero matrices of size $s \times s$. Furthermore, let $\boldsymbol{\omega} = (\omega, \ldots, \omega, 1, \ldots, 1)$ (s ones), $\boldsymbol{H} = z^{-\boldsymbol{n}}$, with $\boldsymbol{n} = (p, \ldots, p, q, \ldots, q)$, $p + q + 1 = \deg \omega$. By (8) we have $\deg \chi(\boldsymbol{\omega}, \boldsymbol{G}) + \eta(\boldsymbol{H}) = s \cdot (\deg \omega - p - q) = s$. Thus the linear solution set $\mathscr{M}(\boldsymbol{\omega}, 0, \boldsymbol{G}, \boldsymbol{H})$ contains at least s solutions being linearly independent over $\mathbb{F}$, which may be obtained from an $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis $\boldsymbol{P}$ as described in Theorem 4.1. In fact, combining s such solutions in a matrix $(N^{\mathrm{T}}, D^{\mathrm{T}})^{\mathrm{T}}$ with N, D being square yields a right-hand matrix Newton–Padé form (MNPF) (N, D) of type (p, q) for the formal matrix-Newton series $\boldsymbol{A}$. Note that, by Corollary 4.6(c), this form is unique up to multiplication on the right with a scalar matrix if and only if $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$ is weakly normal.

If there exists a right-hand MNPF (N, D) of type (p, q) of $\boldsymbol{A}$ satisfying in addition $\det D \neq 0$, then the matrix rational function $N \cdot D^{-1}$ is called a matrix Newton–Padé approximant of type (p, q) of $\boldsymbol{A}$. Note that $\det D \neq 0$ implies in particular that $\boldsymbol{d} := z^{-\boldsymbol{n}}\text{-}\deg \boldsymbol{P}$ must have at least s components less than or equal to 0 due to Corollary 4.6(b). Also, we may have unattainable points, i.e., zeros of $\det D$ out of ω, and so the rational function may no longer agree with $\boldsymbol{A}$ at all interpolation points.

Let $\boldsymbol{Q}$ denote a $(\boldsymbol{\omega}^*, \boldsymbol{G}^*, \boldsymbol{H}^*)$-basis (for instance $\boldsymbol{Q} = \boldsymbol{P}^*$, see Corollary 6.3). Note that $\boldsymbol{\omega}^* = \omega^{s-1} \cdot \boldsymbol{\omega}_0$, $\boldsymbol{\omega}_0 := (1, \ldots, 1, \omega, \ldots, \omega)$ (s ones). Thus, due to Lemma 2.8, $\boldsymbol{Q}_0 := \omega^{1-s} \cdot \boldsymbol{Q}$ is a matrix polynomial. Moreover, since $\boldsymbol{H}^* = z^{-\boldsymbol{n}_0 - (s-1)(p+q)\cdot\boldsymbol{e}}$, $\boldsymbol{n}_0 := (q, \ldots, q, p, \ldots, p)$, we see that $\boldsymbol{Q}_0$ is an $(\boldsymbol{\omega}_0, \boldsymbol{G}^*, z^{-\boldsymbol{n}_0})$-basis with columns building up the right hand MNPF of type (p, q) of the transposed of $\boldsymbol{A}$, or, in other words, the left-hand MNPF of type (p, q) of $\boldsymbol{A}$.

Let $z^{-\boldsymbol{n}_0}\text{-}\deg \boldsymbol{Q}_0 = \boldsymbol{d}_0$. In the case $\boldsymbol{Q} = \boldsymbol{P}^*$ we may conclude from Lemma 6.2(c) that $\boldsymbol{d} + \boldsymbol{d}_0 = \boldsymbol{e}$. Also, from Theorem 4.3 we know that this property remains valid for arbitrary bases $\boldsymbol{Q}$ (after a suitable permutation of columns). Suppose therefore that there exists a right-hand MNPF (N, D), and a left hand MNPF (N^*, D^*), both of type (p, q), satisfying $\det D \neq 0$, and $\det D^* \neq 0$, respectively. It is well known [27] that then the matrix rational functions $N \cdot D^{-1}$ and $(D^*)^{-1} \cdot N^*$ coincide. In addition we see from $\boldsymbol{d} + \boldsymbol{d}_0 = \boldsymbol{e}$ that in fact both vectors $\boldsymbol{d}$ and $\boldsymbol{d}_0$ must contain exactly s components at most 0.

6.2. Invariance under Moebius transforms

For any interpolation problem that includes the point at infinity it is a natural question to ask if this problem is invariant under linear transformations of the extended complex plane. This is indeed the case for the extended M-Padé problem. Note that this invariance property is not immediate for other two-point approximation problems found in the literature. In this section we will prove this invariance in the special case of $\boldsymbol{G}$ being a regular matrix Newton series with respect to the set $\mathbb{F}_0 = \{\alpha_0, \alpha_1, \ldots, \alpha_s\} \subset \mathbb{F}$, and for $\boldsymbol{H}$ having a particular form as discussed in Lemma 3.9(b). The general case follows a similar argument but with a considerable increase in notation.

The problem of finding an $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$-basis $\boldsymbol{P}$ may therefore be restated as follows: define the vectors of nonnegative indices $\boldsymbol{r}_j$ by

$$\operatorname{diag} \boldsymbol{\omega}(z) = (z-\alpha_0)^{\boldsymbol{r}_0} \cdot \cdots \cdot (z-\alpha_s)^{\boldsymbol{r}_s}.$$

Furthermore, denote by $\boldsymbol{A}_j$ the power series expansions of $\boldsymbol{G}$ around α_j, $j = 0, \ldots, s$, with $\det \boldsymbol{A}_j(\alpha_j) \neq 0$. Finally, let $\boldsymbol{H}(z) = z^{\boldsymbol{r}_\infty} \cdot \boldsymbol{A}_\infty$ with $\det \boldsymbol{A}_\infty(\infty) \neq 0$. We are looking for a matrix polynomial $\boldsymbol{P}$, $\det \boldsymbol{P} \neq 0$, and a vector $\boldsymbol{d}$ of integers such that

$$(z-\alpha_j)^{-\boldsymbol{r}_j} \cdot \boldsymbol{A}_j \cdot \boldsymbol{P} = \mathcal{O}(1)_{z\to\alpha_j}, \quad j = 0, \ldots, s, \tag{31}$$

$$z^{\boldsymbol{r}_\infty} \cdot \boldsymbol{A}_\infty \cdot \boldsymbol{P} \cdot z^{-\boldsymbol{d}} = \mathcal{O}(1)_{z\to\infty}, \tag{32}$$

$$|\boldsymbol{r}_0 + \cdots + \boldsymbol{r}_s + \boldsymbol{r}_\infty| = |\boldsymbol{d}|. \tag{33}$$

Suppose for a moment that $|\boldsymbol{r}_0 + \cdots + \boldsymbol{r}_s + \boldsymbol{r}_\infty| = 0$. We see that in general the finite interpolation points may be exchanged without changing essentially the problem, however, the interpolation condition at infinity has a particular form. In fact, from Theorem 4.3 we see that the "ordinary" interpolation condition $z^{\boldsymbol{r}_\infty} \cdot \boldsymbol{A}_\infty \cdot \boldsymbol{P} = \mathcal{O}(1)_{z\to\infty}$ can be verified if and only if $(\boldsymbol{\omega}, \boldsymbol{G}, \boldsymbol{H})$ are normal data. Let us show as well that there is an invariance of our problem with respect to a linear transformation of the extended complex plane.

Consider the change of variable $z = T(\underline{z})$ with T denoting a Moebius transform

$$z = T(\underline{z}) = \frac{a\cdot\underline{z}+b}{c\cdot\underline{z}+d} \quad \text{with } \delta := ad - bc \neq 0$$

(i.e., T is nontrivial). Note that

$$T(\infty) - T(\underline{z}) = \frac{\delta}{c\cdot(c\cdot\underline{z}+d)} \quad \text{for } \underline{z} \neq \infty \tag{34}$$

$$T(\underline{z}_1) - T(\underline{z}_2) = \frac{\delta\cdot(\underline{z}_1 - \underline{z}_2)}{(c\cdot\underline{z}_1+d)\cdot(c\cdot\underline{z}_2+d)} \quad \text{if } \underline{z}_1 \text{ and } \underline{z}_2 \text{ are different from } \infty. \tag{35}$$

In the sequel, all transformed quantities are underlined. Also, we will not consider the trivial case of translation, and suppose therefore that $c \neq 0$, i.e., $T(\infty) \neq \infty$. We introduced the new interpolation points by

$$T(\underline{\alpha}_j) = \alpha_j, \quad j = 1, \ldots, s, \qquad T(\underline{\alpha}_0) = \infty, \qquad T(\infty) = \alpha_0$$

($\alpha_0 = a/c$ may be always considered as an interpolation point by eventually taking $\boldsymbol{r}_0 = \boldsymbol{0}$). Then the points $\underline{\alpha}_0, \ldots, \underline{\alpha}_s, \infty$ are distinct.

Let N be a sufficiently large integer so that each component of $N\boldsymbol{e}+\boldsymbol{r}_\infty$ is greater or equal to zero. From Lemma 3.9(b) together with (32) we may conclude that $\boldsymbol{I}$-deg $\boldsymbol{P} \leqslant N\boldsymbol{e}+\boldsymbol{d}$, i.e., the degree of any polynomial in the jth column of $\boldsymbol{P}$ is bounded by $N+d_j$. Hence

$$\underline{\boldsymbol{P}}(\underline{z}) := \boldsymbol{P}(T(\underline{z})) \cdot (\underline{z}-\underline{\alpha}_0)^{N\boldsymbol{e}+\boldsymbol{d}}$$

is a matrix polynomial. We want to show that it is the basis matrix of an extended M-Padé problem, obtained by applying the coordinate transformation on the data occurring in (31)–(33). We get from (32)

$$\left(\frac{a\underline{z}+b}{c}\right)^{\boldsymbol{r}_\infty} \cdot (\underline{z}-\underline{\alpha}_0)^{-N\boldsymbol{e}-\boldsymbol{r}_\infty} \cdot \underline{\boldsymbol{A}}_0(\underline{z}) \cdot \underline{\boldsymbol{P}}(\underline{z}) \cdot \left(\frac{a\underline{z}+b}{c}\right)^{-\boldsymbol{d}} = \mathcal{O}(1)_{\underline{z}\to\underline{\alpha}_0},$$

where $\underline{\boldsymbol{A}}_0(\underline{z}) := \boldsymbol{A}_\infty(T(\underline{z}))$ is a regular power series around $\underline{\alpha}_0$. Taking into account that $a\underline{z}+b = a(\underline{z}-\underline{\alpha}_0) - \delta/c \neq a\cdot(\underline{z}-\underline{\alpha}_0)$, we see that $\underline{\boldsymbol{P}}$ has order $(N\boldsymbol{e}+\boldsymbol{r}_\infty)$ at $\underline{\alpha}_0$. Moreover, (31) yields for $j=1,\ldots,s$ using (35)

$$\left(\frac{c\cdot(c\underline{\alpha}_j+d)}{\delta} \cdot (\underline{z}-\underline{\alpha}_0)\right)^{\boldsymbol{r}_j} \cdot (\underline{z}-\underline{\alpha}_j)^{-\boldsymbol{r}_j} \cdot \underline{\boldsymbol{A}}_j(\underline{z}) \cdot \underline{\boldsymbol{P}}(\underline{z}) \cdot (\underline{z}-\underline{\alpha}_0)^{-N\boldsymbol{e}-\boldsymbol{d}} = \mathcal{O}(1)_{\underline{z}\to\underline{\alpha}_j},$$

where $\underline{\boldsymbol{A}}_j(\underline{z}) := \boldsymbol{A}_j(T(\underline{z}))$ is a regular power series around $\underline{\alpha}_j$. Therefore, $\underline{\boldsymbol{P}}$ has also order $\boldsymbol{r}_j$ at $\underline{\alpha}_j$, $j=1,\ldots,s$. It remains to consider the point $\alpha_0 = T(\infty)$. Here we get from (31) and (35)

$$\left(-\frac{c^2}{\delta}\cdot\left(1-\frac{\underline{\alpha}_0}{\underline{z}}\right)\right)^{\boldsymbol{r}_0} \cdot \underline{z}^{\boldsymbol{r}_0} \cdot \underline{\boldsymbol{A}}_\infty(\underline{z}) \cdot \underline{\boldsymbol{P}}(\underline{z}) \cdot \underline{z}^{-N\boldsymbol{e}-\boldsymbol{d}} \cdot \left(1-\frac{\underline{\alpha}_0}{\underline{z}}\right)^{-N\boldsymbol{e}-\boldsymbol{d}} = \mathcal{O}(1)_{\underline{z}\to\infty},$$

with $\underline{\boldsymbol{A}}_\infty(\underline{z}) := A_0(T(\underline{z}))$ being a regular power series around infinity. Thus,

$$\underline{\boldsymbol{H}}(\underline{z}) \cdot \underline{\boldsymbol{P}}(\underline{z}) \cdot \underline{z}^{-N\boldsymbol{e}-\boldsymbol{d}} = \mathcal{O}(1)_{\underline{z}\to\infty}, \qquad \underline{\boldsymbol{H}}(\underline{z}) = \underline{z}^{\boldsymbol{r}_0} \cdot \underline{\boldsymbol{A}}_\infty(\underline{z}),$$

and we see immediately that the new order vectors together with the $\underline{\boldsymbol{H}}$-degree of $\underline{\boldsymbol{P}}$ satisfy (33). Thus, also $\underline{\boldsymbol{P}}$ is the solution of an extended M-Padé problem.

From the above considerations it also becomes clear that, instead of allowing for some freedom in the interpolation condition at infinity (32), we could equivalently introduce the parameter $\boldsymbol{d}$ for interpolation at some finite interpolation knot.

7. Future research

If we consider only the case of interpolation at both 0 and ∞ then it is possible to provide an explicit matrix representation of the corresponding linear system of equations. In this case the matrices are of the form of mosaic Toeplitz matrices with each partition being either lower or upper triangular. These generalize in a natural way the Sylvester matrices that appear in Padé approximation.

It is well known that in the context of linear algebra that dividing a problem of linear systems of equations into two smaller systems of linear equations is obtained by the Schur complement method. In our context it is well known that the structures of the problem are not preserved by the Schur complement method. However, in a future publication [10] we will show that our recursion can be viewed as a modified Schur complement method that does preserve the structure.

The recursion that we have given in this paper makes the assumption that we are using exact arithmetic for our computations. Thus we do not address the problems of working with numerical floating point arithmetic. In these cases the problem is one of numerical instabilities. These problems have been addressed in special cases of our recurrence by the weakly stable offdiagonal algorithms in [16, 13] for Padé and Hermite–Padé problems, respectively, and also for row recurrences of rectangular matrix Padé approximants in [38]. The basic tools for these weakly stable algorithms are so called stable points (normal points where the underlying linear system is well conditioned) and near inversion formulae in terms of both an order basis and its dual. The near inversion formula along with a weakly stable version of our recursive computation will be given in a forthcoming paper.

Although we have presented our recursion assuming exact arithmetic we have not considered the problem of the growth of coefficients in exact arithmetic. In this case the natural domain of coefficients will not be a field but rather an integral domain (usually one of multivariate polynomials). In this case one can consider using fraction-free methods for reducing the cost of single arithmetic operations by avoiding unnecessary gcd computations for the coefficients. This remains a topic for future research.

In this paper we have only considered the algebraic problem of computing our interpolants. However, recently also some first convergence results for suitably scaled basis matrices have been obtained in the scalar and in the matrix setting (see [11] and the references therein). For instance, let

$$f(z) := \int_I \frac{\mathrm{d}\mu(x)}{z - x}$$

be some Markov function with I consisting of several real compact intervals, and μ some positive Borel measure being supported on I. Then it is shown in [11, Theorem 3.14] that there exists basis matrices

$$\boldsymbol{P}_n = \begin{pmatrix} p_n & p_{n+1} \\ q_n & q_{n+1} \end{pmatrix}$$

such that the sequence of *modified Padé approximants* (at infinity)

$$\tilde{\pi}_n(z) := (p_n(z), p_{n+1}(z)) \cdot (q_n(z), q_{n+1}(z))^{+}$$

converges to f locally uniformly in $\mathbb{C}\backslash I$ (here we use the pseudoinverse in order to allow for "rectangular denominators"). This has to be compared with the Markov convergence theorem insuring locally uniform convergence only outside the smallest convex set containing I.

References

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

[2] G.A. Baker and P.R. Graves-Morris, *Padé Approximants* (Cambridge Univ. Press, Cambridge, 2nd ed., 1995).

[3] B. Beckermann, Zur Interpolation mit polynomialen Linearkombinationen beliebiger Funktionen, Thesis, Univ. Hannover, 1990.

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

[5] B. Beckermann, A reliable method for computing M-Padé approximants on arbitrary staircases, *J. Comput. Appl. Math.* **40** (1992) 19–42.

[6] B. Beckermann, Nonsymmetric difference operators and polynomials being orthogonal with respect to rectangular matrix valued measures, Publication ANO 335, Université de Lille, 1995.

[7] B. Beckermann, The stable computation of formal orthogonal polynomials, *Numer. Algorithms* **11** (1996) 1–23.

[8] B. Beckermann and G. Labahn, A uniform approach for Hermite Padé and simultaneous Padé approximants and their matrix-type generalizations, *Numer. Algorithms* **3** (1992) 45–54.

[9] B. Beckermann and G. Labahn, A uniform approach for the fast, reliable computation of matrix-type Padé approximants, *SIAM J. Matrix Anal. Appl.* **15** (1994) 804–823.

[10] B. Beckermann, G. Labahn and S. Cabay, Modified Schur complements and fraction-free computation of matrix rational approximants, in preparation.

[11] B. Beckermann and V. Kaliaguine, The diagonal of the Padé table and the approximation of the Weyl function of second order difference operators, Publication ANO 338, Université de Lille, 1995, *Constr. Approximation*, to appear.

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

[13] S. Cabay, A.R. Jones and G. Labahn, Computation of numerical Padé–Hermite and simultaneous Padé systems II: A weakly stable algorithm, *SIAM J. Matrix Anal.* **17** (1996) 268–297.

[14] S. Cabay and G. Labahn, A superfast algorithm for multidimensional Padé systems, *Numer. Algorithms* **2** (1992) 201–224.

[15] S. Cabay, G. Labahn and B. Beckermann, On the theory and computation of non-perfect Padé–Hermite approximants, *J. Comput. Appl. Math.* **39** (1992) 295–313.

[16] S. Cabay and R. Meleshko, A weakly stable algorithm for Padé approximants and the inversion of Hankel matrices, *SIAM J. Matrix Anal. Appl.* **14** (1993) 735–765.

[17] G. Claessens, On the Newton–Padé approximation problem, *J. Approx. Theory* **22** (1978) 150–160.

[18] J. Coates, On the algebraic approximation of functions, *Indag. Math.* **28** (1966) 421–461.

[19] M.G. de Bruin, Some aspects of simultaneous rational approximation, in: *Numerical Analysis and Mathematical Modelling*, Banach Center Publications **24** (PWN-Polish Scientific Publishers, Warsaw, 1990) 51–84.

[20] G. De Samblanx, M. Van Barel and A. Bultheel, Duality in vector Padé Hermite approximation problems, *J. Comput. Appl. Math.* **66** (1996) 153–166.

[21] M.H. Gutknecht, Block structure and recursiveness in rational interpolation, in: E.W. Cheney, C.K. Chui and L.L. Schumaker, Eds., *Approximation Theory VII* (Academic Press, New York, 1993) 93–130.

[22] M.H. Gutknecht, Stable row recurrences for the Padé table and generically superfast look-ahead solvers for non-Hermitean Toeplitz systems, *Linear Algebra Appl.* **188/89** (1993) 351–421.

[23] M.H. Gutknecht, The multipoint Padé table and general recurrences for rational interpolation, in: A. Cuyt, Ed., *Nonlinear Numerical Methods and Rational Approximation* (Kluwer Academic Publishers, Dordrecht, 1994) 109–136.

[24] H. Jager, A multidimensional generalization of the Padé table, *Indag. Math.* **26** (1964) 193–249.

[25] G. Labahn, Inversion algorithms for rectangular-block Hankel matrices, Research Report CS-90-52, Univ. of Waterloo, 1990.

[26] G. Labahn, Inversion components of block Hankel-like matrices, *Linear Algebra Appl.* **177** (1992) 7–48.

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

[28] S. Lang, *Algebra* (Addison-Wesley, Reading, MA, 1984).

[29] K. Mahler, Perfect systems, *Compositio. Math.* **19** (1968) 95–166.

[30] J.H. McCabe and J.A. Murphy, Continued fractions which correspond to power series expansions at two points, *J. Inst. Math. Appl.* **17** (1976) 233–247.

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

[32] V.N. Sorokin and J. Van Iseghem, Algebraic aspects of matrix orthogonality for vector polynomials, Publication ANO 345, Université de Lille, 1995; *J. Approx. Theory*, to appear.

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

[34] M. Van Barel and A. Bultheel, The computation of non-perfect Padé–Hermite approximants, *Numer. Algorithms* **1** (1991) 285–304.
[35] M. Van Barel and A. Bultheel, A new formal approach to the rational interpolation problem, *Numer. Math.* **62** (1992) 87–122.
[36] M. Van Barel and A. Bultheel, A general module theoretic framework for vector M-Padé and matrix rational interpolation, *Numer. Algorithms* **3** (1992) 451–462.
[37] M. Van Barel and A. Bultheel, A look-ahead method to compute vector Padé–Hermite approximants, *Constr. Approx.* **11** (1995) 455–476.
[38] M. Van Barel and A. Bultheel, A look-ahead algorithm for the solution of block Toeplitz systems, Report TW 224, Katholieke Universiteit Leuven, March 1995.
[39] J. Van Iseghem, Vector orthogonal relations, vector QD-algorithm, *J. Comput. Appl. Math.* **19** (1987) 141–150.