THE LINEAR PENCIL APPROACH TO RATIONAL INTERPOLATION

BERNHARD BECKERMANN, MAXIM DEREVYAGIN, AND ALEXEI ZHEDANOV

ABSTRACT. It is possible to generalize the fruitful interaction between (real or complex) Jacobi matrices, orthogonal polynomials and Padé approximants at infinity by considering rational interpolants, (bi-)orthogonal rational functions and linear pencils $zB - A$ of two tridiagonal matrices A, B, following Spiridonov and Zhedanov.

In the present paper, beside revisiting the underlying generalized Favard theorem, we suggest a new criterion for the resolvent set of this linear pencil in terms of the underlying associated rational functions. This enables us to generalize several convergence results for Padé approximants in terms of complex Jacobi matrices to the more general case of convergence of rational interpolants in terms of the linear pencil. We also study generalizations of the Darboux transformations and the link to biorthogonal rational functions. Finally, for a Markov function and for pairwise conjugate interpolation points tending to ∞, we compute explicitly the spectrum and the numerical range of the underlying linear pencil.

1. INTRODUCTION

The connection with Jacobi matrices has led to numerous applications of spectral techniques for self-adjoint operators in the theory of orthogonal polynomials on the real line and Padé approximation. In order to give an idea of these interactions consider a Markov function of the form

$$\varphi(z) = \int_a^b \frac{d\mu(t)}{z-t},$$

where a, b are real numbers and $d\mu(t)$ is a probability measure, that is, $\int_a^b d\mu(t) = 1$. It is well known [1], [31] that one can expand such a Markov function φ into the following continued fraction

$$\varphi(z) = \cfrac{1}{z - b_0 - \cfrac{a_0^2}{z - b_1 - \cfrac{a_1^2}{\ddots}}} = \frac{1}{|z-b_0} - \frac{a_0^2}{|z-b_1} - \frac{a_1^2}{|z-b_2} - \cdots, \tag{1.1}$$

where $b_j, a_j \in \mathbb{R}$, $a_j > 0$. Continued fractions of the form (1.1) are called J-fractions [21, 31]. To the continued fraction (1.1) one can associate a Jacobi matrix A acting in ℓ^2, the space of square summable sequences, and its truncation $A_{[0:n]}$

$$A = \begin{pmatrix} b_0 & a_0 & & \\ a_0 & b_1 & a_1 & \\ & a_1 & b_2 & \ddots \\ & & \ddots & \ddots \end{pmatrix}, \quad A_{[0:n-1]} = \begin{pmatrix} b_0 & a_0 & & \\ a_0 & b_1 & \ddots & \\ & \ddots & \ddots & a_{n-2} \\ & & a_{n-2} & b_{n-1} \end{pmatrix}.$$

2000 *Mathematics Subject Classification.* 47B36, 40A15, 30E10, 47A57.

Key words and phrases. Multipoint Padé approximation, rational interpolation, MP continued fractions, Jacobi matrix, linear pencils.

Then it is known that $\varphi(z) = \langle (zI - A)^{-1} e_0, e_0 \rangle$, and the nth convergent of the above continued fraction is given by

$$\frac{p_n(z)}{q_n(z)} = \langle (zI - A_{[0:n-1]})^{-1} e_0, e_0 \rangle = \frac{1}{\lceil z - a_0} - \cdots - \frac{b_{n-2}^2 \rceil}{\lceil z - a_{n-1}},$$

where the column vector $e_0 = (1, 0, \dots)^\top$ is the first canonical vector of suitable size, q_n are orthogonal polynomials with respect to $d\mu$, and p_n are polynomials of the second kind, see [1, 24, 25]. It is elementary fact of the continued fraction theory that

$$\varphi(z) - \frac{p_n(z)}{q_n(z)} = O\left(\frac{1}{z^{2n+1}}\right)_{z \to \infty}, \tag{1.2}$$

see for instance [1, 4, 21]. Relation (1.2) means that the rational function p_n/q_n is the nth diagonal Padé approximant to φ at infinity. Consequently, the locally uniform convergence of diagonal Padé approximants appears as the strong resolvent convergence of the finite matrix approximations $A_{[0:n]}$. For instance, one knows that $p_n/q_n \to \varphi$ in capacity in the resolvent of A given by the complement of the support of μ, and locally uniformly outside the numerical range of A given by the convex hull of the spectrum of A, see for instance [28]. Besides, it should be mentioned here that an operator approach for proving convergence of Padé approximants for rational perturbations of Markov functions was proposed in [15], see also [14].

If φ is no longer a Markov function but has distinct nth Padé approximants at infinity, we may still recover these approximants as convergents of a continued fraction of type (1.1), but now in general $a_j, b_j \in \mathbb{C}$, $a_j \neq 0$, see [31], that is, A becomes complex symmetric, called a complex Jacobi matrix. There is no longer a natural candidate for the spectrum of A, but it is still possible to characterize the spectrum in terms of some asymptotic behavior of the Padé denominators $q_n(z)$ and the linearized error functions $r_n(z) = q_n(z)\phi(z) - p_n(z)$ [3, 11, 9], see also [7, 15, 14] for more general banded matrices. Convergence outside the numerical range was established in [11], and convergence in capacity in the outer connected component of the resolvent set in [10]. We refer the reader to [8] for some recent summary on complex Jacobi matrices, including some open questions partially solved in [6].

The goal of this paper is to generalize several of the above results to the case of multipoint Padé approximants.

Definition 1.1 ([4]). *The $[n_1|n_2]$ multipoint Padé approximant (or rational interpolant) for a function φ at the points $\{z_k\}_{k=1}^{\infty}$ is defined as the ratio p/q of two polynomials p and $q \neq 0$ of degree at most n_1 and n_2, respectively, such that $\varphi q - p$ vanishes at $z_1, z_2, ..., z_{n_1+n_2+1}$ counting multiplicities.*

It is easy to see that the degree and interpolation conditions lead to a homogeneous system of linear equations, and thus a $[n_1|n_2]$ multipoint Padé approximant exists. Also, one may show uniqueness of the fraction p/q. However, since the denominator may vanish at some of the interpolation points, it may happen that the fraction p/q does not interpolate φ at some point z_k, usually referred to as an unattainable point.

Under some regularity conditions, the $[n-1|n]$ multipoint Padé approximants of φ may be written as nth convergents of a continued fraction of the form

$$\frac{1}{\lceil z - b_0} - \frac{a_0^2(z - z_1)(z - z_2) \rceil}{\lceil z - b_1} - \frac{a_1^2(z - z_3)(z - z_4) \rceil}{\lceil z - b_2} - \dots, \tag{1.3}$$

the odd part of a Thiele continued fraction [4]. Continued fractions of this type are referred to as MP-fractions in [19] and as R_{II}–fractions in [20]. In particular, the authors study in [19, Theorem 4.4] and [20, Theorem 3.5] some analog of Favard's theorem and the link with orthogonal rational functions. Spiridonov and one of the authors [27, 32] showed that such continued fractions are related not to a single Jacobi matrix but to a pencil $zB - A$ with tridiagonal matrices A, B. Various links to bi-orthogonal rational functions have been presented in [32] and [16]. In particular, in [16, Theorem 6.2] the authors present an operator-theoretic proof for the Markov convergence theorem multipoint Padé approximants [18] based on spectral properties of the pencil $zB - A$.

The aim of this paper is to present further convergence results for the continued fraction (1.3), both in the resolvent set and outside the numerical range of the tridiagonal linear pencil $zB - A$. To be more precise, denote by $\ell^2 = \ell^2_{[0:\infty)}$ the Hilbert space of complex square summable sequences $(x_0, x_1, \dots)^\top$ with the usual inner product

$$\langle x, y\rangle = \sum_{j=0}^{\infty} x_j \overline{y}_j, \quad x, y \in \ell^2.$$

We will restrict our attention to the case of tridiagonal matrices A, B with bounded entries, in which case we may identify via usual matrix product the matrices A and B with bounded operators acting in ℓ^2. Notice that many algebraic relations remain true in the unbounded case as well. However, already the simpler case of bounded pencils allows to describe the main ideas of how to generalize results from the classical theory of orthogonal polynomials to the theory of biorthogonal rational functions as well as to the multipoint Padé approximation.

The remainder of the paper is organized as follows: we start from a general bounded MP–fraction and introduce the associated linear pencils together with the rational solutions of some underlying three term recurrence relations in §2.1. In §2.2, by generalizing previous work of Aptekarev, Kaliaguine & Van Assche [3] we show how the asymptotic behavior of these rational solutions allows to decide whether the linear pencil $zB - A$ is boundedly invertible. In particular, we deduce in Corollary 2.5 the pointwise convergence of at least a subsequence of our multipoint Padé approximants towards what is called the m–function (or Weyl function) of the linear pencil. Subsequently, we present in Theorem 2.9 of §2.3 an alternate proof for a Favard-type theorem based on orthogonality properties of associated rational functions, which yields in Corollary 2.11 a simple proof for the fact that the convergents of our continued fractions are indeed multipoint Padé approximants of the m-function of our linear pencil. In §3 we generalize the above-mentioned results of [11, Theorem 3.10], [10, Theorem 3.1], and [10, Theorem 4.4], on the convergence of Padé approximants at infinity in terms of complex Jacobi matrices to the more general case of multi-point Padé approximants in terms of linear pencils $zB - A$. The aim of §4 is to explore LU and UL decompositions of our linear pencil, and the link to biorthogonal rational functions. This naturally leads us to consider generalizations of the Darboux transformations of [12]. Finally, we generalize in §5 the findings described in the begining of this section, namely, if we start with a Markov function and pairwise conjugate interpolation points tending to infinity, then the spectrum of our linear pencil is still the support of the underlying measure, and the numerical range equals its convex hull.

2. Continued fractions, linear pencils, and their resolvent

In this section we show the links between continued fractions in question and linear pencils. Moreover, we prove a Favard type result for the corresponding recurrence relation.

2.1. **Linear pencils.** Let us consider a continued fraction of the form

$$\frac{1\;|}{|\;\beta_0(z)} - \frac{\alpha_0^L(z)\alpha_0^R(z)\;|}{|\;\beta_1(z)} - \frac{\alpha_1^L(z)\alpha_1^R(z)\;|}{|\;\beta_2(z)} - \dots \tag{2.1}$$

where $\beta_n, \alpha_n^L, \alpha_n^R$ are polynomials of degree at most 1 and not identically zero. Next, denote by $C_n(z)$ the nth convergent of this continued fraction obtained by taking only the first n terms in (2.1), then the well-known theory of continued fractions tells us that $C_n(z) = p_n(z)/q_n(z)$, where the polynomials p_n of degree $\leq n-1$ and q_n of degree $\leq n$ are obtained as solutions of the three-term recurrence relation

$$y_{n+1} = \beta_n(z)y_n - \alpha_{n-1}^L(z)\alpha_{n-1}^R(z)y_{n-1}, \quad n = 0, 1, 2, \dots \tag{2.2}$$

by means of the initial conditions (setting $\alpha_{-1}^L = \alpha_{-1}^R = 1$ for convenience)

$$q_0(z) = 1, \quad q_{-1}(z) = 0, \quad p_0(z) = 0, \quad p_{-1}(z) = -1. \tag{2.3}$$

Using (2.2) and (2.3) one easily verifies by recurrence that

$$q_n(z) = \det(zB_{[0:n-1]} - A_{[0:n-1]}), \quad p_n(z) = \det(zB_{[1:n-1]} - A_{[1:n-1]}). \tag{2.4}$$

By Cramer's rule, this implies the following formula for the convergents

$$C_n(z) = \frac{p_n(z)}{q_n(z)} = \langle (zB_{[0:n-1]} - A_{[0:n-1]})^{-1} e_0, e_0 \rangle. \tag{2.5}$$

By induction, one also easily shows the Liouville-Ostrogradsky formula

$$p_{n+1}(z)q_n(z) - p_n(z)q_{n+1}(z) = \prod_{k=0}^{n-1} \alpha_k^L(z)\alpha_k^R(z), \quad n = 0, 1, 2, \dots. \tag{2.6}$$

For a complex number $\phi(z)$, the sequence defined by

$$r_n(z) := \phi(z)q_n(z) - p_n(z), \tag{2.7}$$

gives another solution of (2.2) with initial conditions

$$r_0(z) = \phi(z), \quad r_{-1}(z) = 1. \tag{2.8}$$

We will refer to r_n as linearized error (or function of the second kind) since, from the Pincherle Theorem [21, Theorem 5.7], the continued fraction (2.1) has a limit $\phi(z)$ iff $r_n(z)$ is a minimial solution of the recurrence relation (2.2).

It will be convenient to write the polynomials α_j^L, α_j^R, and β_j occurring in (2.1) in the form of the tridiagonal infinite linear pencil

$$zB - A = \begin{pmatrix} \beta_0(z) & -\alpha_0^R(z) & 0 & 0 & \dots \\ -\alpha_0^L(z) & \beta_1(z) & -\alpha_1^R(z) & 0 & \ddots \\ 0 & -\alpha_1^L(z) & \beta_2(z) & \alpha_2^R(z) & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \tag{2.9}$$

with the two tridiagonal infinite matrices $A = (a_{i,j})_{i,j=0}^{\infty}$ and $B = (b_{i,j})_{i,j=0}^{\infty}$. For a J-fraction, we obtain the linear pencil $z - A$ with a tridiagonal matrix A [1] (see also [8]). In the case of J-fractions it is also known that we may write the eigenvalue equation $Ay = zy$ for some infinite column vector y in terms of normalized counterparts of the monic polynomials $q_n(z)$ (namely the corresponding orthonormal OP). Notice that the product Ay is defined for y not necessarily an element of ℓ^2, since for each component there are only a finite number of non-zero terms. For the linear pencil $zB - A$ we can analogously write the similar eigenvalue equations

$$Aq^R(z) = zBq^R(z), \quad q^L(z)A = zq^L(z)B, \tag{2.10}$$

with an infinite column vector $q^R(z) = (q_0^R(z), q_1^R(z), \dots)^\top$ and an infinite row vector $q^L(z) = (q_0^L(z), q_1^L(z), \dots)$. Here $q_n^L(z)$ and $q_n^R(z)$ are rational functions obtained from

$q_n(z)$ by scaling with a product of linear polynomials. Indeed, defining $q_n^R(z)$, $p_n^R(z)$, and $r_n^R(z)$ via

$$q_n^R(z) = \frac{q_n(z)}{\prod_{k=0}^{n-1} \alpha_k^R(z)}, \quad p_n^R(z) = \frac{p_n(z)}{\prod_{k=0}^{n-1} \alpha_k^R(z)}, \quad r_n^R(z) = \frac{r_n(z)}{\prod_{k=0}^{n-1} \alpha_k^R(z)}, \tag{2.11}$$

leads us to three solutions of the recurrence relation

$$\alpha_n^R(z) y_{n+1}^R - \beta_n(z) y_n^R + \alpha_{n-1}^L(z) y_{n-1}^R = 0, \quad n = 0, 1, 2, \ldots. \tag{2.12}$$

In the similar way, we see that

$$q_n^L(z) = \frac{q_n(z)}{\prod_{k=0}^{n-1} \alpha_k^L(z)}, \quad p_n^L(z) = \frac{p_n(z)}{\prod_{k=0}^{n-1} \alpha_k^L(z)}, \quad r_n^L(z) = \frac{r_n(z)}{\prod_{k=0}^{n-1} \alpha_k^L(z)}, \tag{2.13}$$

are three solutions of the recurrence relation

$$\alpha_n^L(z) y_{n+1}^L - \beta_n(z) y_n^L + \alpha_{n-1}^R(z) y_{n-1}^L = 0, \quad n = 0, 1, 2, \ldots. \tag{2.14}$$

Now, it is immediate to see by taking into account the initial conditions (2.3) that the identities (2.12) and (2.14) reduce to the formal spectral equations (2.10).

It should be also noted that we formally have

$$p^L(z)(zB - A) = -e_0^\top, \quad (zB - A)p^R(z) = -e_0. \tag{2.15}$$

Remark 2.1. There are many degrees of freedom in going from a continued fraction (2.1) to a linear pencil $zB - A$. For instance, for the special case of $\deg \beta_n = 1$ and $\deg \alpha_n^L = 0 = \deg \alpha_n^R$ for all $n \geq 0$, the above approach leads a priori to diagonal B and tridiagonal A without any further symmetry properties. However, by applying an equivalence transformation to (2.1) we can make the polynomials β_n monic, implying that B is the identity matrix. Moreover, we can choose $\alpha_n^L = \alpha_n^R$, i.e., A becomes complex symmetric (also called a complex Jacobi matrix). In this case, $q_n^L = q_n^R$ are known to be the corresponding formal orthonormal polynomials, whereas q_n is the associated monic counterpart. We will return to this scaling and normalization freedom in the last section. □

2.2. m**-functions of linear pencils and the resolvent.** In accordance with the Jacobi case of B being the identity, we define the resolvent set $\rho(A, B)$ of the linear pencil $zB - A$ to be the set of $z \in \mathbb{C}$ such that $zB - A$ has a bounded inverse.

Aptekarev et al. [3, Theorem 1] showed that a bounded tridiagonal matrix has a bounded inverse if and only if the above solutions of the recurrencies (2.12) and (2.14) have a particular asymptotic behavior. In our setting, their findings (see also the slight improvement given in [11, Theorem 2.1]) read as follows.

Theorem 2.2 ([3]). *Suppose that A, B are bounded, and consider for $z \in \mathbb{C}$ the matrix $R(z)$ with entries*

$$R(z)_{j,k} = \begin{cases} r_j^R(z) q_k^L(z) = (q_j^R(z)\phi(x) - p_j^R(z)) q_k^L(z) & \text{if } j \geq k, \\ q_j^R(z) r_k^L(z) = q_j^R(z)(q_k^L(z)\phi(x) - p_k^L(z)) & \text{if } j \leq k. \end{cases}$$

Then $z \in \rho(A, B)$ if and only if there exists $\phi(z) \in \mathbb{C}$ and constants $\gamma(z) > 0$, $\delta(z) \in (0, 1)$ such that

$$|R(z)_{j,k}| \leq \gamma(z)\delta(z)^{|j-k|}, \quad j, k = 0, 1, \ldots. \tag{2.16}$$

In this case, $R(z)_{j,k} = \langle (zB - A)^{-1} e_k, e_j \rangle$, in particular, $\phi(z)$ is uniquely given by

$$\phi(z) = R(z)_{0,0} = \langle (zB - A)^{-1} e_0, e_0 \rangle.$$

For the sake of completeness, we will give below the ideas of the proof of Theorem 2.2. Let us first discuss some immediate consequences.

Remark 2.3. For the particular case of Jacobi matrices (that is $B = I$), the above formulas for the entries of the resolvent, also referred to as Green's functions, have been known for a long time, see for instance the recent book [26, Section 4.4]. Our linear pencil formalism also includes so-called CMV matrices occurring in the study of orthogonal polynomials on the unit circle, see [26, Section 4.2], here A, B are not only tridiagonal but in addition block-diagonal, with unitary blocks. Again, the formulas for the Green's functions given in [26] are a special case of Theorem 2.2. □

A basic object in Theorem 2.2 and in the rest of the paper is following.

Definition 2.4. *The function*

$$m(z) = \langle (zB - A)^{-1} e_0, e_0 \rangle, \quad z \in \rho(A, B) \tag{2.17}$$

will be called the m-function (or Weyl function) of the linear pencil $zB - A$.

Comparing with (2.5) we are left with the central question whether the m-function $p_n(z)/q_n(z) = \langle (zB_{[0:n-1]} - A_{[0:n-1]})^{-1} e_0, e_0 \rangle$ of the finite pencil $zB_{[0:n-1]} - A_{[0:n-1]}$ converges for $n \to \infty$ to the m-function of the infinte pencil $zB - A$.

We learn from Theorem 2.2 that the linearized errors $r_n^L(z) = R(z)_{0,n} = q_n^L(z) m(z) - p_n^L(z)$ and $r_n^R(z) = R(z)_{n,0} = q_n^R(z) m(z) - p_n^R(z)$ tend to zero with a geometric rate

$$\limsup_{n\to\infty} |r_n^L(z)|^{1/n} < 1, \quad \limsup_{n\to\infty} |r_n^R(z)|^{1/n} < 1, \quad z \in \rho(A, B). \tag{2.18}$$

Following exactly the lines of Aptekarev et al. [3, Theorem 2 and Corollary 3] we obtain the following result on point-wise convergence of a subequence.

Corollary 2.5. *We have for $z \in \rho(A, B)$*

$$\limsup_{n\to\infty} |q_n^L(z)|^{1/n} > 1, \limsup_{n\to\infty} |q_n^R(z)|^{1/n} > 1, \liminf_{n\to\infty} |m(z) - \frac{p_n(z)}{q_n(z)}|^{1/n} < 1.$$

Proof. Using (2.11) and (2.13), the Liouville-Ostrogradsky formula (2.6) takes the following form

$$\alpha_n^L(z) q_{n+1}^L(z) r_n^R(z) - \alpha_n^R(z) q_n^L(z) r_{n+1}^R(z) = 1. \tag{2.19}$$

Since $\sup_n \max\{|\alpha_n^L(z)|, |\alpha_n^R(z)|\} < \infty$ by assumption on A, B, relation (2.18) together with the Cauchy-Schwarz inequality implies that

$$\liminf_{n\to\infty} [|q_n^L(z)|^2 + |q_{n+1}^L(z)|^2]^{1/(2n)} > 1,$$

implying our first claim. The second is established using similar techniques, and the third by writing $m(z) - p_n(z)/q_n(z) = r_n^L(z)/q_n^L(z)$. □

By having a closer look at the proof, we see that we have pointwise convergence for a quite dense subsequence, namely for $p_{n+\epsilon_n}/q_{n+\epsilon_n}$ for $n \geq 0$ with suitable $\epsilon_n \in \{0, 1\}$. We will show in Theorem 3.5 below that this point-wise convergence result can be replaced by a uniform convergence result in neighborhoods of an element of $\rho(A, B)$.

In the remainder of this subsection we present the main lines of the proof of Theorem 2.2. The first step consists in showing that our infinite matrix $R(z)$ is a formal left and right inverse for $zB - A$, compare with [31, Section 60 and Section 61]) for the case of complex Jacobi matrices.

Lemma 2.6. *For any value of $\phi(z)$, the formal matrix products $R(z)(zB - A)$ and $(zB - A)R(z)$ give the identity matrix.*

Proof. We will concentrate on the first identity, the second following along the same lines. Write shorter

$$q_{[0:j]}^L = (q_0^L, \ldots, q_j^L, 0, 0, \ldots)$$

and similarly $p^L_{[0:j]}$ and $r^L_{[0:j]}$ for the row vectors built with the other solutions of the recurrence (2.14). Then

$$p^L_{[0:j]}(z)(zB-A) = -e_0^\top + \alpha_j^L(z)p^L_{j+1}(z)e_j^\top + \alpha_j^R p_j^L(z)e_{j+1}^\top, \tag{2.20}$$

$$q^L_{[0:j]}(z)(zB-A) = \alpha_j^L(z)q^L_{j+1}(z)e_j^\top + \alpha_j^R q_j^L(z)e_{j+1}^\top. \tag{2.21}$$

In view of (2.6), (2.11), and (2.13), one obtains

$$(q_j^R(z)p^L_{[0:j]}(z) - p_j^R(z)q^L_{[0:j]}(z))(zB-A) = e_j^\top - q_j^L(z)e_0^\top.$$

In addition, from (2.10) and 2.15 we have that

$$(q^L(z)\phi(z) - p^L(z))(zB-A) = e_0^\top.$$

A combination of the last two equations shows that, for all $j \geq 0$,

$$(R(z)_{j,0}, R(z)_{j,1}, R(z)_{j,2}, \dots)(zB-A) = e_j^\top,$$

as claimed above. □

Proof of Theorem 2.2. Let $z \in \rho(A,B)$. Then, according to Lemma 2.6, $R(z)$ is indeed the matrix representation of the bounded operator $(zB-A)^{-1}$. We get the decay rate (2.16) of the entries of $R(z)$ from [13, Theorem 2.4] using the fact that $R(z)$ is the inverse of a bounded tridiagonal matrix.

Suppose now that $\phi(z) \in \mathbb{C}$ is such that (2.16) is satisfied. Then, using the same arguments as in [3] we have that $R(z)$ represents a bounded operator in ℓ^2,which by Lemma 2.6 is a left and right inverse of $zB - A$. Hence $z \in \rho(A,B)$. □

Remark 2.7. The essential tool in the proof of Theorem 2.2 was the decay rate (2.16) of entries of the inverse of a bounded tridiagonal matrix. In order to specify the rate of convergence, for instance in Corollary 2.5, it is interesting to quote from [13, Theorem 2.4] possible values of $\gamma(z), \delta(z)$ in terms of the condition number

$$\kappa(z) = \|zB-A\| \, \|(zB-A)^{-1}\| \geq 1$$

being obviously continuous in $\rho(A,B)$, compare with [11, Lemma 3.3],

$$\delta(z) = \sqrt{\frac{\kappa(z)-1}{\kappa(z)+1}}, \quad \gamma(z) = \frac{3\,\|(zB-A)^{-1}\|}{\delta(z)^2} \max\Big\{\kappa(z), \frac{(1+\kappa(z))^2}{2\kappa(z)}\Big\}.$$

□

2.3. **Biorthogonal rational functions and a Favard theorem.** Our explicit formulas for the entries of the resolvent allow for a simple proof of biorthogonality for the denominators q_j^R and q_k^L, and in addition an explicit formula for the linear functional of orthogonality discussed by Ismail and Masson [20]. This generalizes the classical case of $B = I$ and a selfadjoint Jacobi matrix A [1] where it is well-known that, for $j \neq k$,

$$\langle q_j(A)e_0, q_k(A)e_0 \rangle = 0.$$

As a consequence, we obtain a simple proof of the fact that the nth convergent of (2.1) is indeed an $[n-1|n]$th multipoint Padé approximant of the m-function.

In this subsection we denote for $k = 0, 1, 2, \dots$ by z_{2k+1} (and by z_{2k+2}) the root of α_k^L (and of α_k^R, respectively), where we put $z_{2k+1} = \infty$ (and $z_{2k+1} = \infty$) if α_k^L (and α_k^R) is of degree 0. Similar to [20] we suppose for convenience that $z_1, z_2, \dots \in \rho(A,B)$. More precisely, we suppose that there exists a domain Γ_{ext} with compact boundary forming a Jordan curve such that

$$z_1, z_2, \dots \in \Gamma_{ext} \subset \text{Clos}(\Gamma_{ext}) \subset \rho(A,B), \tag{2.22}$$

where Clos$(\cdot)$ denotes the closure. The case $z_k = \infty$ needs special care: notice that $\infty \in \rho(A,B)$ if and only if B has a bounded inverse, in which case we will also suppose that $\infty \in \Gamma_{ext}$. The boundary Γ of Γ_{ext} is orientated such that Γ_{ext} is on the right of Γ, implying that

$$g(z) = \frac{1}{2\pi i}\int_\Gamma \frac{g(\zeta)}{z-\zeta}d\zeta$$

for $z \in \Gamma_{ext}$ and any function g being analytic in $\rho(A,B)$ and, if $\infty \in \rho(A,B)$, vanishing at infinity.

We start by establishing an integral formula for the entries of the resolvent.

Lemma 2.8. *Under the assumption* (2.22)*, we have for* $z \in \Gamma_{ext}$ *and* $j,k = 0,1,2,\ldots$

$$R(z)_{j,k} = \langle (zB-A)^{-1}e_k, e_j\rangle = \frac{1}{2\pi i}\int_\Gamma q_j^R(\zeta)q_k^L(\zeta)\frac{m(\zeta)}{z-\zeta}\,d\zeta.$$

Proof. We will consider only the case $j \geq k$, the case $j < k$ is similar. Both the resolvent and $R_{j,k}$ are analytic in $\rho(A,B)$, and vanishing at infinity provided that $\infty \in \rho(A,B)$. Using the explicit formula for $R(z)_{j,k}$ derived in Theorem 2.2, we get for $z \in \Gamma_{ext}$

$$\begin{aligned} R(z)_{j,k} &= \frac{1}{2\pi i}\int_\Gamma r_j^R(\zeta)q_k^L(\zeta)\frac{d\zeta}{z-\zeta} \\ &= \frac{1}{2\pi i}\int_\Gamma q_j^R(\zeta)q_k^L(\zeta)\frac{m(\zeta)}{z-\zeta}\,d\zeta - \frac{1}{2\pi i}\int_\Gamma p_j^R(\zeta)q_k^L(\zeta)\frac{d\zeta}{z-\zeta}. \end{aligned}$$

It remains to show that the last integral equals zero. Denote by Ω a connected component of $\overline{\mathbb{C}} \setminus \mathrm{Clos}(\Gamma_{ext})$. If Ω is bounded, then, by assumption (2.22), all poles of the rational function $\zeta \mapsto p_j^R(\zeta)q_k^L(\zeta)/(z-\zeta)$ are outside of $\mathrm{Clos}(\Omega)$, and hence the integral over $\partial\Omega$ is zero. If Ω is unbounded, then by the above assumption on Γ_{ext} we may conclude that $\infty \notin \rho(A,B)$, implying that all z_ℓ are finite. It follows from (2.11) and (2.13) that all poles of the rational function $\zeta \mapsto p_j^R(\zeta)q_k^L(\zeta)/(z-\zeta)$ are outside of $\mathrm{Clos}(\Omega)$, and this function does vanish at ∞. Hence again the integral over $\partial\Omega$ is zero. □

We are now prepared to state and to give a new constructive proof of the Favard type Theorems [20, Theorem 2.1 and Theorem 3.5] of Ismail and Masson.

Theorem 2.9. *Under the assumption* (2.22)*, define for* $g \in \mathcal{C}(\Gamma)$ *the linear functional*

$$\mathfrak{S}(g) = \frac{1}{2\pi i}\int_\Gamma g(\zeta)m(\zeta)\,d\zeta,$$

then we have the following biorthogonality relations: for any $n \geq 1$ *and for any polynomial* p *of degree* $< n$ *there holds*

$$\mathfrak{S}\left(q_n^R\frac{p}{\alpha_0^L\alpha_1^L\ldots\alpha_{n-1}^L}\right) = 0, \quad \mathfrak{S}\left(\frac{p}{\alpha_0^R\alpha_1^R\ldots\alpha_{n-1}^R}q_n^L\right) = 0.$$

Proof. We again only show the first relation, the second follows by symmetry. Observe first that $\alpha_{n-1}^L(z_{2n-1}) = 0$ implies that $zB - A$ is upper block-diagonal. Since $z_{2n-1} \in \rho(A,B)$ by (2.22), we obtain for the resolvent $(z_{2n-1}B-A)^{-1}$ the block matrix representation

$$\left[\begin{array}{c|c} (z_{2n-1}B_{[0:n-1]} - A_{[0:n-1]})^{-1} & * \\ \hline 0 & (z_{2n-1}B_{[n:\infty]} - A_{[n:\infty]})^{-1} \end{array}\right].$$

In particular, comparing with (2.4) it follows that $q_k(z_{2k-1}) \neq 0$ for $k = 0,1,\ldots,n-1$ (or $\deg q_k = k$ provided that $z_{2k-1} = \infty$), and

$$R(z_{2n-1})_{n,k} = 0, \quad k = 0,1,\ldots,n-1$$

(or $\lim_{z\to\infty} zR(z)_{n,k} = 0$ in the case $z_{2n-1} = \infty$). The first relation implies that

$$\text{span}\left\{\frac{q_k^L}{\alpha_{n-1}^L} : k = 0,1,\ldots,n-1\right\} = \left\{\frac{p}{\alpha_0^L\alpha_1^L\ldots\alpha_{n-1}^L} : \deg p < n\right\},$$

and the second combined with Lemma 2.8 that

$$\mathfrak{S}\Big(q_n^R \frac{q_k^L}{\alpha_{n-1}^L}\Big) = 0, \quad k = 0, 1, ..., n-1,$$

as claimed in Theorem 2.9. □

Remark 2.10. In the statement of Theorem 2.9, one recovers the m-function as a generating function for the linear functional of orthogonality, since

$$z \mapsto \mathfrak{S}_\zeta \left(\frac{1}{z-\zeta} \right) = m(z), \quad z \in \Gamma_{ext}.$$

Suppose in addition that $\infty \in \rho(A,B)$, and thus B has a bounded inverse. Then Cauchy's theorem gives the normalisation $\mathfrak{S}(1) = m'(\infty) = \langle B^{-1} e_0, e_0 \rangle$, and for $\ell \geq 0$

$$\mathfrak{S}_\zeta(\zeta^\ell) = \langle B^{-1}(AB^{-1})^\ell e_0, e_0 \rangle.$$

Similarly, for $z_k \in \Gamma_{ext}$ and $\ell \geq 0$ we have that

$$\frac{m^{(\ell)}(z_k)}{\ell!} = \mathfrak{S}_\zeta \left(\frac{-1}{(\zeta - z_\ell)^{\ell+1}} \right) = -\langle B^{-1}(AB^{-1} - z_k)^{-1-\ell} e_0, e_0 \rangle.$$

Using a partial fraction decomposition, we obtain for any polynomial p (of degree $< 2n$ if $\infty \notin \rho(A,B)$) the even simpler formula

$$\mathfrak{S}(r) = \langle B^{-1} r(AB^{-1}) e_0, e_0 \rangle, \quad r = \frac{p}{\alpha_0^L \alpha_0^R \ldots \alpha_{n-1}^L \alpha_{n-1}^R}.$$

□

The orthogonality relations of Theorem 2.9 allow now to show in a simple way that the convergents of our continued fraction (2.1) are indeed multipoint Padé approximants.

Corollary 2.11. *Under the assumption* (2.22), *for any $n \geq 0$, the rational function p_n/q_n is an $[n-1|n]$ multipoint Padé approximant of the m-function of the pencil $zB - A$ at the points $z_1, ..., z_{2n}$ counting multiplicities.*

Proof. Relation (2.4) shows that p_n, q_n, are polynomials of degree at most $n-1$, and n, respectively, and from the proof of Theorem 2.9 we know that $q_n(z_{2n+1}) \neq 0$, hence q_n is non-trivial.

The interpolation conditions for a Cauchy transform (or more generally for a generating function of a linear functional) are known to translate to orthogonality relations with varying weights, see for instance [28, Lemma 6.1.2]. Since $r_n^R = (mq_n - p_n)/(\alpha_0^R \ldots \alpha_{n-1}^R)$ is analytic in Γ_{ext} (and vanishes at ∞ if $z \in \rho(A,B)$), we only have to show that $\omega := r_n^R/(\alpha_0^L \ldots \alpha_{n-1}^L)$ is analytic in Γ_{ext}, and, provided that $\infty \in \Gamma_{ext}$, its expansion at ∞ starts with a term z^{-n-1}.

Denote by $\widetilde{z}_1, ..., \widetilde{z}_{\ell(k)}$ the finite points out of $z_1, z_3, ..., z_{2k-1}$ with $k \leq n$. If $\ell(k) \geq 1$, define

$$\widetilde{p}(z) = \frac{\alpha_0^L(z) \ldots \alpha_{n-1}^L(z)}{(z - \widetilde{z}_{\ell(1)}) \ldots (z - \widetilde{z}_{\ell(k)})},$$

a polynomial of degree $< n$. Arguing as in Lemma 2.8 and using the Hermite integral formulas for divided differences we find that

$$[\widetilde{z}_{\ell(1)}, ..., \widetilde{z}_{\ell(k)}] r_n^R = \frac{1}{2\pi} \int_\Gamma \frac{r_n^R(\zeta) \widetilde{p}(\zeta)}{\alpha_0^L(\zeta) \ldots \alpha_{n-1}^L(\zeta)} \, d\zeta = \mathfrak{S} \left(\frac{q_n^R \widetilde{p}}{\alpha_0^L \alpha_1^L \ldots \alpha_{n-1}^L} \right) = 0,$$

where in the last step we have applied the orthogonality relation of Theorem 2.9. Hence ω is indeed analytic in Γ_{ext}. If $\infty \in \Gamma_{ext}$, we find by a similar argument for the expansion of ω at ∞

$$\begin{aligned}\omega(z) &= \frac{1}{2\pi}\int_\Gamma \frac{r_n^R(\zeta)}{\alpha_0^L(\zeta)\dots\alpha_{n-1}^L(\zeta)}\frac{d\zeta}{z-\zeta} \\ &= \mathfrak{S}_\zeta\left(\frac{q_n^R(\zeta)}{\alpha_0^L(\zeta)\dots\alpha_{n-1}^L(\zeta)(z-\zeta)}\right) = \sum_{j=0}^{\infty} z^{-j-1}\mathfrak{S}_\zeta\left(\frac{q_n^R(\zeta)\zeta^j}{\alpha_0^L(\zeta)\dots\alpha_{n-1}^L(\zeta)}\right),\end{aligned}$$

which again by Theorem 2.9 starts with the term z^{-n-1}. □

3. Convergence results for multipoint Padé approximants

The aim of this section is to generalize various convergence results for complex Jacobi matrices to the setting of linear pencils.

3.1. **Numerical ranges of linear pencils.** It is well known that zeros of formal orthogonal polynomials lie in the numerical range of the corresponding tridiagonal operator. Moreover, the corresponding sequence of Padé approximants converges locally uniformly outside the closure of the numerical range [11, Theorem 3.10]. In this section, we generalize this machinery to the case of linear pencils and multipoint Padé approximants.

Let us recall that, for a bounded operator T acting in ℓ^2, it's numerical range is defined by

$$\Theta(T) := \{(Ty, y)_{\ell^2} : \|y\| = 1\} \subset \mathbb{C}$$

Clearly, $\Theta(T)$ is a bounded set. By the Hausdorff theorem we have that the spectrum $\sigma(T)$ of T is a subset of the convex set $\overline{\Theta(T)}$ (for instance, see [23, Section 26]). The following definition generalizes the concept of numerical ranges to the linear pencil case.

Definition 3.1 ([23]). *The set*

$$W(A, B) := \{z \in \mathbb{C} : \langle (zB - A)y, y\rangle_{\ell^2} = 0 \textit{ for some } y \neq 0\}$$

is called a numerical range of the linear pencil $zB - A$.

The following proposition is immediate from Definition 3.1.

Proposition 3.2. *All the zeros of q_n and p_n belong to $W(A, B)$.*

Proof. Let us suppose that ξ is a zero of the polynomial q_n. Thus, according to (2.4), there exists an element $y_\xi \in \mathbb{C}^n$ such that

$$(\xi B_{[0:n-1]} - A_{[0:n-1]})y_\xi = 0, \quad \|y_\xi\| = 1.$$

The latter relation implies $\xi \in W(A_{[0:n-1]}, B_{[0:n-1]}) \subset W(A, B)$. Similarly, we have the inclusion of the zeros of p_j to $W(A, B)$. □

In general, for the bounded operators A and B, the set $W(A, B)$ is neither convex nor bounded. However, it turns out that the condition

$$0 \notin \overline{\Theta(B)} \tag{3.1}$$

implies $\sigma(A, B) \subset \overline{W(A, B)}$ [23, Section 26], as well as the representation

$$W(A, B) = \left\{\frac{\langle Af, f\rangle}{\langle Bf, f\rangle} : \ \|f\| = 1\right\}, \tag{3.2}$$

from which we see the boundedness of $W(A, B)$.

Generalizing [11, Theorem 3.10] for complex Jacobi matrices, we are able to prove a result on locally uniform convergence which in some sense generalizes the Gonchar theorem [17].

Theorem 3.3. *Let* (3.1) *be satisfied. Then the sequence of multipoint Padé approximants $m_{[0:n]} := p_{n+1}/q_{n+1}$ converges to the m-function locally uniformly in $\mathbb{C} \setminus \overline{W(A, B)}$.*

Proof. Denote by $D \subset \mathbb{C} \setminus \overline{W(A,B)}$ a closed set with compact boundary. Setting $d := \inf_{\|f\|=1} |\langle Bf, f\rangle| > 0$, we find for $z \in \partial D$ and $\|f\| = 1$ that

$$\|(zB - A)f\| \geq |\langle Bf, f\rangle|\,|z - \frac{\langle Af, f\rangle}{\langle Bf, f\rangle}| \geq d \,\mathrm{dist}(z, W(A,B)),$$

implying that

$$\max_{z\in\partial D} \|(zB-A)^{-1}\| \leq d_1 := \frac{1}{d} \max_{z\in\partial D} \frac{1}{\mathrm{dist}(z, W(A,B))}.$$

Since $W(A_{[0:n-1]}, B_{[0:n-1]}) \subset W(A,B)$, the same argument can be used to estimate the norm of the resolvent of finite subsections

$$\max_{z\in\partial D} \|(zB_{[0:n]} - A_{[0:n]})^{-1}\| \leq d_1. \tag{3.3}$$

Let ψ be a finite sequence, that is, $\psi = (\psi_1, \dots, \psi_k, 0, 0, \dots)^\top$. Then

$$(zB - A)\psi = (zB_{[0:j]} - A_{[0:j]})\psi = \phi$$

for sufficiently large $j \in \mathbb{Z}_+$ and ϕ is also a finite sequence. Further, one obviously has

$$(zB-A)^{-1}\phi = \lim_{j\to\infty} (zB_{[0:j]} - A_{[0:j]})^{-1}\phi. \tag{3.4}$$

Since $zB - A$ is bounded and boundedly invertible, the set of such ϕ's is dense in ℓ^2 and, therefore, due to (3.3) we have that formula (3.4) is also valid for all $\phi \in \ell^2$ implying the pointwise convergence $m_{[0:j]}(\lambda) \to m(\lambda)$ for any $z \in \mathbb{C}\setminus\overline{W(A,B)}$. Now, the statement of the theorem immediately follows from (3.3) and the Vitali theorem [30, Section 5.21]. □

Notice that the concept of a numerical range is valid for operator-valued functions [23]. Thus the presented approach can be also generalized to linear pencils proposed in [5].

3.2. **Uniform convergence of subsequences in neighborhoods.** We start by improving the pointwise convergence result of Corollary 2.5 generalizing [3, Corollary 3]. It was Ambroladze [2, Corollaries 3 and 4] who first observed that, for real Jacobi matrices, a quite dense subsequence of convergents of (2.1) converges uniformly in a neighborhood of any element of the resolvent set. This result has been generalized in [10, Theorem 4.4] to the setting of complex Jacobi matrices. We follow here the lines of the proof presented in [8, Theorem 4.7] since this allows to deduce in the next subsection a result of convergence in capacity in bounded connected components of $\rho(A,B)$.

A central observation in what follows is the following result which for complex Jacobi matrices may be found in [10, Proposition 2.2].

Proposition 3.4. *The family of rational functions*

$$u_n(z) = \frac{q_n(z)}{q_{n+1}(z)} = \frac{q_n^L(z)}{\alpha_n^L(z) q_{n+1}^L(z)} = \frac{q_n^R(z)}{\alpha_n^R(z) q_{n+1}^R(z)}$$

is normal with respect to chordal metric on $\rho(A,B)$.

Proof. We only have to show that u_n is equicontinuous on the Riemann sphere. By the definition of the chordal metric we find for $x, y \in \rho(A,B)$

$$\chi(u_n(x), u_n(y)) = \frac{\left|\alpha_n^L(x) q_{n+1}^L(x) q_n^R(y) - \alpha_n^R(y) q_n^L(x) q_{n+1}^R(y)\right|}{\left\|[q_n^L(x), \alpha_n^L(x) q_{n+1}^L(x)]\right\| \left\|[q_n^R(y), \alpha_n^R(y) q_{n+1}^R(y)]\right\|}.$$

In order to minorize the denominator, we write shorter as in the proof of Lemma 2.6

$$q_{[0:n]}^L = (q_0^L, \dots, q_n^L, 0, 0, \dots), \quad q_{[0:n]}^R = (q_0^R, \dots, q_n^R, 0, 0, \dots)^\top$$

and observe that

$$q^L_{[0:n]}(x)(xB-A) = \Big[\underbrace{0,...,0}_{n}, \alpha^L_n(x)q^L_{n+1}(x), -\alpha^R_n(x)q^L_n(x), 0, ...\Big],$$

implying that

$$\|q^L_{[0:n]}(x)\|^2 \le \|(xB-A)^{-1}\|^2\,(1+|\alpha^R_n(x)|^2)\,(|q^L_n(x)|^2 + |\alpha^L_n(x)q^L_{n+1}(x)|^2).$$

Similarly,

$$(yB-A)q^R_{[0:n]}(y) = \Big[\underbrace{0,...,0}_{n}, \alpha^R_n(y)q^R_{n+1}(y), -\alpha^L_n(y)q^R_n(y), 0, ...\Big]^\top,$$

implying that

$$\|q^R_{[0:n]}(y)\|^2 \le \|(yB-A)^{-1}\|^2\,(1+|\alpha^L_n(y)|^2)\,(|q^R_n(x)|^2 + |\alpha^R_n(y)q^R_{n+1}(y)|^2).$$

Finally,

$$\begin{aligned}&\Big(\alpha^L_n(x)q^L_{n+1}(x)q^R_n(y) - \alpha^R_n(y)q^L_n(x)q^R_{n+1}(y)\Big)\\ &= q^L_{[0:n]}(x)\big[(xB-A)-(yB-A)\big]q^R_{[0:n]}(y) = (x-y)q^L_{[0:n]}(x)Bq^R_{[0:n]}(y),\end{aligned}$$

and a combination of these findings yields that $\chi(u_n(x), u_n(y))$ is bounded above by $|x-y|$ times a quantity which can be bounded for x, y lying in compact subsets of $\rho(A,B)$. □

We are now prepared to generalize [10, Theorem 4.4] to linear pencils.

Theorem 3.5. *For any $\xi \in \rho(A,B)$ there exists a closed neighborhood $V \subset \rho(A,B)$ and $\epsilon_n \in \{0,1\}$ such that $m_{[0:n-1+\epsilon_n]}$ converges to m uniformly in V.*

Proof. Let $v_n = u_n$ and $\epsilon_n = 0$ if $|u_n(\xi)| < 1$, or elsewhere $v_n = 1/u_n$ and $\epsilon_n = 1$. Then

$$|m(z) - m_{[0:n-1+\epsilon_n]}(z)| = \Big|\frac{r^L_{n+\epsilon_n}(z)}{q^L_{n+\epsilon_n}(z)}\Big| = \Big|\frac{\alpha^L_n(z)^{\epsilon_n} r^L_{n+\epsilon_n}(z)\sqrt{1+|v_n(z)|^2}}{\sqrt{|q^L_n(z)|^2 + |\alpha^L_n(z)q^L_{n+1}(z)|^2}}\Big|.$$

Using the equicontinuity of the u_n (and thus the v_n) established in Proposition 3.4, there exists a neighborhood V of ξ such that $|v_n(z)| \le 2$ for all $z \in V$. Applying the Cauchy-Schwarz inequality to (2.19), we obtain for $z \in V$ the upper bound

$$\begin{aligned}&|m(z) - m_{[0:n-1+\epsilon_n]}(z)| \le\\ &\sqrt{5}\sqrt{|r^L_n(z)|^2 + |\alpha^L_n(z)r^L_{n+1}(z)|^2}\sqrt{|r^R_n(z)|^2 + |\alpha^R_n(z)r^R_{n+1}(z)|^2}\end{aligned}$$

and the right-hand side tends to zero with a geometric rate according to Remark 2.7. □

One may construct examples with $B = I$ and selfadjoint A with the spectrum $\mathbb{C} \setminus \rho(A,B)$ consisting of two intervals being symmetric with respect to the origin $\xi = 0$, and $m_{[0:n-1]}$ has a pole at ξ for all odd n. This shows that we may not expect convergence for a subsequence denser than that of Theorem 3.5.

3.3. **Convergence in capacity.** As explained already before, in general one may not expect convergence of $m_{[0:n]}$ to m locally uniformly in $\rho(A,B)$ since there might be so-called spurious poles in $\rho(A,B)$. One strategy of overcoming the problem of spurious poles is to allow for exceptional small sets, as done in [10, Theorem 3.1] for complex Jacobi matrices where convergence in capacity is established. We may generalize these findings for linear pencils, where again we follow the lines of the alternate proof presented in [8, Theorem 4.7].

Theorem 3.6. *Let V be a closed connected subset of $\rho(A,B)$ with compact boundary, then there exist $\epsilon_n \in \{0,1\}$ such that $m_{[0:n-1+\epsilon_n]}$ converges to m in capacity in V.*

If (3.1) is satisfied and $V \not\subset \overline{W(A,B)}$ then we obtain convergence in capacity of the whole subsequence.

Proof. Let again be $v_n = u_n^{1-2\epsilon_n}$ with $\epsilon_n \in \{0,1\}$ to be fixed later, and consider the sets

$$V_\epsilon := \{z \in V : |v_n(z)| \geq 1/\epsilon\}.$$

The arguments in the proof of Theorem 3.5 show that $m_{[0:n-1+\epsilon_n]}$ converges to m uniformly in $V \setminus V_\epsilon$. It remains thus to show that the capacity of V_ϵ tends to zero for $\epsilon \to 0$.

We choose ϵ_n in order to insure that the normal family (v_n) does not have a partial limit being equal to the constant ∞ in the connected component of $\rho(A,B)$ containing V: this can be done for instance by choosing a fixed $\xi \in V$ and to take ϵ_n as in Theorem 3.5, namely $\epsilon_n = 0$ if $|u_n(\xi)| < 1$, and elsewhere $\epsilon_n = 1$. However, under the assumptions of the second part of the statement, by taking $\xi \in V \setminus \overline{W(A,B)}$ it follows from the proof of Theorem 3.3 that

$$\sup_n |u_n(\xi)| = \sup_n |e_n^\top (\xi B_{[0:n]} - A_{[0:n]})^{-1} e_n| < \infty,$$

and hence here we may take the constant sequence $\epsilon_n = 0$.

It is now a well-known fact on normal families (see for instance [10, Lemma 2.4] or the proof of [8, Theorem 4.7]) that for normal meromorphic families $(v_n)_n$ with partial limits different from ∞ there exist monic polynomials ω_n of degree bounded independent of n such that

$$C := \sup_n \max_{z \in V} |\omega_n(z) v_n(z)| < \infty.$$

This enables us insure that

$$V_\epsilon \subset \{z \in V : \frac{C}{|\omega_n(z)|} \geq 1/\epsilon\} \subset \{z \in \mathbb{C} : |\omega_n(z)| \leq \epsilon C\}.$$

Since the capacity increases for increasing sets, and since the capacity of the right-hand lemniscate can be explicitly computed to be $(\epsilon C)^{1/\deg \omega_n}$, the assertion follows. □

4. Biorthogonal rational functions and bi-diagonal decompositions

In this section we give an operator interpretation of the Darboux transformations of rational solutions of the difference equations in question (for the orthogonal polynomials case see [12]). In other words, we present a scheme for constructing biorthogonal rational functions.

4.1. LU-factorizations.

Let us try to factorize the linear pencil $zB - A$ as follows

$$zB - A = L(z)D(z)U(z) \tag{4.1}$$

where $D(z) = \mathrm{diag}(d_0(z), d_1(z), ...)$ is a diagonal matrix, and L, U are bidiagonal matrices of the forms

$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ -v_0^L & 1 & 0 & \\ 0 & -v_1^L & 1 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -v_0^R & 0 & \cdots \\ 0 & 1 & -v_1^R & \ddots \\ 0 & 0 & \ddots & \ddots \\ \vdots & & \ddots & \ddots \end{pmatrix}.$$

Comparing coefficients gives

$$-\alpha_n^L = -v_n^L d_n, \quad -\alpha_n^R = -v_n^R d_n, \quad d_0 = \beta_0, \quad \beta_n = d_n + \frac{\alpha_{n-1}^L \alpha_{n-1}^R}{d_{n-1}}.$$

Thus $d_0(z) = q_1(z)/q_0(z)$ by (2.3), and by recurrence using (2.2) one deduces that

$$d_n(z) = \frac{q_{n+1}(z)}{q_n(z)}, \quad v_n^L(z) = \frac{\alpha_n^L(z)q_n(z)}{q_{n+1}(z)}, \quad v_n^R(z) = \frac{\alpha_n^R(z)q_n(z)}{q_{n+1}(z)}.$$

Hence, the decomposition (4.1) exists if and only iff $q_n(z) \neq 0$ for all $n \geq 0$. In particular, from Proposition 3.2 we obtain existence of such a factorization for $z \notin W(A, B)$.

The decomposition (4.1) gives us the possibility to define Christoffel type transformations.

Proposition 4.1. *Under assumption* (2.22)*, let* $x_0 \in \Gamma_{ext}$ *such that the decomposition (4.1) exists for* $z = x_0$*. Define for* $n \geq 0$ *the functions rational in* x

$$Q_n^L(x_0, x) = \frac{q_n^L(x) - v_n^L(x_0)q_{n+1}^L(x)}{x_0 - x}, \; Q_n^R(x_0, x) = \frac{q_n^R(x) - v_n^R(x_0)q_{n+1}^R(x)}{x_0 - x}.$$

Then we have the orthogonality relations

$$\frac{1}{2\pi i} \int_\Gamma Q_j^L(x_0, x)Q_k^R(x_0, x)(x_0 - x)m(x)dx = \delta_{j,k}/d_j(x_0), \tag{4.2}$$

where $\delta_{j,k}$ *is the Kronecker delta.*

Proof. Denote by $I_{j,k}$ the expression on the left-hand side of (4.2). We only consider the case $0 \leq j \leq k$, the other case follows by symmetry. By definition of $\mathcal{Q}_j^L(x_0, x)$ and $\mathcal{Q}_k^R(x_0, x)$ and by Lemma 2.8 we obtain

$$\begin{aligned} I_{jk} \;\; = \;\; & R(x_0)_{k,j} - v_k^R(x_0)R(x_0)_{k+1,j} \\ & - v_j^L(x_0)R(x_0)_{k,j+1} + v_j^L(x_0)v_k^R(x_0)R(x_0)_{k+1,j+1}. \end{aligned}$$

For $j < k$, we may apply Theorem 2.2 and obtain after factorization

$$I_{jk} = (q_j^L(x_0) - v_j^L(x_0)q_{j+1}^L(x_0))(r_k^R(x_0) - v_k^R(x_0)r_{k+1}^R(x_0)) = 0$$

by definition of $v_j^L(x_0)$. If $j = k$, we get slightly different formulas from Theorem 2.2 and obtain after some simplifications

$$I_{jj} = q_j^L(x_0)(r_j^R(x_0) - v_j^R(x_0)r_{j+1}^R(x_0)) = \frac{q_j(x_0)}{q_{j+1}(x_0)} = \frac{1}{d_j(x_0)},$$

where in the second equality we have applied (2.19). □

Remark 4.2. Clearly, the functions $\alpha_0^L \dots \alpha_n^L Q_n^L(x_0, \cdot)$ and $\alpha_0^R \dots \alpha_n^R Q_n^R(x_0, \cdot)$ are polynomials of degree $\leq n$. □

Proposition 4.1 tells us that the Christoffel transformation leads to multiplication of the biorthogonality measure $m(x)$ by a linear factor $(x_0 - x)$. This process can be repeated. Indeed, after the Christoffel transformation we again obtain a pair of biorthogonal rational functions satisfying a generalized eigenvalue equation with a new pair of the Jacobi matrices $\tilde{A}$, $\tilde{B}$ [32]. We can thus apply the Christoffel transformation to these new functions factorizing the linear pencil $x_1\tilde{B} - \tilde{A}$ in a similar way as in (4.1). Then the weight function $m(x)(x_0 - x)$ is multiplied by a linear factor $x_1 - x$ with $x_1 \neq x_0$. Repeating this process, let us introduce the polynomial $\pi_N(x) = (x_0 - x)(x_1 - x)\dots(x_{N-1} - x)$ with $x_i \neq x_j$, for $i \neq j$ and construct the functions

$$Q_n^L(x_0, x_1, \dots, x_{N-1}; x) = \frac{A_{n,N}^L(x)}{\pi_N(x)B_{n,N}}$$

where

$$A_{n,N}^L(x) = \det \begin{bmatrix} q_n^L(x) & q_{n+1}^L(x) & \dots & q_{n+N}^L(x) \\ q_n^L(x_0) & q_{n+1}^L(x_0) & \dots & q_{n+N}^L(x_0) \\ \dots & \dots & \dots & \dots \\ q_n^L(x_{N-1}) & q_{n+1}^L(x_{N-1}) & \dots & q_{n+N}^L(x_{N-1}) \end{bmatrix}, \quad (4.3)$$

$$B_{n,N}^L = \det \begin{bmatrix} q_{n+1}^L(x_0) & q_{n+2}^L(x_0) & \dots & q_{n+N}^L(x_0) \\ \dots & \dots & \dots & \dots \\ q_{n+1}^L(x_{N-1}) & q_{n+2}^L(x_{N-1}) & \dots & q_{n+N}^L(x_{N-1}) \end{bmatrix}, \quad (4.4)$$

and similar expressions for $Q_n^R(x_0, ..., x_{n-1}; x)$, $A_{n,N}^R(x)$ and $B_{n,N}^R(x)$. Note that if two or more of the parameters x_i coincide, say $x_1 = x_0$, then we may apply a simple limiting process leading to appearance of derivatives in corresponding determinants. Then it is easy to show that these functions satisfy the biorthogonality relation

$$\begin{aligned} &\frac{1}{2\pi i}\int_\Gamma Q_j^L(x_0, x_1, \dots, x_{N-1}; x) Q_k^R(x_0, x_1, \dots, x_{N-1}; x)\pi_j(x)m(x)dx \\ &= \delta_{j,k}/d_j(x_0, x_1, \dots, x_{j-1}), \end{aligned} \quad (4.5)$$

with some constants $d_j(x_0, x_1, \dots, x_{j-1})$. Formula (4.5) is a direct generalization of the Christoffel formula for the orthogonal polynomials, see, e.g., [29, §2.5].

4.2. UL**-decomposition.** For $z \in \rho(A, B)$, let us find a decomposition

$$zB - A = U(z)D(z)L(z) \quad (4.6)$$

with a diagonal matrix $D(z) = \text{diag}(d_0(z), d_1(z), ...)$, and bidiagonal matrices

$$U = \begin{pmatrix} 1 & -u_0^R & 0 & \cdots \\ 0 & 1 & -u_1^R & \ddots \\ 0 & 0 & \ddots & \ddots \\ \vdots & & \ddots & \ddots \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ -u_0^L & 1 & 0 & \\ 0 & -u_1^L & 1 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

By comparing coefficients we have

$$-\alpha_n^L = -u_n^L d_{n+1}, \; -\alpha_n^R = -u_n^R d_{n+1}, \; \beta_n = d_n + u_n^L u_n^R d_{n+1} = d_n + \frac{\alpha_n^L \alpha_n^R}{d_{n+1}}.$$

It turns out that this decomposition is unique after fixing an arbitrary value for d_0. Indeed, let $y_{-1} = d_0, y_0 = 1$, and consider y_n defined by the recurrence relation (2.2). Then it follows that

$$d_n = \frac{\alpha_{n-1}^L \alpha_{n-1}^R y_{n-1}}{y_n}, \quad u_n^L = \frac{y_{n+1}}{\alpha_n^R y_n}, \quad u_n^R = \frac{y_{n+1}}{\alpha_n^L y_n},$$

where from (2.3) we learn that

$$y_n(z) = (1 - m(z)d_0(z))q_n(z) + d_0(z)r_n(z) = q_n(z) - d_0(z)p_n(z) \quad (4.7)$$

for all $n \geq -1$. Thus the decomposition (4.6) exists if and only if $d_0(z) \in \mathbb{C}$ is chosen such that $y_n(z) \neq 0$ and $\alpha_n^L(z)\alpha_n^R(z) \neq 0$ for all $n \geq 0$. For the special case $d_0(z) = 0$ we may compare with the LU decomposition of the preceding subsection and get $u_n^R = 1/v_n^L$ and similarly $u_n^L = 1/v_n^R$. Also, for the special case $m(z)d_0(z) = 1$ one may show that $y_n(z) = d_0(z)r_n(z) \neq 0$ provided that $z \notin W(A, B)$.

Suppose that the above factorization exists for $z = x_0$, and define the Geronimus type transformations by the following formulas

$$\mathcal{Q}_n^L(x_0, x) = q_n^L(x) - u_{n-1}^R(x_0)q_{n-1}^L(x), \; \mathcal{Q}_n^R(x_0, x) = q_n^R(x) - u_{n-1}^L(x_0)q_{n-1}^R(x),$$

and $\mathcal{Q}_0^L(x_0, x) = \mathcal{Q}_0^R(x_0, x) = 1$.

Proposition 4.3. *Under assumption* (2.22)*, let $x_0 \in \Gamma_{ext}$, $d_0(x_0) \neq 0$, such that the above factorization* (4.6) *exists for $z = x_0$. Consider for $g \in \mathcal{C}(\Gamma)$ the linear functional*

$$\widetilde{\mathfrak{S}}(g) = \frac{1}{2\pi i}\int_\Gamma g(\zeta)\frac{m(\zeta)}{x_0 - \zeta}\,d\zeta + \Big(\frac{1}{d_0(x_0)} - m(x_0)\Big)g(x_0),$$

then we obtain for $j \neq k$ the biorthogonality relations

$$\widetilde{\mathfrak{S}}\Big(\mathcal{Q}_j^R(\cdot,x_0)\mathcal{Q}_k^L(\cdot,x_0)\Big) = 0, \quad \widetilde{\mathfrak{S}}\Big(\mathcal{Q}_j^R(\cdot,x_0)\mathcal{Q}_j^L(\cdot,x_0)\Big) = \frac{1}{d_j(x_0)}. \tag{4.8}$$

Proof. We only look at the case $j \geq k \geq 0$, the other case follows by symmetry. Let us compute the (j,k)th entry of the product $L(x_0)(x_0B - A)^{-1}$ (which formally is perhaps expected to be equal to the upper triangular matrix $D(x_0)^{-1}U(x_0)^{-1}$ but turns out to be a full matrix). Using Lemma 2.8 observing that $x_0 \in \Gamma_{ext}$ we get for $j > 0$,

$$\begin{aligned} & \langle L(x_0)(x_0B-A)^{-1}e_k, e_j\rangle \\ = \;& \langle (x_0B-A)^{-1}e_k, e_j\rangle - u_{j-1}^L(x_0)\langle (x_0B-A)^{-1}e_k, e_{j-1}\rangle \\ = \;& \frac{1}{2\pi i}\int_\Gamma \mathcal{Q}_j^R(\zeta,x_0)q_k^L(\zeta)\frac{m(\zeta)}{x_0-\zeta}d\zeta, \end{aligned}$$

and we observe that the same conclusion is true for $j = 0$. If now $j > k$, we may rewrite the last expression as

$$\langle L(x_0)(x_0B-A)^{-1}e_k, e_j\rangle = \Big(r_j^R(x_0) - u_{j-1}^L(x_0)r_{j-1}^R(x_0)\Big)q_k^L(x_0).$$

Noticing that $u_{j-1}^L(x_0) = y_j^R(x_0)/y_{j-1}^R(x_0)$ with $y_n^R = y_n/(\alpha_0^R...\alpha_{n-1}^R)$, we get according to (4.7)

$$\begin{aligned} & r_j^R(x_0) - u_{j-1}^L(x_0)r_{j-1}^R(x_0) \\ = \;& r_j^R(x_0) - \frac{y_j^R(x_0)}{d_0(x_0)} - u_{j-1}^L(x_0)\left(r_{j-1}^R(x_0) - \frac{y_{j-1}^R(x_0)}{d_0(x_0)}\right) \\ = \;& \Big(m(x_0) - \frac{1}{d_0(x_0)}\Big)\mathcal{Q}_j^R(x_0,x_0). \end{aligned}$$

Thus for all $g \in \text{span}\{q_k^L : k = 0,...,j-1\} = \text{span}\{\mathcal{Q}_k^L(\cdot,x_0) : k = 0,...,j-1\}$ we conclude that $\widetilde{\mathfrak{S}}(\mathcal{Q}_j^R(\cdot,x_0)g) = 0$, and, by definition of $\widetilde{\mathfrak{S}}$ and Theorem 2.2,

$$\begin{aligned} & \widetilde{\mathfrak{S}}(\mathcal{Q}_j^R(\cdot,x_0)\mathcal{Q}_j^L(\cdot,x_0)) = \widetilde{\mathfrak{S}}(\mathcal{Q}_j^R(\cdot,x_0)q_j^L) \\ = \;& \langle L(x_0)(x_0B-A)^{-1}e_j, e_j\rangle + \Big(\frac{1}{d_0(x_0)} - m(x_0)\Big)\mathcal{Q}_j^R(x_0,x_0)q_j^L(x_0) \\ = \;& \mathcal{Q}_j^R(x_0,x_0)\Big(r_j^L(x_0) + (\frac{1}{d_0(x_0)} - m(x_0))q_j^L(x_0)\Big) \\ = \;& \mathcal{Q}_j^R(x_0,x_0)\frac{y_j(x_0)}{d_0(x_0)\alpha_0^L(x_0)...\alpha_{j-1}^L(x_0)} = \frac{1}{d_j(x_0)}, \end{aligned}$$

the last claim being evident for $j = 0$, and for $j > 0$ according to (2.6) and (4.7)

$$\begin{aligned} \frac{\mathcal{Q}_j^R(x_0,x_0)}{d_0(x_0)\alpha_0^L(x_0)...\alpha_{j-1}^L(x_0)} &= \frac{p_jq_{j-1} - p_{j-1}q_j}{y_{j-1}\alpha_0^L...\alpha_{j-1}^L\alpha_0^R...\alpha_{j-1}^R} \\ &= \frac{1}{y_{j-1}\alpha_{j-1}^L\alpha_{j-1}^R} = \frac{y_j(x_0)}{d_j(x_0)} \end{aligned}$$

where for simplicity we have dropped in the intermediate expression the argument x_0. $\square$

Remark 4.4. Formula (4.8) means that the (bi)orthogonality measure $\tilde{m}(x)$ for the transformed rational functions $\mathcal{Q}_j^L(x_0, x), \mathcal{Q}_k^R(x_0, x)$ consists of a regular part $m(x)/(x_0 - x)$ on Γ plus a point mass at $x = x_0$, with mass $M_0 = 2\pi i(1/d_0(x_0) - m(x_0))$, where $d_0(x_0) \neq 0$ is a free parameter. A similar situation occurs in the case of ordinary orthogonal polynomials, where the mass of the point mass in the Geronimus transformation can be freely chosen [12]. □

Remark 4.5. Proposition 4.3 for $x_0 \to \infty$ (after multiplication with x_0) has been considered before in [16, Theorem 2.2]. □

5. An example

In order to illustrate the above findings and to give a non-trivial example, we study in this section the properties of a symmetric linear pencils related to a Markov function of the form

$$\varphi(z) = \int_a^b \frac{d\mu(t)}{z - t},$$

with a probability measure μ with support included in some compact real interval $[a, b]$. Here, the entries $A_{j,k}, B_{j,k}$ of the linear pencil $zB - A$ for symmetric interpolation points

$$z_1 = \overline{z_2}, z_3 = \overline{z_4}, \ldots \in \mathbb{C} \setminus [a, b] \tag{5.1}$$

are obtained by developing φ into an even part of a Thiele continued fraction. Before going into details, we recall from the beginning of §1 the special case of interpolation at infinity $z_1 = z_2 = z_3 = \ldots = \infty$. Here the expansion of φ into a J-fraction generates a pencil $zB - A$ with a real Jacobi matrix A and with $B = I$ the identity, and it is known that the spectrum of the linear pencil $zB - A$ (and thus of A) is given by the support of the underlying measure μ, and the numerical range equals to its convex hull $[a, b]$. The aim of this section is to show that these properties remain valid for more general sets of interpolation points.

Returning to the task of developing φ into the continued fraction in question, the following result has been shown in [16, Lemma 3.1 and Remark 3.3], by making the link with Nevalinna functions. The proof given in [16] uses the assumption $|\operatorname{Im} z_j| \geq \delta > 0$ and it can be immediately generalized to our setting.

Proposition 5.1. *Suppose that* (5.1) *holds, and that μ has an infinite number of points of increase such that φ is not a rational function. Then there exist probability measures $\mu_0 = \mu, \mu_1, \mu_2, \ldots$ such that, for all $j \geq 0$,*

$$\varphi_j(z) = \frac{1}{zB_{j,j} - A_{j,j} - B_{j+1,j}^2 (z - z_{2j+1})(z - z_{2j+2})\varphi_{j+1}(z)}$$

with the Markov functions

$$\varphi_j(z) = \int_a^b \frac{d\mu_j(t)}{z - t},$$

and the real numbers

$$B_{j,j} = \frac{\int_a^b \frac{d\mu_j(t)}{|z_{2j+1} - t|^2}}{\left|\int_a^b \frac{d\mu_j(t)}{z_{2j+1} - t}\right|^2} > 1, \; A_{j,j} = \frac{\int_a^b \frac{t\, d\mu_j(t)}{|z_{2j+1} - t|^2}}{\left|\int_a^b \frac{d\mu_j(t)}{z_{2j+1} - t}\right|^2}, \; B_{j+1,j} = \sqrt{B_{j,j} - 1} > 0.$$

Hence, our Markov function φ for the symmetric interpolation points (5.1) induces a linear tridiagonal pencil $zB - A$ if we set according to (2.9),

$$\begin{aligned}
\beta_j(z) &= zB_{j,j} - A_{j,j}, \\
-\alpha_j^L(z) &= zB_{j+1,j} - A_{j+1,j} = B_{j+1,j}(z - z_{2j+1}), \\
-\alpha_j^R(z) &= zB_{j,j+1} - A_{j,j+1} = B_{j+1,j}(z - \overline{z_{2j+1}}) = B_{j,j+1}(z - z_{2j+2}).
\end{aligned}$$

We collect some elementary properties of this pencil in the following two propositions.

Proposition 5.2. *Suppose that* (5.1) *holds. Then the above tridiagonal matrices A, B are Hermitian and bounded.*

Proof. It follows from (5.1) and the explicit formulas given in Proposition 5.1 that A and B are hermitian, and B is real. In order to show that B is bounded, it is sufficient to show that its entries are uniformly bounded, where in our case it is sufficient to consider the diagonal ones. Let us first establish the minorization

$$\left|\int_a^b \frac{d\mu_j(t)}{z-t}\right| \geq \frac{\mathrm{dist}(z,[a,b])}{\max\{|z-a|^2,|z-b|^2\}}, \quad z\in\mathbb{C}\setminus[a,b]. \tag{5.2}$$

For a proof of (5.2) we suppose that $\mathrm{Re}\, z \geq (a+b)/2$, the other case is similar. Since $t\mapsto \mathrm{Im}(1/(z-t))$ does not change sign on $[a,b]$, we get

$$\left|\mathrm{Im}\int_a^b \frac{d\mu_j(t)}{z-t}\right| = \int_a^b \left|\mathrm{Im}\,\frac{1}{z-t}\right| d\mu_j(t) \geq \frac{|\,\mathrm{Im}\, z|}{|z-a|^2}.$$

Hence our claim (5.2) follows provided that $|\,\mathrm{Im}\, z| = \mathrm{dist}(z,[a,b])$. Otherwise, we have that $\mathrm{Re}(z-t)\geq \mathrm{Re}(z-b)>0$ for all $t\in[a,b]$, and hence

$$|\mathrm{Re}\,\varphi_j(z)| \geq \frac{\mathrm{Re}(z-b)}{|z-a|^2},$$

and the claim follows by observing that $|z-b| = \mathrm{dist}(z_{2j+1},[a,b])$.

Combining (5.2) with the definition of $B_{j,j}$ given in Proposition 5.1 we conclude that

$$B_{j,j} \leq \frac{\max(|z_{2j+1}-a|^4, |z_{2j+1}-b|^4)}{\mathrm{dist}(z_{2j+1},[a,b])^4}$$

the right-hand side being bounded according to assumption (5.1). Thus B is bounded.

Similarly, one shows that the diagonal entries $A_{j,j}$ of A are uniformly bounded. In order to discuss the off-diagonal entries of A, we choose a fixed point $z\in\mathbb{C}\setminus[a,b]$ having a positive distance from the set of the interpolation points z_j, and get with the help of Proposition 5.1

$$\begin{aligned}|A_{j+1,j}|^2 = |A_{j,j+1}|^2 &= |z_{2j+1}|^2 B_{j+1,j}^2 \\ &\leq \frac{1}{\varphi_{j+1}(z)}\frac{|z_{2j+1}|^2}{|z-z_{2j+1}|^2}\left(|zB_{j,j}-A_{j,j}| + \frac{1}{|\varphi_j(z)|}\right)\end{aligned}$$

the right-hand side being bounded uniformly for $j\geq 0$ according to (5.2). Hence, A is also bounded. $\square$

Proposition 5.3. *Suppose that* (5.1) *holds. Then for all $y=(y_0,y_1,...)^\top\in\ell^2$ there holds*

$$\langle By,y\rangle \geq |y_k|^2 \quad \textit{if } y_0=...=y_{k-1}=0. \tag{5.3}$$

Furthermore, for the numerical range of Definition 3.1 there holds $W(A,B)\subset[a,b]$.

Proof. In order to show (5.3), let $y=(y_0,y_1,y_2,...)^\top\in\ell^2$. We write as before $y_{[0:n]} = (y_0,y_1,...,y_n,0,0,...)^\top\in\ell^2$, and notice that $\langle By_{[0:n]},y_{[0:n]}\rangle\to\langle By,y\rangle$ for $n\to\infty$ since B is bounded by Proposition 5.2. Then for $y_0=...=y_{k-1}=0$ and $n\geq k$ using the relation $B_{j,j}=1+B_{j+1,j}^2$ we get

$$\begin{aligned}\langle By_{[0:n]},y_{[0:n]}\rangle &= \sum_{j=k}^n B_{j,j}|y_j|^2 + 2\sum_{j=k}^{n-1} B_{j+1,j}\,\mathrm{Re}(y_j\overline{y_{j+1}}) \\ &= |y_k|^2 + \sum_{j=k}^{n-1}|B_{j+1,j}y_j + y_{j+1}|^2 + B_{n+1,n}^2|y_n|^2 \geq |y_k|^2,\end{aligned}$$

implying that (5.3) holds. Since also $\langle Ay_{[0:n]}, y_{[0:n]}\rangle \to \langle Ay, y\rangle$ for $n \to \infty$, we get using (5.3) that $W(A, B)$ is included in the closure of the union of the numerical ranges $W(A_{[0:n]}, B_{[0:n]})$ of all finite sections. Hence it is sufficient to show that $A_{[0:n]} - aB_{[0:n]}$ and $bB_{[0:n]} - A_{[0:n]}$ are positive definite for all $n \geq 0$, or, the determinants of these matrices are positive. From (2.4) we know that

$$q_{n+1}(b) = \det(bB_{[0:n]} - A_{[0:n]}), \quad (-1)^{n+1} q_{n+1}(a) = \det(A_{[0:n]} - aB_{[0:n]}),$$

moreover, q_{n+1} has the leading coefficient $\det(B_{[0:n]})$ which is > 0 since $B_{[0:n]}$ is positive definite by (5.3). Thus the positivity of both determinants inclusion $W(A_{[0:n]}, B_{[0:n]}) \subset [a, b]$ follows from the observation that all $n + 1$ roots of q_{n+1} , that is, the poles of a rational interpolant of a Markov function are lying in the open interval (a, b), see [28, Lemma 6.1.2]. □

The positive definiteness of finite sections of B also for not necessarily bounded $[a, b]$ has been shown already in [16, Proposition 4.2], where the authors also establish (5.3).

Notice that property (5.3) in general does not imply that condition (3.1) is true. However, only the latter condition allows us to conclude that the spectrum of the pencil $zB - A$ is included in $[a, b]$. There is a special case where we may say more.

Theorem 5.4. *Beside* (5.1)*, suppose in addition that* $z_{2j+1} = \overline{z_{2j+2}} \to \infty$ *as* $j \to \infty$. *Then the operator* B *is a compact perturbation of the identity, and condition* (3.1) *holds.*

In particular, the spectrum of $zB - A$ *is given by the support of the measure* μ*, and, outside the spectrum,* φ *coincides with the* m*–function of the linear pencil* $zB - A$*.*

Proof. We have shown in the proof of Proposition 5.2 that $|A_{j+1,j}|^2 = |z_{2j+1}|^2(B_{j,j} - 1) = |z_{2j+1}|^2 B_{j+1,j}^2$ is bounded for $j \to \infty$, and hence

$$\lim_{j\to\infty} B_{j+1,j} = \lim_{j\to\infty} B_{j,j+1} = 0, \quad \lim_{j\to\infty} B_{j,j} = 1,$$

showing that B is a compact perturbation of the identity, and B has its numerical range included in $[0, +\infty)$ by (5.3). Hence, if (3.1) does not hold, then 0 would be an eigenvalue of B, with corresponding eigenvector $y \in \ell^2, y \neq 0$. Inserting this y into (5.3) with k such that $y_k \neq 0$ gives a contradiction.

It follows from the text after (3.1) together with Proposition 5.3 that the spectrum $\sigma(A, B)$ of the linear pencil $zB - A$ is included in $[a, b]$. Also, by construction and Corollary 2.11, p_n/q_n interpolates both φ and m in z_{2n}, implying that these functions are equal for $z = z_{2n}$ and for all n, and analytic in $\mathbb{C} \setminus [a, b]$ including ∞. Since these points accumulate at ∞, we conclude that $m = \varphi$ outside $[a, b]$. Finally, the inclusion supp$(\mu) \subset \sigma(A, B)$ follows from the fact that φ is not analytic in any domain containing points of the support of μ.

Given $z \in \mathbb{C} \setminus [a, b]$, by choosing a contour Γ surrounding $[a, b]$ but not the interpolation points z_j nor z we get from Lemma 2.8 the formula

$$\langle (zB - A)^{-1} e_k, e_j\rangle = \frac{1}{2\pi i} \int_\Gamma q_j^R(\zeta) q_k^L(\zeta) \frac{m(\zeta)}{z - \zeta} d\zeta = \int_a^b q_j^R(t) q_k^L(t) \frac{d\mu(t)}{z - t},$$

where for the second identity we have used Fubini and the fact that $\varphi = m$ on Γ. We denote by $R(z)$ the infinite matrix with entries $R(z)_{j,k}$ given by the above right-hand integral, which is clearly well defined for any z outside the support of μ. From Lemma 2.6 we know that $R(z)$ is a formal left and right inverse of $(zB - A)$, and the desired conclusion $z \notin \sigma(A, B)$ follows as in the proof of Theorem 2.2 by showing that $R(z)$ is bounded.

For this last step, we consider the UL decomposition of B discussed in Remark 4.5 and in [16, Theorem 2.2]: let U be an upper bidiagonal matrix with ones on the diagonal, and the quantities $B_{j+1,j}$ on the main upper diagonal, then U represents a bounded operator on ℓ^2 according to Proposition 5.2. Moreover, we have that $B = UU^*$, and, with B, also U has a bounded inverse. Hence it will be sufficient to show that

$$|\langle U^* R(z) U y, y\rangle| \leq \frac{\langle y, y\rangle}{\mathrm{dist}(z, \mathrm{supp}(\mu))} \tag{5.4}$$

for all $y = (y_0, y_1, ..., y_n, 0, 0, ...)^\top \in \ell^2$ and for all n. Comparing with Proposition 4.3 we find that

$$\langle U^* R(z) U e_k, e_j\rangle = \int_a^b \mathcal{Q}_j^R(\infty, t)\mathcal{Q}_k^L(\infty, t)\frac{d\mu(t)}{z-t}$$

where

$$\mathcal{Q}_n^L(\infty, x) = q_n^L(x) - B_{n,n-1}q_{n-1}^L(x),\ \mathcal{Q}_n^R(\infty, x) = q_n^R(x) - B_{n,n-1}q_{n-1}^R(x),$$

and $\mathcal{Q}_0^L(\infty, x) = \mathcal{Q}_0^R(\infty, x) = 1$, and finally

$$\frac{1}{2\pi i}\int_\Gamma \mathcal{Q}_j^R(\infty, \zeta)\mathcal{Q}_k^L(\infty, \zeta) m(\zeta) d\zeta = \int_a^b \mathcal{Q}_j^R(\infty, t)\mathcal{Q}_k^L(\infty, t)\, d\mu(t) = \delta_{j,k}.$$

In addition, since q_n has real coefficients, it also follows from (5.1) that, for $t \in \mathbb{R}$,

$$q_j^R(t) = \overline{q_j^L(t)}, \quad \mathcal{Q}_j^R(\infty, t) = \overline{\mathcal{Q}_j^L(\infty, t)},$$

implying that

$$\begin{aligned} |\langle U^* R(z) U y, y\rangle| &= \left| \int_a^b \left| \sum_{j=0}^n y_j \mathcal{Q}_j^L(\infty, t)\right|^2 \frac{d\mu(t)}{z-t} \right| \\ &\leq \frac{1}{\mathrm{dist}(z, \mathrm{supp}(\mu))}\int_a^b \left| \sum_{j=0}^n y_j \mathcal{Q}_j^L(\infty, t)\right|^2 d\mu(t) = \frac{\langle y, y\rangle}{\mathrm{dist}(z, \mathrm{supp}(\mu))}, \end{aligned}$$

as claimed in (5.4). □

Remark 5.5. The assumption $z_{2j+1} = \overline{z_{2j+2}} \to \infty$ as $j \to \infty$ is very restrictive and can be relaxed. For instance, if

$$\limsup_{j\to\infty} \frac{\max(|z_{2j+1} - a|, |z_{2j+1} - b|)}{\mathrm{dist}(z_{2j+1}, [a, b])} < \sqrt[4]{2},$$

then it follows from the proof of Proposition 5.2 that $\sup_j B_{j+1,j} < 1$. As a consequence, the operator U from the proof of Theorem 5.4 and thus B is a compact perturbation of a boundedly invertible operator. This implies that (3.1) and hence the second part of the statement of Theorem 5.4 is still true. □

In the setting of Theorem 5.4, we may therefore apply our findings of Theorem 3.3, Theorem 3.5, or Theorem 3.6 in order to study the convergence of the multipoint Padé approximants towards the Markov function φ, compare with [16, Theorem 6.2].

Finally, returning to the discussion of Remark 2.1 concerning the degrees of freedom of representing multipoint Padé approximants via linear pencils, it is not difficult to see that the two linear pencils $zB - A$ and $\Delta D(zB - A)D^{-1}\Delta$ for diagonal D, Δ with non-zero diagonal entries generate the same continued fraction (2.1). Notice that the matrix D does not affect the diagonal entries and can be therefore be considered as to be a balancing factor for the offdiagonal entries, whereas Δ allows to scale the entries. In terms of the continued fraction (2.1), a scaling corresponds to considering an equivalence transformation of (2.1), and different normalizations can be found in the literature concerning the special cases of J-fractions, T-fractions or Thiele continued fractions. A balancing, however, leaves invariant the continued fraction (2.1) and just addresses the question how to factorize the products $\alpha_j^L\alpha_j^R$.

It is always possible to choose a scaling such that the resulting matrices A, B become bounded. However, such a scaling might produce a matrix B having no longer a bounded inverse, or satisfying no longer the condition (3.1). We also know from [11, Theorem 2.3] that, for fixed z, the balancing which is best for obtaining $z \in \rho(A, B)$ is the one which makes $zB - A$ to be complex symmetric (i.e., a complex Jacobi matrix). In the special case of Theorem 5.4, we have chosen a balancing factor to make B real symmetric, and a scaling such that A, B are bounded and B has a bounded inverse.

A study of best scaling or balancing for general linear pencils is beyond the scope of this paper. For future research it might be interesting to consider a (formal) factorization $z_0 B - A = M_1(z_0) M_2(z_0)$ for some fixed z_0 (as done in § 4) and to discuss the convergence of multi-point approximants in terms of spectral properties of $z \mapsto M_1(z_0)^{-1}(zB - A) M_2(z_0)^{-1}$, since this latter quantity does not depend on scaling or balancing (but depends on how to choose the factors $M_j(z_0)$).

REFERENCES

[1] N.I. Akhiezer, *The classical moment problem*, Oliver and Boyd, Edinburgh, 1965.

[2] A. Ambroladze, On exceptional sets of asymptotic relations for general orthogonal polynomials, *J. Approx. Theory* **82** (1995) 257-273.

[3] A.I. Aptekarev, V.A. Kaliaguine, W. Van Assche, Criterion for the resolvent set of nonsymmetric tridiagonal operators, *Proc. Amer. Math. Soc.* **123** (1995) 2423-2430.

[4] G.A. Baker, P. Graves-Morris, *Padé approximants. Parts I and II*, Encyclopedia of Mathematics and its Applications, 13, 14. Addison-Wesley Publishing Co., Reading, Mass., 1981.

[5] D. Barrios Rolania, G. Lopez Lagomasino, Asymptotic behavior of solutions of general three term recurrence relations, Advances in Computational Mathematics **26** (2007) 937.

[6] B. Beckermann, M. Castro, On the determinacy of complex Jacobi matrices, *Math. Scand.* **95** (2004), 285-298.

[7] B. Beckermann, A. Osipov, Some spectral properties of infinite band matrices, *Numerical Algorithms* **34** (2003) 173-185.

[8] B. Beckermann, Complex Jacobi matrices, *J. Comput. Appl. Math.* **127** (2001) 17-65.

[9] B. Beckermann, On the classification of the spectrum of second order difference operators, *Math. Nachr.* **216** (2000), 45-59.

[10] B. Beckermann, On the convergence of bounded J–fractions on the resolvent set of the corresponding second order difference operator, *J. Approx. Theory* **99** (1999) 369-408.

[11] B. Beckermann, V. Kaliaguine, The diagonal of the Padé table and the approximation of the Weyl function of second order difference operators, *Constructive Approximation* **13** (1997) 481-510.

[12] M.I. Bueno, F. Marcelán, Darboux transformation and perturbation of linear functionals, *Linear Algebra Appl.* **384** (2004) 215-242.

[13] S. Demko, W. F. Moss, P. W. Smith, Decay rates for inverses of band matrices, *Math. Comp.* **43** (1984) 491-499.

[14] M. Derevyagin, Generalized Jacobi operators in Krein spaces, *J. Math. Anal. Appl.* **349** (2009) 568-582.

[15] M. Derevyagin, V. Derkach, On the convergence of Padé approximations for generalized Nevanlinna functions, *Trans. Moscow Math. Soc.* **68** (2007) 119-162.

[16] M.S. Derevyagin, A.S. Zhedanov, An operator approach to multipoint Padé approximations, *J. Approx. Theory* **157** (2009) 70-88.

[17] A.A. Gonchar, On uniform convergence of diagonal Padé approximants, *Math. USSR Sb.* **46** (1983) 539-559.

[18] A. Gonchar and G. Lopez, On Markov's theorem for multipoint approximants for functions of Stieltjes type, *Math. USSR Sb.* **105** (1978) 512-524. English translation in *Math. USSR Sb.* **34** (1978) 449-459.

[19] E. Hendriksen and O. Njastad, A Favard Theorem for Rational Functions, *J. Math. Anal. Applics.* **142** (1989) 508-520.

[20] M.E.H. Ismail and D.R. Masson, Generalized orthogonality and continued fractions, *J. Approx. Theory* **83** (1995) 1-40.

[21] W.B. Jones and W.J. Thron, *Continued fractions. Analytic theory and applications*. Encyclopedia of Mathematics and its Applications, 11. Addison-Wesley Publishing Co., Reading, Mass., 1980.

[22] L. Lorentzen, H. Waadeland, Continued fractions with applications, North Holland, Amsterdam, 1992.

[23] A.S. Markus, Introduction to the spectral theory of polynomial operator pencils, Translations of Mathematical Monographs, AMS, 1988.

[24] E. M. Nikishin, V. N. Sorokin, Rational Approximations and Orthogonality, Translations of Mathematical Monographs 92, AMS, Providence, R.I., 1991.

[25] B. Simon, The classical moment problem as a self-adjoint finite difference operator, *Adv. in Math.* **137** (1998) 82-203.
[26] B. Simon, Orthogonal polynomials on the unit circle, Part 1, AMS, Providence, RI, 2005.
[27] V.Spiridonov, A.Zhedanov, To the theory of biorthogonal rational functions, RIMS Kokyuroku **1302** (2003), 172–192, http://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/42746
[28] H. Stahl and V. Totik, General orthogonal polynomials, Cambridge University Press, 1992.
[29] G. Szegő, Orthogonal Polynomials, fourth edition, AMS, 1975.
[30] E. Titchmarsh, The theory of functions, second edition, Oxford Science Publications, 1939.
[31] H. S. Wall, Analytic Theory of Continued Fractions, Chelsea, Bronx NY, 1973.
[32] A. Zhedanov, Biorthogonal rational functions and the generalized eigenvalue problem, *J. Approx. Theory* **101** (1999) 303-329.

BERNHARD BECKERMANN, LABORATOIRE PAINLEVÉ UMR 8524 (ANO-EDP), UFR MATHÉMATIQUES – M3, UST LILLE, F-59655 VILLENEUVE D'ASCQ CEDEX, FRANCE
E-mail address: `bbecker@math.univ-lille1.fr`

MAXIM DEREVYAGIN, DEPARTMENT OF NONLINEAR ANALYSIS, INSTITUTE OF APPLIED MATHEMATICS AND MECHANICS, R.LUXEMBURG STR. 74, 83114 DONETSK, UKRAINE, AND LABORATOIRE PAINLEVÉ, UFR MATHÉMATIQUES – M3, UST LILLE, F-59655 VILLENEUVE D'ASCQ CEDEX, FRANCE,
E-mail address: `derevyagin.m@gmail.com`

ALEXEI ZHEDANOV, INSTITUTE FOR PHYSICS AND ENGINEERING, R.LUXEMBURG STR. 72, 83114 DONETSK, UKRAINE
E-mail address: `zhedanov@yahoo.com`