Finite sections method for Hessenberg matrices.

Leonel Robert
Department of Mathematics
University of Toronto
lrobert@math.utoronto.ca

Luis Santiago
Facultad de Matemáticas y Computación
Universidad de La Habana
lsmcu@yahoo.com

Abstract

Let D be a Hessenberg matrix and $\mathcal{D}$ the closed operator associated to it. In this work we study the approximation of the resolvent of $\mathcal{D}$ by the resolvents of the truncated matrices D_n of D. A concept of determinacy is introduced analogous to the one for Jacobi matrices. We apply our results to some questions in rational approximation.

keywords: finite sections, Hessenberg matrices, Jacobi matrices, othogonal polynomials, asymptotics.

0 Introduction.

A Hessenberg matrix is an infinite matrix $D = (d_{i,j})_{i,j=0}^{\infty}$ such that $d_{i,j} = 0$ for $j > i+1$; in this paper we assume additionaly that it satisfies $d_{i,i+1} \neq 0$. This paper is devoted to the approximation of the resolvent of a Hessenberg matrix D by the resolvents of its finite sections D_n. Specifically, we consider a closed operator $\mathcal{D} : l_2 \to l_2$ associated to D by

$$\begin{aligned} domain(\mathcal{D}) &= \{x \in l_2 : Dx \in l_2\}, \\ \mathcal{D}x &= Dx \end{aligned}$$

and we look for conditions to ensure the strong convergence of operators

$$\Pi_n^* R(z, D_n)^* \Pi_n \to R(z, \mathcal{D})^*, \tag{1}$$

where $D_n = (d_{i,j})_{i,j=0}^{n-1}$ are the finite sections of D, $R(z, D_n)$, $R(z, \mathcal{D})$ are the resolvent functions of D_n and $\mathcal{D}$ and $\Pi_n : l_2 \to \mathbb{C}^n$ is the projection $\Pi_n x = (x_0, x_1, \ldots, x_{n-1})^t$.

As we will show below, one way to rephrase (1) is to say that the *finite sections method* converges for $\overline{D}^t - \overline{z}I$. There is an extensive literature about the finite sections method. An account of the results in this area can be found in [3], with emphasis on the case of Toeplitz operators.

Convergence (1) is important in problems arising in rational approximation, particularly in the convergence of Padé approximants and of asymptotics for orthogonal polynomials. When D is a Jacobi matrix, $(R(z, D_n))_{0,0}$ is the n-th Padé approximant of $(R(z, \mathcal{D}))_{0,0}$, and if D is real and determinate, Stieljes' Theorem states that (1) holds for every $z \in \mathbb{C} \backslash \mathbb{R}$ (see [12]). When D is a perturbation of a real Jacobi matrix, an arbitrary complex Jacobi matrix or a banded matrix, convergence (1) has been studied by a number of authors, such as G. Lopez, B. Beckermann, V. Kaliaguin, et al, in papers [2], [1], [7]. In all these works extensive use is made of the connection between these matrices, orthogonal polynomials and Padé approximants. In this work we generalize some of the results proved in the previously cited papers. Our methods are closely related to theirs, particularly to [2].

In [13] we develop a general setting for the study of orthogonal polynomials and of infinite simultaneous Padé approximation. This has motivated our interest in proving convergence of type (1) for Hessenberg matrices, since most of the relevant objects of that theory can be written in terms of a Hessenberg matrix and its finite sections. The paper [13] contains a number of applications for the theorems proved here.

This work is organized as follows. Section 1 introduces the definitions and notation that will be used throughout the paper. Most of these definitions are borrowed from the theory of orthogonal polynomials. In particular, a generalization of the concept of determinate Jacobi matrix to Hessenberg matrices is presented and an Invariability Theorem is proven.

In Section 2 we prove the main theorems on strong convergence of the resolvents of the finite sections of a determinate Hessenberg matrix to its resolvent function.

Section 3 contains applications to obtain asymptotics for orthogonal polynomials. We will be interested in what happens to orthogonal polynomials when perturbing their recurrence coefficients. Theorems 6 and 7 generalize results of P. Nevai and W. van Assche for orthogonality in the real line ([15],[11]). The theorems of Section 2 are also applied to obtain convergence of the simultaneous Padé approximants associated to a band matrix. This extends previous results of V. Kaliaguin ([7]).

We are grateful to Dr. Bernhard Beckermann for his useful suggestions during the preparation of this work.

1 Hessenberg matrices.

We call an infinite matrix $D = (d_{i,j})_{i,j=0}^{\infty}$ with $d_{i,j} \in \mathbb{C}$ such that $d_{i,j} = 0$ for $j > i+1$ and $d_{i,i+1} \neq 0$, a lower Hessenberg matrix.

1.1 Hessenberg matrices as operators in l_2.

Let us denote by l_2 the Hilbert space of infinite column vectors $x = (x_0, x_1, \dots)^t$, $x_i \in \mathbb{C}$ satisfying

$$\sum_{i=0}^{\infty} |x_i|^2 < \infty,$$

and denote by C_0 the dense subspace of l_2 formed by the vectors with a finite number of nonzero components.

Consider the action of the infinite Hessenberg matrix D on an infinite vector defined by the usual matrix product Dx. This is well defined since each component of Dx is computed by a finite sum of products. In this way we can associate to D an operator $\mathcal{D} : l_2 \to l_2$

$$\begin{aligned} domain(\mathcal{D}) &= \{x \in l_2 : Dx \in l_2\}, \\ \mathcal{D}x &= Dx. \end{aligned}$$

We also define the operator $\mathcal{H}$ with domain C_0 such that $\mathcal{H}x = \overline{D}^t x$, $x \in C_0$.

Theorem 1. *The operator $\mathcal{D}$ is closed and defined on its maximal domain of definition. Moreover, one has that $\mathcal{D} = \mathcal{H}^*$.*

Proof. The operator $\mathcal{H}$ is densely defined in C_0. Therefore $\mathcal{H}$ has an adjoint (Theorem III.5.28 of [8]). It follows from the identity

$$\langle Dx, y \rangle = \langle x, \overline{D}^t y \rangle, \qquad x \in l_2, y \in C_0$$

that $\mathcal{D}$ is the adjoint of $\mathcal{H}$, and thus it is closed. □

Remark. Notice that, unlike $\mathcal{D}$, the operator $\mathcal{H}$ is not in general closed or even closable. It is closable only when $\mathcal{D}$ is densely defined.

The same argument as in the theorem above shows that to any infinite matrix $A = (a_{i,j})_{i,j=0}^{\infty}$ with $a_{i,j} = 0$ for $j > i + k$ for some $k \in \mathbb{Z}$, one can associate a closed operator $\mathcal{A}$ defined on its maximal domain. On the other hand, to every bounded operator $\mathcal{T}$ can be associated the infinite matrix $T = (t_{i,j})$ with $t_{i,j} = \langle \mathcal{T} e_i, e_j \rangle$, where $\{e_i\}_{i=0}^{\infty}$ denotes the canonical basis of

l_2; that is, $e_i = (\delta_{i,n})_{n=0}^{\infty}$. We will make frequent use of this correspondence between matrices and operators throughout this paper (we use caligraphic fonts to denote operators). One has to be careful with this correspondence, since composition of operators does not always correspond to products of matrices. However, we do have the following:

If A is a Hessenberg matrix corresponding to the maximal closed operator $\mathcal{A}$ and $\mathcal{B}$ is a bounded operator with matrix B such that range$(\mathcal{B}) \subset$ domain$(\mathcal{A})$ then $\mathcal{A}\mathcal{B}$ is bounded with matrix AB.

Here is the proof: $\mathcal{A}\mathcal{B}$ is closed and defined everywhere, so it is bounded. Let $x \in l_2$, then $\mathcal{A}\mathcal{B}x = \mathcal{A}(Bx) = ABx$, hence the matrix of $\mathcal{A}\mathcal{B}$ is AB. Though not explicitly stated, we will make frequent use of this fact.

Now that $\mathcal{D}$ is closed, this allows us to talk about its spectrum $\sigma(\mathcal{D})$, the resolvent set $\Omega(\mathcal{D})$, the resolvent function $R(z, \mathcal{D}) = (zI - \mathcal{D})^{-1}$ and related concepts.

We end this subsection discussing the relation between upper Hessenberg matrices and cyclic operators.

Let $\mathcal{A}$ be an operator in l_2 with dense domain S, such that $\mathcal{A}S \subset S$ and $\mathcal{A}$ has a cyclic vector $v \in S$; that is, $span\{\mathcal{A}^i v, i = 0, \dots\}$ is dense in l_2.

Theorem 2. *The operator $\mathcal{A}$ is unitary equivalent to an extension of the operator $\mathcal{H}$ associated to the transpose of a Hessenberg matrix.*

Proof. By the Gram-Schmidt process applied to the sequence $\{\mathcal{A}^i v\}_{i=0}^{\infty}$ find an orthonormal sequence $\{v_i\}_{i=0}^{\infty}$. Since v is cyclic, $\{v_i\}_{i=0}^{\infty}$ is an orthonormal basis of l_2. Define the unitary operator $\mathcal{U}$ such that $\mathcal{U}e_i = v_i$. Then the domain of $\mathcal{U}^*\mathcal{A}\mathcal{U}$ contains C_0 and $\mathcal{U}^*\mathcal{A}\mathcal{U}e_i \in span\{e_j, j = 0, \dots, i+1\}$. Thus, the matrix of $\mathcal{U}^*\mathcal{A}\mathcal{U}$ with respect to the canonical basis is the transpose of a Hessenberg matrix. □

Suppose that the domain of $\mathcal{A}$ is $span\{\mathcal{A}^i v, i = 0, \dots\}$, then $\mathcal{A}$ is unitary equivalent to $\mathcal{H}$. Define the sequence of finite dimensional subspaces $S_n = span\{\mathcal{A}^i v, i = 0 \dots, n\}$ and let $\mathcal{P}_n$ be the orthogonal projection in S_n. Let us write $\mathcal{A}_n = P_n \mathcal{A}|_{S_n}$. Then the convergence (1) implies (we change $\overline{z}$ by z)

$$P_n^* R(z, \mathcal{A}_n) P_n \to (R(z, \mathcal{A}^*))^*.$$

When the above convergence holds we say that the finite section methods converges for $\mathcal{A} - zI$ and the sequence of projections $\mathcal{P}_n$. If $\mathcal{A}$ is closable and $\mathcal{A}_{min}$ denotes its closure then the right member in the above limit becomes $R(z, \mathcal{A}_{min})$.

In the sequel we restrict our analysis to Hessenberg matrices, particularly to lower Hessenberg matrices. However, the above remark shows how the results to be proven in Section 2 can be reinterpreted as providing a suitable sequence of projections $\mathcal{P}_n$ to apply the finite sections method to cyclic operators.

1.2 Orthogonal polynomials.

To the infinite Hessenberg matrix D we associate a sequence of polynomials $\{p_i(z)\}_{i=0}^{\infty}$ defined by the recurrence relations, $p_0(z) = 1$ and

$$zp_n(z) = \sum_{i=0}^{n+1} d_{n,i} p_i(z),$$

which can be written in vector form as

$$D\mathbf{p} = z\mathbf{p}, \tag{2}$$

where $\mathbf{p} = (p_0, p_1, \dots)^t$.

Since $\deg p_n = n$, these polynomials form a basis of $\mathbb{C}[z]$ and we can define a complex valued linear functional $\Lambda : \mathbb{C}[z, \overline{z}] \to \mathbb{C}$ by

$$\Lambda(p_n \overline{p_m}^t) = \delta_{n,m}.$$

The functional Λ is positive, that is, $\Lambda(p\overline{p}) > 0$ when $p \neq 0$. Thus, $\langle p, q\rangle = \Lambda(p\overline{q})$ is a scalar product in $\mathbb{C}[z]$ and it is clear that the sequence $\{p_n\}_{n=0}^{\infty}$ is a sequence of orthonormal polynomials with respect to this scalar product.

Example. *Suppose that D is a real, symmetric and tridiagonal matrix (a real Jacobi matrix) with $d_{i,i+1} > 0$. Then, by (2), the sequence $\{p_n\}_{n=0}^{\infty}$ satisfies a three term recurrence relation, and by Favard's Theorem there exists a positive measure μ with infinite support and finite moments such that*

$$\Lambda(p(z, \overline{z})) = \int_{-\infty}^{\infty} p(t, t) d\mu(t), \tag{3}$$

and thus

$$\int_{-\infty}^{\infty} p_n(t) p_m(t) d\mu(t) = \delta_{n,m}.$$

Hence $\{p_n\}_{n=0}^{\infty}$ is a sequence of orthonormal polynomials in the real line.

Let us extend the functional Λ to infinite matrices by applying it to each entry (this is clearly an extension if we identify the space of polynomials with the infinite matrices that have zeroes in all entries, except possibly in the upper left corner). It can be checked that left and right multiplying scalar matrices, enter in and out of the parenthesis. That is, whenever the product $AP(z,\overline{z})B$ is well defined, it follows that $\Lambda(AP(z,\overline{z})B) = A\Lambda(P(z,\overline{z}))B$, where $P(z,\overline{z})$ is an infinite matrix with entries in $\mathbb{C}[z,\overline{z}]$ and A and B are infinite scalar matrices. This will allow us to take advantage of the matrix calculus and of a compact notation in proving the next formulas. Also notice that if A, B are Hessenberg matrices and X is an arbitrary infinite matrix, then the matrix products AX and XB^t are well defined and $(AX)B^t = A(XB^t)$.

Define the infinite matrix

$$\widetilde{D}(z) = \Lambda_t\left(\frac{\mathbf{p}(t)-\mathbf{p}(z)}{t-z}\overline{\mathbf{p}(t)}^t S^t\right),$$

where $(S)_{i,j} = (\delta_{i,j+1})$ is the infinite shift matrix. ¿From the orthogonality of the polynomials follows at once that $\widetilde{D}(z)$ is lower triangular. This matrix plays an important role in the study of D because $\widetilde{D}(z)S$ is a right inverse of $D - zI$ and it is closely related to the resolvent of $\mathcal{D}$. Namely, we have the following matrix formulas

$$(D - zI)\widetilde{D}(z) = S^t, \quad (4)$$

$$R(z,D) = -\widetilde{D}(z)S + \mathbf{p}(z)\mathbf{q}(z)^t, \quad (5)$$

where $R(z,D)$ denotes the matrix associated to the operator $R(z,\mathcal{D})$ and $\mathbf{q}(z) = e_0^t R(z,D)$.

In order to prove (4) we take $D - zI$ inside the parenthesis in

$$(D - zI)\Lambda_t\left(\frac{\mathbf{p}(t)-\mathbf{p}(z)}{t-z}\overline{p(t)}^t S^t\right),$$

and use the fact that $(D - zI)\mathbf{p}(t) = (t-z)\mathbf{p}(t)$. To prove (5) we need the following lemma.

lemma 1. *Let W be a Hessenberg matrix and A an arbitrary infinite matrix. Then the equation*

$$WX = A \quad (6)$$

has a unique solution X_0 with the first row equal to zero. Every other solution is of the form $X = X_0 + \mathbf{v}\mathbf{x}^0$, where $\mathbf{v}$ is the unique infinite vector such that $W\mathbf{v} = 0$, $v_0 = 1$, and $\mathbf{x}^0$ is the first row of X.

Proof. By (4) it is clear that $X_0 = \widetilde{W}(0)SA$ is a solution of the given equation with the first row equal to zero. Also, observe that every solution of $Wx = 0$ is of the form $x = \mathbf{v}u$ with $u \in \mathbb{C}$. Thus if $WX = 0$ this implies that every column of X is of the form $\mathbf{v}u_i$ and $X = \mathbf{v}\mathbf{u}^t$.

Let X be a solution of (6). We have $W(X - X_0) = 0$, thus $X - X_0 = \mathbf{v}\mathbf{u}^t$. Comparing the first row in both sides we get $\mathbf{u}^t = x^0$, the first row of X. □

Now (5) follows from the previous lemma applied to the matrix identities

$$(D - zI)\widetilde{D}(z)S = I, \quad (zI - D)R(z, D) = I,$$

for $z \in \Omega(\mathcal{D})$. Here we take $W = D - zI$ and $A = I$.

When D is a real Jacobi matrix, the elements of $\widetilde{D}(z)$ are just the so called "shifted polynomials", $(\widetilde{D}(z))_{n,i} = p^i_{n-i}(z)$. To keep consistent with this case, we use the same notation for the entries of a general $\widetilde{D}(z)$.

1.3 Determinate Hessenberg matrices.

A real Jacobi matrix is said to be determinate if the moment problem associated to the functional Λ is determinate; that is, there is a unique measure μ such that (3) holds. This is equivalent to

$$\sum_{n=0}^{\infty}(|p_n(z)|^2 + |p^1_n(z)|^2) = \infty, \tag{7}$$

for some $z \in \mathbb{C}$.

The Theorem of Invariability states that if (7) holds for some $z_0 \in \mathbb{C}$ then it holds for every $z \in \mathbb{C}$. Moreover, if the Jacobi matrix J' is a bounded perturbation of J then they are both determinate or indeterminate (see [12]).

For complex Jacobi matrices, (7) is taken as the definition of determinacy. B. Beckerman studied determinacy in this case in its connection to proper matrices ([2]). Again, a theorem of invariability holds, as is proven by M. Castro in [10].

Here we use an analogue of (7) to define determinacy for Hessenberg matrices. We show how our definition agrees with the one for complex Jacobi matrices and we prove an invariability theorem.

Definition 1. *Let $\widetilde{\mathcal{D}}(z)$ denote the operator associated to the matrix $\widetilde{D}(z)$. We say that the Hessenberg matrix D is indeterminate if and only if $\widetilde{\mathcal{D}}(z)$ is a Hilbert-Schmidt operator for some $z \in \mathbb{C}$.*

Recall that a bounded operator $\mathcal{A}$ of matrix $A = (a_{i,j})_{i,j=0}^{\infty}$ is said to be a Hilbert-Schmidt operator if $\sum_{i,j=0}^{\infty} |a_{i,j}|^2 < \infty$ ([5]). Since $p^n_0(z) = d^{-1}_{n-1,n}$

(this is not difficult to prove from the definition), a sufficient condition for a Hessenberg matrix to be determinate is $\sum_{n=0}^{\infty} |d_{n-1,n}^{-2}| = \infty$. In particular, every bounded Hessenberg matrix is determinate.

When D is a Jacobi matrix one has the formula

$$\widetilde{D}(z)S - S^t\widetilde{D}(z)^t = \mathbf{p}^1(z)\mathbf{p}^t(z) - \mathbf{p}(z)(\mathbf{p}^1(z))^t, \tag{8}$$

with $\mathbf{p}^1(z) = \widetilde{D}(z)e_1$. This formula can be deduced from the definition of $\widetilde{D}(z)$ applied to this case. It follows from this formula that $\widetilde{\mathfrak{D}}(z)$ being a Hilbert-Schmidt operator is equivalent to

$$\sum_{n=0}^{\infty}(|p_n(z)|^2 + |p_n^1(z)|^2) < \infty,$$

and thus for Jacobi matrices our definition of determinacy agrees with the standard one.

In order to prove invariability of determinacy we need the following lemma.

lemma 2. *Let D_1, D_2 be Hessenberg matrices such that $D_2 = D_1 - \Delta$ with Δ a lower triangular matrix. Then*

$$\widetilde{D}_2(z) = \widetilde{D}_1(z) + \widetilde{D}_1(z)S\Delta\widetilde{D}_2(z).$$

This is the so called "comparison equation" which has been noted before for real Jacobi matrices by W. van Assche in [15]. In [15] this formula is used to prove several perturbation theorems.

Proof. We have

$$(D_1 - zI)(I - \widetilde{D}_1(z)S\Delta)\widetilde{D}_2(z) = (D_2 - zI)\widetilde{D}_2(z) = S^t.$$

Hence $(I - \widetilde{D}_1(z)S\Delta)\widetilde{D}_2(z)$ and $\widetilde{D}_1(z)$ are both solutions of the equation $(D_1 - zI)X = S^t$ with the same first row equal to e_0, and by lemma 1 they are equal. □

Theorem 3. *Determinacy of Hessenberg matrices is invariant under lower triangular bounded perturbations.*

Proof. Write $D_2 = D_1 - \Delta$ with Δ lower triangular and bounded and $\widetilde{\mathfrak{D}}_1(z)$ Hilbert-Schmidt. Then by the comparison equation

$$(I - \widetilde{D}_1(z)S\Delta)\widetilde{D}_2(z) = \widetilde{D}_1(z).$$

The operator $I - \widetilde{\mathfrak{D}}_1(z)S\Delta$ is a compact perturbation of the identity. Therefore, if $0 \in \sigma(I - \widetilde{\mathfrak{D}}_1(z)S\Delta)$ then $0 \in \sigma_p(I - \widetilde{\mathfrak{D}}_1(z)S\Delta)$. But $I - \widetilde{D}_1(z)S\Delta$ is a lower triangular matrix with ones on the main diagonal. Hence $I - \widetilde{\mathfrak{D}}_1(z)S\Delta$ is injective and we conclude that it is invertible. The matrices $(I - \widetilde{D}_1(z)S\Delta)^{-1}\widetilde{D}_1(z)$ and $\widetilde{D}_2(z)$ are both lower triangular and satisfy the same equation $(D_2 - zI)X = S^t$. Hence they are equal; that is,

$$\widetilde{D}_2(z) = (I - \widetilde{D}_1(z)S\Delta)^{-1}\widetilde{D}_1(z).$$

Now since $\widetilde{\mathfrak{D}}_1(z)$ is Hilbert-Schmidt, so is $\widetilde{\mathfrak{D}}_2(z)$. □

Corollary 1. *If for some $z_0 \in \mathbb{C}$, $\widetilde{\mathfrak{D}}(z_0)$ is a Hilbert-Schmidt operator, then $\widetilde{\mathfrak{D}}(z)$ is Hilbert-Schmidt for every $z \in \mathbb{C}$.*

Proof. If we take $D_2 = D_1 - (z - z_0)I$ then it is easy to see that $\widetilde{D}_2(z_0) = \widetilde{D}_1(z)$. Now the result follows from the previous theorem. □

2 Finite sections method for Hessenberg matrices.

We denote by Π_n the $n \times \infty$ matrix

$$\Pi_n = (e_0, e_1, \ldots, e_{n-1})^t.$$

To every $n \times n$ matrix M we can associate the infinite matrix $\Pi_n^t M \Pi_n$ coincident with M in the upper left corner and with zeroes everywhere else. In the sequel, when talking about a finite matrix as infinite, for example $R(z, D_n)$, we always mean the infinite matrix obtained through Π_n. Also, for an infinite matrix A, we write $(A)_n = \Pi_n A \Pi_n^t$ for its finite sections. The same conventions apply for finite and infinite vectors.

For a sequence $\{\mathcal{T}_n\}_{n=0}^{\infty}$ of bounded operators in l_2, we write $\mathcal{T}_n \to \mathcal{T}$ meaning convergence in the strong topology of operators; that is, pointwise convergence. Whenever the convergence is understood in a different sense, this will be specified.

A well known theorem of Kantorovich states that if $\mathcal{A}$, $\mathcal{A}_n$ are bounded operators with $\mathcal{A}_n \to \mathcal{A}$ and $\|\mathcal{A}_n^{-1}\| < M$, then $\mathcal{A}$ is invertible and $\mathcal{A}_n^{-1} \to \mathcal{A}^{-1}$. In particular, this theorem can be applied to $(A)_n$, the finite sections of A. In the next two theorems we proof analogues of this result for determinate (but not necessarily bounded) Hessenberg matrices. These theorems generalize classical ones for real Jacobi matrices and more recent ones for complex Jacobi matrices and banded matrices (see Theorem 4.3, Corollary 4.4 of [2]).

We start with a lemma proving that for Hessenberg matrices one has that $(\mathcal{D}^{-1})_n - D_n^{-1}$ is always of rank one. We use the notation

$$\begin{aligned} \pi_n(z) &= \frac{1}{p_n(z)}(e_n^t \widetilde{D}(z)S)_n, \\ r_{i,j}(z) &= e_j^t R(z,D)e_i. \end{aligned}$$

lemma 3. *We have the following matrix identities*

$$(R(z,\mathcal{D}))_n - R(z,D_n) = \mathbf{p_n}(\mathbf{q}_n - \pi_{\mathbf{n}}) \qquad \text{for } z \in \Omega(D_n)\cap\Omega(\mathcal{D}) \quad (9)$$

$$(R(z,D_n))_m - R(z,D_m) = \mathbf{p_m}((\pi_{\mathbf{n}})_m - \pi_{\mathbf{m}}) \ \ \text{for } z \in \Omega(D_n)\cap\Omega(D_m) \quad (10)$$

$$\begin{aligned} &\Pi_{n-k+1}((R(z,\mathcal{D}))_n - R(z,D_n)) = \qquad (11)\\ &= \frac{1}{r_{n-k,n}(z) + p_{k-1}^{n-k+1}(z)} \Pi_{n-k+1} R(z,\mathcal{D}) e_{n-k}(e_n^t R(z,\mathcal{D}))_n \end{aligned}$$

for $z \in \Omega(\mathcal{D})\cap\Omega(D_n)$ *and* $n \geq k-1 \geq 0$.

Proof. From (5) of Section 1 we have

$$(R(z,D))_n = -(\widetilde{D}\,(z)\,S)_n + \mathbf{p_n}(z)\mathbf{q}_n(z). \qquad (12)$$

From (4) we have

$$(D_n - zI_n)\,(\widetilde{D}(z)S + e_{n-1}d_{n-1,n}(e_n^t\widetilde{D}(z)S)_n = I_n,$$

hence

$$(D_n - zI_n)\,((\widetilde{D}(z)S)_n - \mathbf{p_n}(z)\pi_{\mathbf{n}}(z)) = I_n,$$

so that

$$R(z,D_n) = -(\widetilde{D}(z)S)_n + \mathbf{p_n}(z)\pi_{\mathbf{n}}(z). \qquad (13)$$

Now (9) and (10) are readily obtained using (12) and (13).

Starting from (5) and performing some straightforward calculations, the following identities can be obtained:

$$\begin{aligned} \mathbf{p_{n-k+1}}(z)q_{n-k}(z) &= \Pi_{n-k+1}R(z,\mathcal{D})e_{n-k} & n \geq k-1 \geq 0 \\ p_n(z)q_{n-k}(z) &= r_{n-k,n}(z) + p_{k-1}^{n-k+1}(z) & n \geq k-1 \geq 0 \end{aligned}$$

$$\mathbf{q}_n(z) - \pi_{\mathbf{n}}(z) = \frac{1}{p_n(z)}(e_n^t R(z,\mathcal{D}))_n. \qquad (14)$$

We use these identities to write the right side of (9) like in (11):

$$\begin{aligned}\Pi_{n-k+1}\mathbf{p}_n(\mathbf{q}_n-\pi_{\mathbf{n}}) &= \mathbf{p}_{\mathbf{n-k+1}}(z)\frac{q_{n-k}(z)}{q_{n-k}(z)}\frac{(e_n^t R(z,\mathcal{D}))_n}{p_n(z)} \\ &= \frac{\Pi_{n-k+1}R(z,\mathcal{D})e_{n-k}(e_n^t R(z,\mathcal{D}))_n}{r_{n-k,n}(z)+p_{k-1}^{n-k+1}(z)}.\end{aligned}$$

□

Definition 2. *For an infinite matrix A and an infinite sequence of indices $S \subset \{0,1,\dots\}$ define the sets:*

$$\begin{aligned}\Theta_S(A) &= \{z : \limsup_{n\in S} \|R(z,(A)_n)\| < \infty\}, \\ \widetilde{\Gamma}(A) &= \overline{\bigcup_n \Gamma((A)_n)} = \overline{\{\langle Ax, x\rangle : \|x\| = 1, x \in C_0\}}, \\ \widetilde{\Gamma}_{ess}(A) &= \bigcap_{\mathcal{K}} \widetilde{\Gamma}(A+K),\end{aligned}$$

where $\Gamma((A)_n)$ denotes the numerical range of the truncated matrix $(A)_n$ and $\mathcal{K}$ ranges through all the compact operators with lower triangular matrix.

The set $\Theta_S(A)$ is open. This follows from its definition and the inequality

$$\|(B+\epsilon I)^{-1}\| \leq \frac{\|B^{-1}\|}{1-|\epsilon|\|B^{-1}\|}, \tag{15}$$

with $\epsilon < (\|B^{-1}\|)^{-1}$, applied to $B = (A)_n - zI_n$. When $S = \{1,2\dots\}$ we ommit the subscript and write simply $\Theta(A)$. We have $(\widetilde{\Gamma}(A))^c \subset \Theta(A)$, which follows from

$$\|R(z,(A)_n)\| \leq \frac{1}{d(z,\Gamma((A)_n))}. \tag{16}$$

Theorem 4. *Suppose we have one of the two cases:*

a) D is determinate and $S = \{0,1,\dots\}$,

b) $\sum_{n\in S} |d_{n-1,n}|^{-2} = \infty$.

Then the following hold:

i)

$$\Theta_S(D) = \{z : R(z,D_n)^* \to R(z,\mathcal{D})^*, n \in S\} \subset \Omega(\mathcal{D}).$$

For $x \in l_2$ the convergence

$$\lim_{n\in S} R(z,D_n)^* x = R(z,\mathcal{D})^* x$$

is uniform in compact subsets of $\Theta_S(D)$.

ii) If $z \in \Theta_S(D) \cap \Omega(\mathcal{D}+\mathcal{K})$ *and* $\mathcal{K}$ *is compact, then* $z \in \Theta_S(D+K)$.

iii) $\Omega(\mathcal{D})\backslash\widetilde{\Gamma}_{ess}(D) \subset \Theta_S(D)$.

Proof. i) Banach-Steinhaus' Theorem implies uniform boundness of the sequence $R(z, D_n)$ from its strong convergence, thus we have inclusion in one direction.

Next we want to prove that the uniform boundness of $\|R(z, D_n)\|$ for $n \in S$, n large enough, implies that $z \in \Omega(\mathcal{D})$ and $R(z, D_n)^* \to R(z, \mathcal{D})^*$. We split the argument in two parts.

Claim 1. If $z \in \Theta_S(D)$ *and* D *satisfies a) or b) then* $\mathcal{D} - zI$ *is injective.*

Suppose that $\mathbf{p}(z) \in l_2$. Comparing the first row in both sides of (13) we get that $e_0^t R(z, D_n) = \pi_{\mathbf{n}}(z)$, then recall the definition of $\pi_{\mathrm{n}}(z)$ to get

$$(e_n^t \widetilde{D}(z) S)_n = p_n(z) e_0^t R(z, D_n)$$

Taking norm in both sides of this equality and using the uniform boundness of $\|R(z, D_n)\|$ we get that $\sum_{n\in S} \|e_n^t \widetilde{D}(z)\| < \infty$. This leads to a contradiction whether we have a) or b) (recall that $p_0^n(z) = d_{n-1,n}^{-1}$ are the elements along the main diagonal of $\widetilde{D}(z)S$).

Thus, in both cases we must have $\mathbf{p}(z) \notin l_2$. Since any nonzero vector v such that $(D - zI)v = 0$ must be a scalar multiple of $\mathbf{p}(z)$, this implies that $\mathcal{D} - zI$ is injective.

Claim 2. If $z \in \Theta_S(D)$ *and* $\mathcal{D} - zI$ *is injective, then* $z \in \Omega(\mathcal{D})$ *and* $R(z, D_n)^* \to R(z, \mathcal{D})^*$, $n \in S$.

Let us take norms on both sides of equation (10). We obtain the inequality $\|\mathbf{p}_m(z)\| \|((\pi_n(z))_m - \pi_m(z))\| < M$. This, together with $\|\mathbf{p}(z)\| = \infty$, yields that $\{\pi_n(z)\}_{n\in S}$ is convergent in l_2.

For $x \in C_0$ the right side of

$$x^t \left(R(z, D_n))_m - R(z, D_m)\right) = x^t \mathbf{p}_m(z) \left((\pi_n(z))_m - \pi_m(z)\right)$$

converges to zero as $n, m \to \infty$. This implies that $R(z, D_n)^* x$ is convergent for $x \in C_0$, $n \in S$. Let $x \in l_2$ and take x_0 in C_0 such that $\|x - x_0\| < \varepsilon/M$. The right side of

$$\begin{aligned} \left((R(z, D_n))_m^* - R(z, D_m)^*\right) x &= \left((R(z, D_n))_m^* - R(z, D_m)^*\right)(x - x_0) + \\ &+ \left((R(z, D_n))_m^* - R(z, D_m)^*\right) x_0 \end{aligned}$$

has norm less than ε when $n, m \to \infty$. Thus $\{R(z, D_n)^*\}_{n\in S}$ converges in the strong topology to some bounded operator $\mathcal{R}^*$. The identity (13) implies the matrix formula $R = -\widetilde{D}(z)S + \mathbf{p}(z)e_0^t R$. Matrix $-\widetilde{D}(z)S$ is a right inverse for $zI - D$ (see 4), hence R is a right inverse for $zI - D$ too. This means that $zI - \mathcal{D}$ has a bounded right inverse, thus it is surjective. It is also injective, so we conclude that $zI - \mathcal{D}$ is invertible and $\mathcal{R} = R(z, \mathcal{D})$. This completes the proof of the claim.

It remains to prove that convergence is uniform on compact subsets of $\Theta_S(D)$. Let $K \subset \Theta_S(D)$ be a compact set. For every $z_0 \in K$ we can choose an open neighborhood $V(z_0)$ of z_0 such that $\|R(z, D_n)\| < M(z_0)$ for every $z \in V(z_0)$ and $n > N(z_0)$, $n \in S$ (use the inequality (15)). Choosing a finite covering of K from these neighborhoods we prove that the sequence $\{R(z, D_n)\}_{n \in S}$ is uniformly bounded in K for $n > N$. Now we use Vitali's Theorem to obtain the uniform convergence of $R(z, D_n)^* x$ for every $x \in l_2$ uniformly in K.

ii) We use the following simple lemma in Hilbert spaces.

lemma 4. *If $\mathcal{K}_n$, $\mathcal{K}$ are compact operators such that $\|\mathcal{K} - \mathcal{K}_n\| \to 0$ and $\mathcal{T}$, $\mathcal{T}_n$ are bounded operators such that $\mathcal{T}_n \to \mathcal{T}$, then*

$$\|\mathcal{T}_n \mathcal{K}_n - \mathcal{T}\mathcal{K}\| \to 0 \quad \textit{and} \quad \|\mathcal{K}_n \mathcal{T}_n - \mathcal{K}\mathcal{T}\| \to 0.$$

The proof is simple and it is omitted. The lemma is first proved when $\mathcal{K}$ is of finite rank and then it is completed by using the fact that compact operators are norm limits of finite rank operators.

The operator $I - \mathcal{K}R(z, \mathcal{D}) = (zI - (\mathcal{D} + \mathcal{K}))R(z, \mathcal{D})$ is invertible because it is Fredholm and injective. $\{R(z, D_n)^*\}_{n \in S}$ converges strongly to $R(z, \mathcal{D})^*$, so by the previous lemma $\|(\mathcal{K})_n R(z, D_n) - \mathcal{K}R(z, \mathcal{D})\| \to 0$. Thus $I - (\mathcal{K})_n R(z, D_n)$ is invertible for n big enough. Now the result follows from the equality

$$R(z, (\mathcal{D} + \mathcal{K})_n) = R(z, D_n)\left(I - (\mathcal{K})_n R(z, D_n)\right)^{-1}.$$

iii) Let $z \in (\widetilde{\Gamma}(D + K)^c \bigcap \Omega(D))$ for some compact operator K. Then $z \in \Theta_S(D + K)$ and since $D - zI$ is a compact perturbation of $D + K - zI$ we obtain by ii) that $z \in \Theta_S(D)$. □

For proper complex Jacobi matrices, it is shown in [2] that i) holds for any subsequence S without the need of b). Let us see how this is true for determinate Jacobi matrices.

Suppose that D is a Jacobi determinate matrix and that $z \in \Theta_S(D)$. If $\|\mathbf{p}(z)\|^2 < \infty$ then we derive from (13) that $\widetilde{\mathcal{D}}(z)$ is bounded. But then by (8) the boundness of $\widetilde{\mathcal{D}}(z)$ is equivalent to the indeterminacy of D. Thus $\|\mathbf{p}(z)\|^2 = \infty$. The reasoning continues as in the proof of Theorem 4 i).

Imposing some extra conditions on D we can improve Theorem 4. The norms of the functions $r_{n-k,n}(z)$ in the formula (11) are bounded by $\|R(z,\mathcal{D})\|$. Thus if $p_k^n(z) \to \infty$ for some $z \in \Omega(\mathcal{D})$, as $n \to \infty$, $n \in S$, then

$$(p_{k-1}^{n-k+1}(z) + r_{n-k,n}(z))^{-1} = p_{k-1}^{n-k+1}(z)^{-1}\left(1 + \frac{r_{n-k,n}(z)}{p_{k-1}^{n-k+1}(z)}\right)^{-1} \to 0,$$

and using (11) we get

$$\|(\Pi_{n-k+1}R(z,D_n))^* - R(z,\mathcal{D})^*\| \to 0.$$

In particular, since $p_0^n(z) = d_{n-1,n}^{-1}$,

$$\lim_{n\in S} d_{n-1,n} = 0 \Rightarrow \Theta_S(D) = \Omega(\mathcal{D}).$$

If $|d_{n-1,n}|$ is a bounded sequence, we can improve Theorem 4. Define the sets

$$\begin{aligned} Z_S(D) &= \{z : z \in \sigma(D_n), n \in S\}, \\ Z_S^\infty(D) &= \{z : \exists\{z_{n_k}\}, z_{n_k} \in \sigma(D_{n_k}), z_{n_k} \to z, n_k \to \infty, n_k \in S\}. \end{aligned}$$

It follows from (15) and (16) that we have the inclusions

$$Z_S(D) \subset \widetilde{\Gamma}(D), \quad Z_S^\infty(D) \subset (\Theta_S(D))^c.$$

Since $\widetilde{\Gamma}(D)$ is a convex set, its complement has at most two connected components. Define $\Omega^\infty(\mathcal{D})$ to be the union of the connected components of $\Omega(\mathcal{D})$ which have nonempty intersection with $(\widetilde{\Gamma}(D))^c$ (again there are at most two of them). Finally, write $v(K,D_n)$ for the number of eigenvalues of D_n in $K \subset \mathbb{C}$.

Theorem 5. *Suppose that $\widetilde{\Gamma}(D) \neq \mathbb{C}$ and that $\{|d_{n-1,n}|\}_{n\in S}$ is a bounded sequence. Then we have*

i) For every compact set $K \subset \Omega^\infty(\mathcal{D})$, there is an infinite subsequence $S_1 \subset S$ such that $K\backslash\Theta_{S_1}(D)$ is at most finite.

ii)$\Omega^\infty(\mathcal{D})\backslash Z_S^\infty(D) \subset \Theta_S(D)$.

iii) For every compact set $K \subset \Omega^\infty(D)$, the sequence $\{v(K,D_n)\}_{n\in S}$ is uniformly bounded.

Proof. Before we go into the main argument, we need to prove some preliminary facts.

Claim 1. The matrix D is determinate. This is clear from the fact that $\{|d_{n-1,n}^{-1}|\}_{n\in S}$ is bounded away from zero.

Claim 2. There is an infinite subsequence of indices $S_0 \subset S$ such that $\lim_{n\in S_0} d_{n-1,n} = d$, for some $d \in \mathbb{C}$. This is clear.

Claim 3. In every connected component of $\Omega^\infty(\mathcal{D})$ there are points z_1 such that

$$\|R(z_1,\mathcal{D})\| < \frac{1}{2\max(1,|d|)}. \tag{17}$$

Let us prove this assertion. Since $\widetilde{\Gamma}(D)$ is not the whole plane, every connected component of its complement contains a semiplane. Using (16) we can choose points far enough from $\widetilde{\Gamma}(D)$ such that (17) holds for every D_n. But D is determinate, thus the strong convergence of $R(z,D_n)$ in $(\widetilde{\Gamma}(D))^c$ implies (17) for $R(z,\mathcal{D})$.

Now let $K \subset \Omega^\infty(D)$ be a compact set and assume without lost of generality that it is in one of the connected components of $\Omega^\infty(D)$. Choose a connected open set V such that $K \subset V \subset \overline{V} \subset \Omega^\infty(D)$ and $z_1 \in V$ for some z_1 satisfying (17). The sequence $\{r_{n-1,n}(z)d_{n-1,n}\}_{n\in S_0}$ is uniformly bounded in $\overline{V}$. Hence for some infinite subsequence $S_1 \subset S_0$ we have $(1 + r_{n-1,n}(z)d_{n-1,n}) \to r(z)$ uniformly in $\overline{V}$. The analytic function $r(z)$ cannot be identically zero since $r(z_1) \neq 0$ (use (17)). Thus for any $z \in K$, except in a finite set of points,

$$|(d_{n-1,n}^{-1} + r_{n-1,n}(z))^{-1}| = |(d_{n-1,n}(1 + r_{n-1,n}(z)d_{n-1,n})^{-1}| < M_z$$

for some constant M_z and $n \in S_1$ large enough. Now i) follows from (11) with $k = 0$.

ii)Let $z_0 \in \Omega^\infty(\mathcal{D})\backslash Z_S^\infty(D)$. We are going to prove that for every infinite subsequence $S_0 \subset S$ we can find an infinite subsequence $S_1 \subset S_0$ such that $z_0 \in \Theta_{S_1}(D)$ and from this it will follow that $z \in \Theta_S(D)$. Take $V(z_0)$ a neighborhood of z_0 with $\overline{V(z_0)} \subset \Omega^\infty(\mathcal{D})\backslash Z_S^\infty(D)$. By i) applied to $\overline{V(z_0)}$ there is some $S_1 \subset S_0$ such that $z \in \Theta_{S_1}(D)$ for all except a finite set of points of $V(z_0)$. In particular, $\{\|R(z,D_n)\|\}_{n\in S_1}$ is uniformly bounded for n large enough and $z \in \Gamma$ in some smooth closed Jordan curve enclosing z_0. Since $z \notin Z_S^\infty(D)$, every function $R(z,D_n)$, $n \in S$ is analytic in $V(z_0)$ for n large enough. Thus we have

$$R(z_0,D_n) = \frac{1}{2\pi i}\int_\Gamma \frac{1}{z-z_0}R(z,D_n)dz,$$

and from this follows that $\{R(z_0,D_n)\}_{n\in S_1}$ is bounded for n large enough.

iii) In order to prove the boundness of the sequence $v(K, D_n)$ for $n \in S$ we use the same technique as before. That is, we prove that to every subsequence $S_0 \subset S$ we can find $S_1 \subset S_0$ infinite, such that $v(K, D_n)$ is bounded for $n \in S_1$. But this follows from the discussion in the proof of i) and the chain of inclusions

$$\begin{aligned} \sigma(D_n) \cap \Omega(D) = \{z : p_n(z) = 0\} \cap \Omega(D) \quad &\subset \quad \{z : p_n(z) q_{n-1}(z) = 0\} = \\ &= \quad \{z : d_{n-1,n}^{-1} + r_{n-1,n}(z) = 0\}. \end{aligned}$$

□

The following proposition contains a very weak form of approximation of $R(z, D_n)$ to $R(z, \mathcal{D})$ in any point of $\Omega(\mathcal{D})$. This is a well known fact for Jacobi matrices.

Proposition 1. *Let $z \in \Omega(\mathcal{D})$ and $k \geq 0$. There is an infinite sequence of indices S such that* $\lim_{n \in S}(R(z, D_n))_k = (R(z, \mathcal{D}))_k$.

Proof. Let us suppose that $z \in \Omega(\mathcal{D}) \cap \Omega(D_n)$. By (14) we have that

$$(\mathbf{q}(z) - \pi_{\mathbf{n}}(z))_k = (p_n(z))^{-1} (e_n^t R(z, \mathcal{D}))_k.$$

Suppose the norm of right side is separated from 0 as n tends to infinite. Then $|p_n(z)|^2 < \frac{1}{\epsilon} \|(e_n^t R(z, \mathcal{D}))_k\|^2$ for some $\epsilon > 0$ and every n. This implies that $\mathbf{p}(z) \in l_2$, which contradicts $z \in \Omega(\mathcal{D})$. Thus $\lim_{n \in S}(\pi_n(z))_k = (\mathbf{q}(z))_k$ for some sequence of indices S. Now use (9) to complete the proof. □

3 Applications.

This section contains applications of the above results to problems in rational approximation. Specifically, to asymptotics of orthogonal polynomials and convergence of simultaneous Padé approximants. The link between the finite sections method and these subjects has been known, and exploited, for some time, see for example C. Brezinski, A. Magnus, in [4], [9]. Other problems in rational approximation allow the same approach, as for example the reconstruction of planar domains by their moments ([6], [14]).

3.1 Asymptotics of polynomials.

Let $\{p_n(z)\}_{n=0}^{\infty}$ be a sequence of polynomials with coefficients in $\mathbb{C}$ such that $p_0 = 1$ and $\deg p_n = n$. To this sequence we can associate the Hessenberg matrix D such that

$$D\mathbf{p} = z\mathbf{p},$$

where $\mathbf{p} = (p_0, p_1, \dots)^t$. Inversely, as we saw in Section 1, we can associate to a Hessenberg matrix D the unique sequence of polynomials satisfying this equation with $p_0 = 1$.

In this subsection we are going to study the asymptotics of the polynomials $p_n(z)$ when their associated Hessenberg matrix is perturbed. Sequences like $\{p_n(z)\}_{n=0}^{\infty}$ can arise when considering orthogonal polynomials with respect to some scalar product. More specifically, consider a scalar product $\langle \cdot, \cdot \rangle$ on the polynomials with coefficients in $\mathbb{C}$, $\langle \cdot, \cdot \rangle : \mathbb{C}[z] \times \mathbb{C}[z] \to \mathbb{C}$ such that $\langle 1, 1 \rangle = 1$. Applying the Gram-Schmidt orthogonalization process to the basis z^n we obtain a sequence of orthonormal polynomials $\{p_n(z)\}_{n=0}^{\infty}$ such that $p_0 = 1$ and $\deg p_n = n$.

Write $\{p_{1,n}\}_{n=0}^{\infty}$, $\{p_{2,n}\}_{n=0}^{\infty}$ for two sequences of polynomials and D_1, D_2 for their respective Hessenberg matrices. The Hessenberg matrices that we will consider in the sequel are assumed to be determinate, so that the results of Section 2 can be applied.

The next theorem is well known for real and complex Jacobi matrices ([11], [15], [2]), where a converse is also valid. (Theorem 3.7 of [2])

Theorem 6. *Let D_1, D_2 be Hessenberg matrices with $D_1 - D_2 = \Delta$ a lower triangular matrix representing a compact operator in l_2. Then we have*

$$\lim_{n\to\infty} \left(\frac{p_{1,n-1}(z)}{p_{1,n}(z)} - \frac{p_{2,n-1}(z)}{p_{2,n}(z)} \right) \frac{1}{d_{n-1,n}} \to 0,$$

uniformly on compact subsets of $\Theta(D_1) \cap \Omega(\mathcal{D}_2)$. Under the conditions described in Theorems 4 and 5 of Section 2, the same result holds for subsequences of indices.

Proof. Let $z \in \Theta(D_1) \cap \Omega(D_2)$. We have

$$R(z, D_{1,n}) - R(z, D_{2,n}) = R(z, D_{2,n})(\Delta)_n R(z, D_{1,n}).$$

Since $\Theta(D_1) \cap \Omega(\mathcal{D}_2) = \Theta(D_1) \cap \Theta(D_2)$ (use Theorem 4 ii)), $R(z, D_{1,n})^*$ and $R(z, D_{2,n})^*$ converge strongly. Also Δ_n is a sequence of compact operators that converges in the norm topology to Δ. Thus, the right member of the above equality is a compact operator which by lemma 4 converges in the norm topology to

$$R(z, \mathcal{D}_1) - R(z, \mathcal{D}_2) = R(z, \mathcal{D}_2) \Delta R(z, \mathcal{D}_1).$$

Compact operators take weakly convergent sequences to convergent sequences, thus we have that $\langle (R(z, D_{1,n}) - R(z, D_{2,n})) e_n, e_n \rangle \to 0$. Using (13) one sees that

$$\begin{aligned}\langle R(z,D_{1,n})e_n,e_n\rangle &= \frac{p_{1,n-1}(z)}{p_{1,n}(z)}\frac{1}{d_{n-1,n}},\\ \langle R(z,D_{2,n})e_n,e_n\rangle &= \frac{p_{2,n-1}(z)}{p_{2,n}(z)}\frac{1}{d_{n-1,n}}.\end{aligned}$$

This yields the desired result. □

Denote by $B_k(l_2)$ the the k-Schatten class of bounded operators with finite norm $\|\cdot\|_k$. For $\mathcal{T}\in B_k(l_2)$ we denote by $\det_k(I+\mathcal{T})$ the Fredholm determinant of $I+\mathcal{T}$. In the sequel we will make use of some of the properties of Fredholm determinants. A reference for these and further facts is [5].

Let D_1, D_2 be Hessenberg matrices such that $D_1-D_2=\Delta$ is a lower triangular matrix representing an operator $\Delta\in B_l(l_2)$. For $k\geq l$ define the functions $\phi^k_{-1}:\Omega(\mathcal{D}_1)\to\mathbb{C}$, $\phi^k_{-1,n}:\Omega(D_{1,n})\to\mathbb{C}$ (so called perturbation determinants) by

$$\begin{aligned}\phi^k_{-1}(z) &= \det_k\left(I+\Delta R(z,\mathcal{D}_1)\right),\\ \phi^k_{-1,n}(z) &= \det_k\left(I+\Delta_n R(z,D_{1,n})\right),\end{aligned}$$

which are analytic in $\Omega(\mathcal{D}_1)$ and $\Omega(D_{1,n})$ respectively. Notice that for $k=1$ one has $\phi^1_{-1,n}(z)=\frac{p_{2,n}(z)}{p_{1,n}(z)}$.

Let us define the analytic operator-valued functions $\Phi^k(z)$, $\Phi^k_n(z)$ by

$$\begin{aligned}\Phi^k(z) &= \phi^k_{-1}(z)R(z,\mathcal{D}_2) && z\in\Omega(\mathcal{D}_1)\cap\Omega(\mathcal{D}_2),\\ \Phi^k_n(z) &= \phi^k_{-1,n}(z)R(z,D_{2,n}) && z\in\Omega(D_{1,n})\cap\Omega(D_{2,n}).\end{aligned}$$

Then we have

$$\Phi^k(z)=R(z,\mathcal{D}_1)\phi^k_{-1}(z)\left(I+\Delta R(z,\mathcal{D}_1)\right)^{-1}.$$

But from the Carleman inequality (Theorem XI.9.26, [5]), we have

$$\begin{aligned}\left\|\phi^k_{-1}(z)\left(I+\Delta R(z,\mathcal{D}_1)\right)^{-1}\right\| &\leq \exp\left(C\left\|\Delta R(z,\mathcal{D}_1)\right\|^k_k\right)\\ &\leq \exp\left(C\left\|\Delta\right\|^k_k\left\|R(z,\mathcal{D}_1)\right\|^k\right),\end{aligned}$$

for some constant C. Thus

$$\|\Phi^k(z)\|\leq\|R(z,\mathcal{D}_1)\|\exp\left(C\left\|\Delta\right\|^k_k\left\|R(z,\mathcal{D}_1)\right\|\right),$$

which implies that $\Phi^k(z)$ can be analytically extended to $\Omega(\mathcal{D}_1)$. In the same way it can be seen that $\Phi^k_n(z)$ is analytic in $\Omega(D_{1,n})$.

Theorem 7. *We have*

i) $\phi^k_{-1,n}(z) \to \phi^k_{-1}(z)$,

ii) $\left(\Phi^k_n(z)\right)^* \to \left(\Phi^k(z)\right)^*$,

uniformly in compact subsets of $\Theta(D_1)$.

Proof. i) It is enough to prove the convergence

$$\|\Delta_n R(z, D_{1,n}) - \Delta R(z, \mathcal{D}_1)\|_k \to 0, \tag{18}$$

uniformly in compact subsets of $\Theta(D_1)$ and then the result follows from the continuity of the determinant in $B_k\,(l_2)$ (lemma XI.9.16, [5]).

Lemma 4 of Section 2 still holds if we take the operators $\mathcal{K}_n$ in $B_k(l_2)$ and use the norm $\|\cdot\|_k$ of $B_k(l_2)$ instead of the spectral norm $\|\cdot\|$. By Theorem 4 we have that $R(z, D_{1,n})^* \to R(z, \mathcal{D}_1)^*$ uniformly on compact sets of $\Theta(D_1)$ and also $\|\Delta_n - \Delta\|_k \to 0$ (see lemma XI.9.11 of [5]). Thus the convergence in (18) follows.

ii) Using the Carleman inequality and reasoning as we did before for $\Phi^k\,(z)$, we can see that the sequence of operators $\Phi^k_n\,(z)$ is uniformly bounded on compact subsets of $\Theta(D_1)$. Therefore the sequence $\Phi^k_n(z)^* x$ is normal for every $x \in l_2$, and by Vitali's Theorem it is enough if we prove its convergence in an infinite set with accumulation points.

Without lost of generality we can suppose that $\Theta(D_1)$ is not empty. Since D_2 is a compact perturbation of D_1 the set $\Theta(D_1) \cap \Omega(D_2)$ differs from $\Theta(D_1)$ in at most a denumerable set of isolated points. Hence it is open and not empty. Let $z \in \Theta(D_1) \cap \Omega(\mathcal{D}_2)$. Then, since $(I + \Delta R(z, \mathcal{D}_1))^{-1} = I - \Delta R(z, \mathcal{D}_2)$, we have

$$\begin{aligned}
(\Phi^k(z))^* x &= \overline{\phi^k_{-1}(z)}((I + \Delta R(z, \mathcal{D}_1))^{-1})^* R(z, \mathcal{D}_1)^* x, \\
(\Phi^k_n(z))^* x &= \overline{\phi^k_{-1,n}(z)}((I + \Delta_n R(z, D_{1,n}))^{-1})^* R(z, D_{1,n})^* x.
\end{aligned}$$

Each factor of the second equality converges to the corresponding one of the first equality. Convergence of the rightmost factor was proved in Theorem 4 of Section 2, convergence of the other two factors was proved in i). □

When D_1, D_2 are Jacobi matrices the function $\phi^1_{-1}(z)$ is known as Szego's function or Jost's function, and sometimes it is defined by the formula ([15])

$$\phi^1_{-1}(z) = 1 + \sum_{j=0}^{\infty} (\Delta_{j,j} q_j(z) + \Delta_{j+1,j} q_{j+1}(z)) p_{2,j}(z).$$

Let's see how this formula still holds in the general case. We denote the first row of $\Phi^1(z)$ by $\phi^1(z)$; $\mathbf{q}_1(z)$ denotes the first row of $R(z, D_1)$.

Theorem 8. *Let $D_1 - D_2 = \Delta$ be a lower triangular matrix representing an operator of trace class. Then*

$$\phi_{-1}^1(z) = 1 + \sum_{j=0}^{\infty}\sum_{i=j}^{\infty} \Delta_{ij} q_{1,i}(z) p_{2,j}(z), \tag{19}$$

$$\phi^1(z) = \mathbf{q}_1(z) + \sum_{j=0}^{\infty}\sum_{i=j}^{\infty} \Delta_{ij} q_{1,i}(z) e_j^t \widetilde{D}_2(z), \tag{20}$$

for every $z \in \Omega(\mathcal{D}_1)$. The right side of (20) converges in l_2 and uniformly in compact subsets of $\Omega(\mathcal{D}_1)$.

Proof. Recall from the comparison equation that

$$e_n^t \widetilde{D}_2(z) = e_n^t \widetilde{D}_1(z) + e_n^t \widetilde{D}_1(z) S \Delta \widetilde{D}_2(z), \tag{21}$$

which, dividing by $p_{1,n}$ and equating the first component on both sides, yields

$$\begin{aligned} \frac{p_{2,n}(z)}{p_{1,n}(z)} &= 1 + \pi_{1,n}(z) S \Delta \mathbf{p}_2(z) \\ &= 1 + \sum_{j=0}^{n-1}\sum_{i=j}^{n-1} \Delta_{ij} \frac{p_{1,n-(i+1)}^{i+1}(z)}{p_{1,n}(z)} p_{2,j}(z). \end{aligned} \tag{22}$$

Let us assume that Δ is finite; that is, $(\Delta)_m = \Delta$ for some m. Then for $n > m$ the left side of (22) is

$$\phi_{-1,n}^1(z) = \det(I + \Delta R(z, D_{1,n})) = \det(I + \Delta (R(z, D_{1,n}))_m).$$

Let $z \in \Omega(\mathcal{D}_1)$, by Proposition 1 there is a subsequence of indices such that $(R(z, D_{1,n}))_m \to (R(z, \mathcal{D}_1))_m$. Taking limits, the left side of (22) tends to $\det(I + \Delta(R(z, \mathcal{D}_1))_m) = \det(I + \Delta R(z, \mathcal{D}_1))$, thus we get

$$\phi_{-1}^1(z) = 1 + \sum_{j=0}^{m}\sum_{i=j}^{m} \Delta_{ij} q_{1,i}(z) p_{2,j}(z).$$

for $z \in \Omega(\mathcal{D}_1)$.

Let Δ be of trace class and Δ_n its finite sections. Then we have

$$(\phi_{-1}^1)^{(n)}(z) = 1 + \sum_{j=0}^{n}\sum_{i=j}^{n} \Delta_{ij} q_{1,i}(z) p_{2,j}(z),$$

where $(\phi^1_{-1})^{(n)}$ is the ϕ^1_{-1}-function associated to the perturbation by Δ_n. Since Δ_n tends to Δ in the norm of $B_1(l_2)$, we have that $(\phi^1_{-1})^{(n)} \to \phi^1_{-1}$. Taking limits on both sides of the last equality we get that the right side must be convergent too. This implies (19).

In order to prove (20) we start from (21), divide both sides by $p_{1,n}(z)$ and drop the first component on both sides to get

$$\frac{p_{2,n}(z)}{p_{1,n}(z)}\pi_{2,n}(z) = \pi_{1,n}(z) + \pi_{1,n}(z)S\Delta\widetilde{D}_2(z).$$

Following the same limiting argument used to prove (19), we get (20). □

We have proved asymptotics for orthogonal polynomials under conditions on the Hessenberg matrix; that is, conditions on the parameters of the recurrence relation. A general treatment of the perturbations of real Jacobi matrices can be found in ([15]). There the function $\phi^1_{-1}(z)$ together with the formulas of Theorem 8 are discussed and their connection with scattering theory is developed. In [2], B. Beckermann proves some asymptotics analogous to ours for complex Jacobi matrices and for bounded Hessenberg matrices. For bounded operators, these asymptotics are also studied by M. Putinar in [14]. We have used here the same tecnique for the proofs.

We end this subsection discussing the case of orthogonal polynomials in the circle. Let $\{p_n\}_{n=0}^{\infty}$ be a sequence of orthonormal polynomials in the circle with positive principal coefficients

$$\int_{\Gamma} p_n\overline{p_m}d\mu = \delta_{n,m},$$

for a positive measure μ supported in $\Gamma = \{z \in \mathbb{C} : z\overline{z} = 1\}$ with infinite support. It can be proved that in this case the associated matrix D satisfies $D\overline{D}^t = I$, so that $\overline{D}^t$ represents an isometry in l_2 and hence is bounded of norm 1. Moreover, we have that

$$\Theta(D) = (\Gamma(D))^c = \{z \in \mathbb{C} : |z| > 1\}.$$

If we take $\Upsilon = (d_{01}, d_{12}, ...)$ the diagonal matrix formed with the upper nonzero diagonal of D and write $D = \Upsilon S + \Delta$ where Δ is lower triangular, then it follows from $D\overline{D}^t = I$ that

$$\Upsilon S\overline{\Delta}^t + \Delta S^t\Upsilon^t + \Delta\overline{\Delta}^t = I - \Upsilon^2.$$

It is well known that Szego's condition on the measure μ is equivalent to $\sum_i |1 - d_{i,i+1}| < \infty$. This implies that Υ is invertible and the trace of the left side in the last equality is finite. Hence the trace of $\Delta\overline{\Delta}^t$ is finite and Δ is a Hilbert-Schmidt operator. Moreover, it follows from the above equality that $\Upsilon S\overline{\Delta}^t$ is equal to the lower triangular part of $\Delta\overline{\Delta}^t$. Thus the sums of of the absolute values of the elements in each diagonal of Δ are convergent.

It follows from the discussion above that D is a perturbation of ΥS by a Hilbert-Schmidt lower triangular matrix and thus the theorems of this section apply taking $D_1 = \Upsilon S$ and $D_2 = D$. After some calculations one obtains

$$
\begin{aligned}
p_{1,n}(z) &= \prod_{j=0}^{n} d_{j,j+1} z^n, \quad p_{2,n}(z) = p_n(z), \\
\phi^2_{-1,n}(z) &= \frac{p_n(z)}{\left(\prod_{j=0}^{n} d_{j,j+1}\right) z^n} \exp(-\sum_{j=0}^{n} d_{j,j}),
\end{aligned}
$$

so that

$$\frac{p_n(z)}{z^n} \to C\phi^2_{-1}(z) \text{ uniformly in compact subsets of} \{z \in \mathbb{C} : |z| > 1\},$$

where $C = \exp(\sum_{j=0}^{\infty} d_{j,j}) \prod_{j=0}^{\infty} d_{j,j+1}$. Notice that the shifted polynomials are no longer orthogonal with respect to a measure supported in the unit circle. However our results imply

$$\frac{p^i_{n-i}(z)}{z^{n-i}} \to C_i\phi^2_i(z) \text{ uniformly on compact subsets of } \{z \in \mathbb{C} : |z| > 1\},$$

where $C_i = \exp(\sum_{j=i}^{\infty} d_{j,j}) \prod_{j=i}^{\infty} d_{j,j+1}$.

3.2 Simultaneous Padé approximants.

Let D be a banded Hessenberg matrix with $m+2$ nonzero diagonals. Define the infinite matrix of formal series

$$[zI - D]^{-1} = \sum_{i=0}^{\infty} \frac{D^i}{z^{i+1}},$$

and the formal series $q_i(z) = e_0^t [zI - D]^{-1} e_i$, $i = 0, \ldots, m$.

It is known that the rational functions $\pi_{n,i}(z) = e_0^t R(z, D_n) e_i$, $i = 0, \ldots, m$, are the simultaneous Padé approximants of the functions $q_i(z)$ associated to the sequence of proper multi-indices

$$(\underbrace{r+1, r+1, \ldots, r+1}_{j \text{ times}}, \underbrace{r, r, \ldots, r}_{m-j \text{ times}}),$$

where $n = mr + j$, $0 \leq j < m$. The analytic counterparts of the formal series $q_i(z)$ are the analytic functions $q_i(z) = e_0^t R(z, D)e_i$. References for these and further facts are [12] and [7].

The theorems of Section 2 imply convergence of approximants.

Theorem 9. *If D is determinate, then $\pi_{n,i} \to q_i(z)$ uniformly on compact subsets of $\Theta(D)$. We have the estimate*

$$|q_i(z) - \pi_{n,i}(z)| < O((\sum_{k=0}^{n} |p_k(z)^2|)^{-1}), \quad z \in \Theta(D). \tag{23}$$

Under the conditions described in Theorems 4 and 5 of Section 2, the same results hold for subsequences of indices.

Notice that for every $z \in \Omega(D)$ one has $\sum_{k=0}^{\infty} |p_k(z)^2| = \infty$. It is a known fact that when $\{d_{n-1,n}\}_{n=1}^{\infty}$ is bounded, the sequence $\{|p_n(z)|\}_{n=0}^{\infty}$ has geometric growth in $\Theta(D)$.

Proof. Convergence of the approximants follows at once from Theorem 4 i). Taking norms in both sides of (10) yields the estimate (23). □

References

[1] D. Barrios, G. López, A. Martinez and E. Torrano, On the domain of convergence and poles of complex J-fractions, J. Approx. Theory 93 (1998) 177-200.

[2] B.Beckermann, Complex Jacobi Matrices, J. Comput. Appl. Math 127 (2001)17-65.

[3] A. Bottcher, Infinite matrices and projection methods, Lectures on operator theory and its applications (Waterloo, ON, 1994), 1–72, Fields Inst. Monogr., 3, Amer. Math. Soc., Providence, RI, (1996).

[4] C. Brezinski, "Padé type approximation and General Orthogonal Polynomials", Birkhauser, Basel (1980).

[5] N. Dunford and J.T. Schwartz, "Linear operators", Parts and II, Interscience, New York (1967).

[6] B. Gustafsson, C. He, P. Milanfar, M. Putinar, Reconstructing planar domains from their moments, Inverse Problems 16 (2000), 1053–1070.

[7] V. Kaliaguine, The operator moment problem, vector continued fractions and an explicit form of the Favard theorem for vector orthogonal polynomials, J. Comput. Appl. Math. 65 (1995) 181–193.

[8] T. Kato, "Perturbation theory for linear operators", Springer-Verlag, New York (1966).

[9] A. P. Magnus, Toeplitz matrix techniques and convergence of complex weight Padé approximants, J. Comput. App. Math. 19 (1987) 23–38.

[10] M. Castro Smirnova, Determinacy of Bounded Complex Perturbations of Jacobi matrices, J. Approx. Theory, 104 (2000) no. 6, 345–351.

[11] P. Nevai, W. Van Assche, Compact perturbations of orthogonal polynomials, Pacific J. Math. 153 (1992), no. 1, 163–184.

[12] E. M. Nikishin and V. N. Sorokin, "Rational Approximation and Orthogonality", Translation of Mathematical Monographs 92, Am. Math. Soc., Providence, RI (1991).

[13] L. Robert, L. Santiago, General Orthogonal Polynomials, preprint, (2001).

[14] M. Putinar, On a diagonal Padé approximation in two variables, Numer. Math., 93 (2002), 131–152

[15] W. Van Assche, Constructive methods in the analysis of orthogonal polynomials, Katholieke Universiteit Leuven, Belgium, (Doctoral Dissertation) (1992).