Orthogonal Rational Functions and Structured Matrices

Marc Van Barel *Dario Fasino*
Luca Gemignani *Nicola Mastronardi*

Report TW 350, November 2002

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Orthogonal Rational Functions and Structured Matrices

Marc Van Barel *Dario Fasino*
Luca Gemignani *Nicola Mastronardi*

Report TW 350, November 2002

Department of Computer Science, K.U.Leuven

Abstract

The space of all proper rational functions with prescribed poles is considered. Given a set of points z_i in the complex plane and the weights w_i, we define the discrete inner product

$$\langle \phi, \psi \rangle := \sum_{i=0}^{n} |w_i|^2 \overline{\phi(z_i)} \psi(z_i).$$

In this paper we derive a method to compute the coefficients of a recurrence relation generating a set of orthonormal rational basis functions with respect to the discrete inner product. We will show that these coefficients can be computed by solving an inverse eigenvalue problem for a matrix having a specific structure. In case where all the points z_i lie on the real line or on the unit circle, the computational complexity is reduced by an order of magnitude.

Keywords : orthogonal rational functions, structured matrices, diagonal-plus-semiseparable matrices, inverse eigenvalue problems, recurrence relation
AMS(MOS) Classification : 42C05, 65F18, 65D15.

ORTHOGONAL RATIONAL FUNCTIONS AND STRUCTURED MATRICES

MARC VAN BAREL*, DARIO FASINO†, LUCA GEMIGNANI‡, AND NICOLA MASTRONARDI§

Abstract. The space of all proper rational functions with prescribed poles is considered. Given a set of points z_i in the complex plane and the weights w_i, we define the discrete inner product

$$\langle \phi, \psi \rangle := \sum_{i=0}^{n} |w_i|^2 \overline{\phi(z_i)} \psi(z_i).$$

In this paper we derive a method to compute the coefficients of a recurrence relation generating a set of orthonormal rational basis functions with respect to the discrete inner product. We will show that these coefficients can be computed by solving an inverse eigenvalue problem for a matrix having a specific structure. In case where all the points z_i lie on the real line or on the unit circle, the computational complexity is reduced by an order of magnitude.

Key words. orthogonal rational functions, structured matrices, diagonal-plus-semiseparable matrices, inverse eigenvalue problems, recurrence relation

AMS subject classifications. 42C05, 65F18, 65D15

1. Introduction and motivation.

Proper rational functions are an essential tool in many areas of engineering, as system theory and digital filtering, where polynomial models are inappropriate, due to their unboundedness at infinity. In fact, for physical reasons the transfer functions describing linear time-invariant systems often have to be bounded on the real line. Furthermore, approximation problems with rational functions are in the core of, e.g., the partial realization problem [20], model reduction problems [4, 5, 11], robust system identification [5, 24].

Recently a strong interest has been brought to a variety of rational interpolation problems where a given function is to be approximated by means of a rational function with prescribed poles (see [6, 7, 32] and the references given therein). By linearization, such problems naturally lead to linear algebra computations involving structured matrices. Exploiting the close connections between the functional problem and its matrix counterparts, generally allows us to take advantage of the special structure of these matrices to speed up the approximation scheme. For example, in [25] efficient algorithms are designed for rational function evaluation and interpolation from their connection with displacement structured matrices.

* Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3001 Heverlee, Belgium, e-mail: Marc.VanBarel@cs.kuleuven.ac.be. The research of the first author is partially supported by the Fund for Scientific Research (FWO), "SMA: Structured matrices and their applications", grant G#0078.01, "ANCILA: Asymptotic analysis of the convergence behavior of iterative methods in numerical linear algebra", grant #G.0176.02, the K.U.Leuven research project "SLAP: Structured linear algebra package", grant OT-00-16, and the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, grant IPA V/22. The scientific responsibility rests with the authors.

†Dipartimento di Matematica e Informatica, Università degli Studi di Udine, Viale delle Scienze 208, 33100 Udine, Italy, e-mail: fasino@dimi.uniud.it

‡Dipartimento di Matematica, Università di Pisa, Via F. Buonarroti 2, 56127 Pisa, Italy, e-mail: gemignan@dm.unipi.it. The research of the third author is partially supported by G.N.C.S. and MIUR funds.

§Istituto per le Applicazioni del Calcolo "M. Picone", sez. Bari, Consiglio Nazionale delle Ricerche, Via G. Amendola, 122/I, I-70126 Bari, Italy, e-mail:N.Mastronardi@area.ba.cnr.it

The purpose of this paper is to devise a procedure to construct a set of proper rational functions with prescribed poles, that are orthogonal with respect to a discrete inner product. Orthogonal rational functions are useful in solving multipoint generalizations of classical moment problems and associated interpolation problems; see [7] for further references on this topic. We also mention the recent appearance in the numerical analysis literature of quadrature formulas that are exact for sets of rational functions having prescribed poles, see e.g., [8, 19]. Such formulas provide a greater accuracy than standard quadrature formulas, when the poles are chosen in such a way to mimic the poles present in the integrand. The construction of Gauss-type quadrature formulas is known to be a task closely related to that of orthogonalizing a set of prescribed basis functions. In the polynomial case, this fact was explored by L. Reichel [26, 27]. Indeed, in these papers the construction of polynomial sequences that are orthogonal with respect to a discrete inner product by means of their three-term recurrence relation, is tied with the solution of an inverse eigenvalue problem for symmetric tridiagonal matrices, that is equivalent to the construction of Gauss quadrature formulas.

In this paper, we adapt the technique laid down in [27] for polynomial sequences to a specific set of proper rational functions. The goal is the computation of an orthonormal basis of the linear space $\mathcal{R}_n$ of proper rational functions $\phi(z) = n(z)/d(z)$ w.r.t. a discrete inner product $\langle \cdot , \cdot \rangle$. Here $\deg(n(z)) \leq \deg(d(z)) \leq n$ and $d(z)$ has a prescribed set $\{y_1, \ldots, y_n\}$, $y_i \in \mathbb{C}$, of possible zeros; moreover, we set $\langle \phi, \psi \rangle := \sum_{i=0}^{n} |w_i|^2 \overline{\phi(z_i)} \psi(z_i)$, for $\phi(z), \psi(z) \in \mathcal{R}_n$. Such computation arises in the solution of least squares approximation problems with rational functions with prescribed poles. Moreover, it is also closely related with the computation of an orthogonal factorization of Cauchy-like matrices whose nodes are the points z_i and y_i [16, 14].

We prove that an orthonormal basis of $(\mathcal{R}_n, \langle \cdot , \cdot \rangle)$ can be generated by means of a suitable recurrence relation. When the points z_i as well as the points y_i are all real, fast $O(n^2)$ Stieltjes-like procedures for computing the coefficients of such relation were first devised in [14, 16]. However, like the polynomial (Vandermonde) case [26], these fast algorithms result to be quite sensitive to roundoff errors so that the computed functions are far from orthogonal. Therefore, in this paper we propose a different approach based on the reduction of the considered problem to the following inverse eigenvalue problem (DS-IEP): Find a matrix S of order $n + 1$ whose lower triangular part is the lower triangular part of a rank 1 matrix, and a unitary matrix Q of order $n + 1$ such that $Q^H \vec{w} = \|\vec{w}\| \vec{e}_1$ and $Q^H D_z Q = S + D_y$. Here and below $\vec{w} = [w_0, \ldots, w_n]^T$, $D_z = \mathrm{diag}[z_0, \ldots, z_n]$ and $D_y = \mathrm{diag}[y_0, \ldots, y_n]$, where y_0 can be chosen arbitrarily. Moreover, we denote by $\mathcal{S}_k$ the class of $k \times k$ matrices S whose lower triangular part is the lower triangular part of a rank 1 matrix. If both S and S^H belong to $\mathcal{S}_k$, then S is called a *semiseparable* matrix.

A quite similar reduction to an inverse eigenvalue problem for a tridiagonal symmetric matrix (T-IEP) or for a unitary Hessenberg matix (H-IEP) was also exploited in the theory on the construction of orthonormal polynomials w.r.t. a discrete inner product (see [29, 21, 3, 13, 26, 2, 28, 17] for a survey of the theory and applications on T-IEP and H-IEP). This theory can be generalized to orthonormal *vector* polynomials. We refer the interested reader to [1, 30, 31, 33, 9, 34]. Since invertible semiseparable matrices are the inverses of tridiagonal ones [18], we find that DS-IEP gives a generalization of T-IEP and, in particular, it reduces to T-IEP in the case where $y_i, z_i \in \mathbb{R}$ and all prescribed poles y_i are equal.

We devise a method for solving DS-IEP which fully exploits its recursive properties. This method proceeds by applying a sequence of carefully chosen Givens rotations to update the solution at the k-th step by adding a new data $(w_{k+1}, z_{k+1}, y_{k+1})$. The unitary matrix Q can thus be determined in its factored form as a product of $O(n^2)$ Givens rotations at the cost of $O(n^2)$ arithmetic operations (ops). The complexity of forming the matrix S depends on the structural properties of its upper triangular part and, in general, it requires $O(n^3)$ ops. In the case where all the points z_i lie on the real axis, we show that S is a semiseparable matrix so that the computation of S can be carried out using $O(n^2)$ ops only. In addition to that, the class $\mathcal{S}_{n+1}$ results to be close under bilinear rational (Moebius) transformations of the form $z \to (\alpha z+\beta)/(\gamma z+\delta)$. Hence, by combining these two facts together, we are also able to prove that the process of forming S can be performed at the cost of $O(n^2)$ ops whenever all points z_i belong to a generalized circle (ordinary circles and straight lines) in the complex plane.

This paper is organized in the following way. In Section 2 we reduce the computation of a sequence of orthonormal rational basis functions to the solution of an inverse eigenvalue problem for matrices of the form $\mathrm{diag}[y_0, \ldots, y_n] + S$, with $S \in \mathcal{S}_{n+1}$. By exploiting this reduction, we also determine relations for the recursive construction of such functions. Section 3 provides our method for solving DS-IEP in the general case whereas the more specific situations corresponding to points lying on the real axis, on the unit circle or on a genreic circle in the complex plane are considered in Section 4. In Section 5 we present and discuss numerical experiments confirming the effectiveness and the accuracy of the proposed method and, finally, conclusions and further developments are drawn in Section 6.

2. The computation of orthonormal rational functions and its matrix framework. In this section we will study the properties of a sequence of proper rational functions with prescribed poles, that are orthonormal with respect to a certain discrete inner product. We will also design an algorithm to compute such a sequence via a suitable recurrence relation. The derivation of this algorithm follows from reducing the functional problem into a matrix setting to the solution of an inverse eigenvalue problem involving structured matrices.

2.1. The functional problem. Given the complex numbers $y_1, y_2, \ldots, y_n$ all different from each other. Let us consider the vector space $\mathcal{R}_n$ of all proper rational functions having possible poles in $y_1, y_2, \ldots, y_n$:

$$\mathcal{R}_n := \mathrm{span}\{1, \frac{1}{z-y_1}, \frac{1}{z-y_2}, \ldots, \frac{1}{z-y_n}\}.$$

The vector space $\mathcal{R}_n$ can be equipped with the inner product $\langle \cdot, \cdot \rangle$ defined below:

Definition 2.1 (Bilinear form). *Given the complex numbers $z_0, z_1, \ldots, z_n$ which together with the numbers y_i are all different from each other, and the "weights" $0 \neq w_i$, $i = 0, 1, \ldots, n$, we define a bilinear form $\langle \cdot, \cdot \rangle : \mathcal{R}_n \times \mathcal{R}_n \to \mathbb{C}$ by*

$$\langle \phi, \psi \rangle := \sum_{i=0}^{n} |w_i|^2 \overline{\phi(z_i)} \psi(z_i).$$

Since there is no proper rational function $\phi(z) = n(z)/d(z)$ with $\deg(n(z)) \leq \deg(d(z)) \leq n$ different from the zero function such that $\phi(z_i) = 0$ for $i = 0, \ldots, n$, this bilinear form defines a positive definite inner product in the space $\mathcal{R}_n$.

The aim of this paper is to develop an efficient algorithm for the solution of the following functional problem:

PROBLEM 1 (Computing a sequence of orthonormal rational basis functions). *Construct an orthonormal basis*

$$\vec{\alpha}_n(z) := [\alpha_0(z), \alpha_1(z), \ldots, \alpha_n(z)]$$

of $(\mathcal{R}_n, \langle \cdot , \cdot \rangle)$ *satisfying the properties*

$$\begin{aligned} \alpha_j(z) \in \mathcal{R}_j \setminus \mathcal{R}_{j-1} & \qquad (\mathcal{R}_{-1} := \emptyset) \\ \langle \alpha_i, \alpha_j \rangle = \delta_{i,j} & \qquad \textit{(Kronecker delta)} \end{aligned}$$

for $i, j = 0, 1, 2, \ldots, n$.

We will show later that the computation of such an orthonormal basis $\vec{\alpha}_n(z)$ is equivalent to the solution of an inverse eigenvalue problem for matrices of the form $\mathrm{diag}[y_0, \ldots, y_n] + S$, where $S \in \mathcal{S}_{n+1}$.

2.2. The inverse eigenvalue problem. Let $D_y = \mathrm{diag}[y_0, \ldots, y_n]$ be the diagonal matrix whose diagonal elements are $y_0, y_1, \ldots, y_n$, where y_0 can be chosen arbitrarily; analogously, set $D_z = \mathrm{diag}[z_0, \ldots, z_n]$. Recall that $\mathcal{S}_k$ is the class of $k \times k$ matrices S whose lower triangular part is the lower triangular part of a rank 1 matrix. Furthermore, denote by $\|\vec{w}\|$ the Euclidean norm of the vector $\vec{w} = [w_0, w_1, \ldots, w_n]^T$.

Our approach to solving Problem 1 mainly relies upon the equivalence between that problem and the following inverse eigenvalue problem (DS-IEP):

PROBLEM 2 (Solving an inverse eigenvalue problem). *Given the numbers* w_i, z_i, y_i, *find a matrix* $S \in \mathcal{S}_{n+1}$ *and a unitary matrix* Q *such that*

$$Q^H \vec{w} = \|\vec{w}\| \vec{e}_1, \tag{2.1}$$

$$Q^H D_z Q = S + D_y. \tag{2.2}$$

Observe that, if (Q, S) is a solution of Problem 2, then S can not have zero rows and columns. By contradiction, if we suppose that $S\vec{e}_j = \vec{0}$, where $\vec{e}_j$ is the j-th column of the identity matrix I_{n+1} of order $n+1$, then $D_z Q \vec{e}_j = Q D_y \vec{e}_j = y_{j-1} Q \vec{e}_j$, from which it would follow $y_{j-1} = z_i$ for a certain i.

Results concerning the existence and the uniqueness of the solution of Problem 2 were first proven in the papers [14, 15, 16] for the specific case where $y_i, z_i \in \mathbb{R}$ and S is a semiseparable matrix. In particular, under such auxiliary assumptions, it was shown that the matrix Q is simply the orthogonal factor of a QR decomposition of a Cauchy-like matrix built from the nodes y_i and z_i. Next we give a generalization of the results of [14, 15, 16] to cover with the more general situation considered here.

THEOREM 2.2. *Problem 2 has at least one solution. If* (Q_1, S_1) *and* (Q_2, S_2) *are two solutions of Problem 2, then there exists a unitary diagonal matrix* $F = \mathrm{diag}[1, \mathrm{e}^{\mathrm{i}\theta_1}, \ldots, \mathrm{e}^{\mathrm{i}\theta_n}]$ *such that*

$$Q_2 = Q_1 F, \quad S_2 = F^H S_1 F.$$

Proof. It is certainly possible to find two vectors $\vec{u} = [u_0, \ldots, u_n]^T$ and $\vec{v} = [v_0, \ldots, v_n]^T$ with $v_i, u_i \neq 0$ and $u_i v_0 / (z_i - y_0) = w_i$, for $0 \leq i \leq n$. Indeed, it is sufficient to set, for example, $v_i = 1$ and $u_i = w_i(z_i - y_0)$. Hence, let us consider the nonsingular Cauchy-like matrix $C \equiv (u_{i-1} v_{j-1} / (z_{i-1} - y_{j-1}))$ and let $C = QR$ be a QR-factorization of C. From $D_z C - C D_y = \vec{u}\vec{v}^T$ one easily finds that

$$Q^H D_z Q = R D_y R^{-1} + Q \vec{u} \vec{v}^T R^{-1} = D_y + S,$$

where

$$S = RD_yR^{-1} - D_y + Q\vec{u}\vec{v}^T R^{-1} \in \mathcal{S}_{n+1}.$$

Moreover, $Q\vec{e}_1 = CR^{-1}\vec{e}_1 = \vec{w}/\|\vec{w}\|$ by construction. Hence, the matrices Q and $S = Q^H D_z Q - D_y$ solve Problem 2.

Concerning uniqueness, assume that (Q, S) is a solution of Problem 2 with $S \equiv (s_{i,j})$ and $s_{i,j} = \widetilde{u}_{i-1}\widetilde{v}_{j-1}$ for $1 \leq j \leq i \leq n+1$. As $S\vec{e}_1 \neq \vec{0}$, it follows that $\widetilde{v}_0 \neq 0$ and, therefore, we may assume $\widetilde{v}_0 = 1$. Moreover, from (2.2) it is easily found that

$$D_z Q\vec{e}_1 = Q\vec{\widetilde{u}} + y_0 Q\vec{e}_1,$$

where $\vec{\widetilde{u}} = [\widetilde{u}_0, \ldots, \widetilde{u}_n]^T$. From (2.1) we have

$$\vec{\widetilde{u}} = Q^H(D_z - y_0 I_{n+1})\frac{\vec{w}}{\|\vec{w}\|}. \tag{2.3}$$

Relation (2.2) can be rewritten as

$$Q^H D_z Q = \vec{\widetilde{u}}\vec{\widetilde{v}}^T + U = \vec{\widetilde{u}}\vec{\widetilde{v}}^T + RD_yR^{-1},$$

where U is an upper triangular matrix with diagonal entries y_i and $U = RD_yR^{-1}$ gives its Jordan decomposition, defined up to a suitable scaling of the columns of the upper triangular eigenvector matrix R. Hence, we find that

$$D_z QR - QRD_y = Q\vec{\widetilde{u}}\vec{\widetilde{v}}^T R = \vec{u}\vec{v}^T$$

and, therefore, $QR = C \equiv (u_{i-1}v_{j-1}/(z_{i-1} - y_{j-1}))$ is a Cauchy-like matrix with $\vec{u} = Q\vec{\widetilde{u}}$ uniquely determined by (2.3). This means that all the eligible Cauchy-like matrices C are obtained one from each other by a multiplication on the right by a suitable diagonal matrix. In this way, from the essential uniqueness of the orthogonal factorization of a given matrix, we may conclude that Q is uniquely determined up to multiplication on the right by a unitary diagonal matrix F having fixed its first diagonal entry equal to 1. Finally, the result for S immediately follows from using again relation (2.2). $\square$

The above theorem says that the solution of Problem 2 is essentially unique up to a diagonal scaling. Furthermore, once the weight vector $\vec{w}$ and the points z_i are fixed, then the determinant of S results to be a rational function in the variables $y_0, \ldots, y_n$ whose numerator is not identically zero. Hence, we can show that, for almost any choice of $y_0, \ldots, y_n$, the resulting matrix S is nonsingular. The paper [16] dealt with this *regular* case, in the framework of the orthogonal factorization of real Cauchy matrices. In particular, it is shown there that the matrix S is nonsingular when all the nodes y_i, z_i are real and there exists an interval, either finite or infinite, containing all nodes y_i and none of the nodes z_i.

In what follows we assume that $S^{-1} = H$ exists. It is well known that the inverse of a matrix whose lower triangular part is the lower triangular part of a rank 1 matrix is an irreducible Hessenberg matrix [18]. Hence, we will use the following notation: The matrix $H = S^{-1}$ is upper Hessenberg with subdiagonal elements $b_0, b_1, \ldots, b_{n-1}$; for $j = 0, \ldots, n-1$, the j-th column H_j of H has the form

$$H_j^T =: [\vec{h}_j^T, b_j, \vec{0}], \quad b_j \neq 0.$$

The outline of the remainder of this section is as follows. First we assume that we know a unitary matrix Q and the corresponding matrix S solving Problem 2. Then we provide a recurrence relation between the columns Q_j of Q and, in addition to that, we give a connection between the columns Q_j and the values at the points z_i attained by certain rational functions satisfying a similar recurrence relation. Finally, we show that these rational functions form a basis we are looking for.

2.3. Recurrence relation for the columns of Q. Let the columns of Q denoted as follows:

$$Q =: [Q_0, Q_1, \ldots, Q_n] .$$

THEOREM 2.3 (Recurrence relation). *For $j = 0, 1, \ldots, n$, the columns Q_j satisfy the recurrence relation*

$$b_j(D_z - y_{j+1}I_{n+1})Q_{j+1} = Q_j + ([Q_0, Q_1, \ldots, Q_j]\, D_{y,j} - D_z\, [Q_0, Q_1, \ldots, Q_j])\, \vec{h}_j,$$

with $Q_0 = \vec{w}/\|\vec{w}\|$, $Q_{n+1} = 0$ and $D_{y,j} = \operatorname{diag}[y_0, \ldots, y_j]$.

Proof. Since $Q^H\vec{w} = \vec{e}_1\|\vec{w}\|$, it follows that $Q_0 = \vec{w}/\|\vec{w}\|$. Multiplying relation (2.2) to the left by Q, we have

$$D_zQ = Q(S + D_y).$$

Multiplying this to the right by $H = S^{-1}$, gives us

$$D_zQH = Q(I_{n+1} + D_yH). \tag{2.4}$$

Considering the j-th column of the left and right-hand side of the equation above we have the claim. □

2.4. Recurrence relation for the orthonormal rational functions. In this section we define an orthonormal basis $\vec{\alpha}_n(z) = [\alpha_0(z), \alpha_1(z), \ldots, \alpha_n(z)]$ for $\mathcal{R}_n$ using a recurrence relation built by means of the information contained in the matrix H.

DEFINITION 2.4 (Recurrence for the orthonormal rational functions). *Let us define $\alpha_0(z) = 1/\|\vec{w}\|$ and*

$$\alpha_{j+1}(z) = \frac{\alpha_j(z) + ([\alpha_0(z), \ldots, \alpha_j(z)]\, D_{y,j} - z\, [\alpha_0(z), \ldots, \alpha_j(z)])\, \vec{h}_j}{b_j(z - y_{j+1})},$$

for $0 \le j \le n - 1$.

In the next theorem, we prove that the rational functions $\alpha_j(z)$ evaluated in the points z_i are connected to the elements of the unitary matrix Q. This will allow us to prove in Theorem 2.6 that the rational functions $\alpha_j(z)$ are indeed the orthonormal rational functions we are looking for. In what follows, we use the notation $D_w = \operatorname{diag}[w_0, \ldots, w_n]$.

THEOREM 2.5 (Connection between $\alpha_j(z_i)$ and the elements of Q). *Let*

$$\vec{\alpha}_j = [\alpha_j(z_0), \ldots, \alpha_j(z_n)]^T \in \mathbb{C}^{n+1}, \quad 0 \le j \le n.$$

For $j = 0, 1, \ldots, n$, we have $Q_j = D_w\vec{\alpha}_j$.

Proof. Filling in z_i for z in the recurrence relation for $\alpha_{j+1}(z)$, we get

$$b_j(D_z - y_{j+1}I_{n+1})\vec{\alpha}_{j+1} = \vec{\alpha}_j + ([\vec{\alpha}_0, \ldots, \vec{\alpha}_j]\, D_{y,j} - D_z\, [\vec{\alpha}_0, \ldots, \vec{\alpha}_j])\, \vec{h}_j.$$

Since $Q_0 = \vec{w}/\|\vec{w}\| = D_w\vec{\alpha}_0$, the theorem is proved by finite induction on j, comparing the preceding recurrence with the one in Theorem 2.3. □

Now it is easy to prove the orthonormality of the rational functions $\alpha_j(z)$.

THEOREM 2.6 (Orthonormality of $\vec{\alpha}_n(z)$). *The functions $\alpha_0(z), \ldots \alpha_n(z)$ form an orthonormal basis for $\mathcal{R}_n$ with respect to the inner product $\langle \cdot, \cdot \rangle$. Moreover, we have $\alpha_j(z) \in \mathcal{R}_j \setminus \mathcal{R}_{j-1}$.*

Proof. Firstly, we prove that $\langle \alpha_i, \alpha_j \rangle = \delta_{i,j}$. This follows immediately from the fact that $Q = D_w[\vec{\alpha}_0, \ldots, \vec{\alpha}_n]$ and Q is unitary. Now we have to prove that $\alpha_j(z) \in \mathcal{R}_j \setminus \mathcal{R}_{j-1}$. This is clearly true for $j = 0$ (recall that $\mathcal{R}_{-1} = \emptyset$). Suppose it is true for $j = 0, 1, 2, \ldots, k < n$. From the recurrence relation, we derive that $\alpha_{k+1}(z)$ has the form

$$\alpha_{k+1}(z) = \frac{\text{rational function with possible poles in } y_0, y_1, \ldots, y_k}{(z - y_{j+1})}.$$

Also $\lim_{z\to\infty} \alpha_{k+1}(z) \in \mathbb{C}$ and, therefore, $\alpha_{k+1}(z) \in \mathcal{R}_{k+1}$. Note that simplification by $(z - y_{k+1})$ does not occur in the previous formula for $\alpha_{k+1}(z)$ because $Q_{k+1} = D_w\vec{\alpha}_{k+1}$ is linearly independent of the previous columns of Q. Hence, $\alpha_{k+1}(z) \in \mathcal{R}_{k+1} \setminus \mathcal{R}_k$. □

In the next theorem, we give an alternative relation among the rational functions $\alpha_j(z)$.

THEOREM 2.7 (Alternative relation). *We have*

$$z\vec{\alpha}_n(z) = \vec{\alpha}_n(z)(S + D_y) + \alpha_{n+1}(z)\vec{s}_n, \tag{2.5}$$

where $\vec{s}_n$ is the last row of the matrix S and the function $\alpha_{n+1}(z)$ is given by

$$\alpha_{n+1}(z) = c\prod_{j=0}^{n}(z - z_j)/\prod_{j=1}^{n}(z - y_j)$$

for some constant c.

Proof. Let H_n be the last column of $H = S^{-1}$, and define

$$\alpha_{n+1}(z) = \vec{\alpha}_n(z)(zI_{n+1} - D_y)H_n - \alpha_n(z). \tag{2.6}$$

Thus, the recurrence relation given in Definition 2.4 can also be written as

$$\vec{\alpha}_n(z)(zI_{n+1} - D_y)H = \vec{\alpha}_n(z) + \alpha_{n+1}(z)\vec{e}_{n+1}^T.$$

Multiplying to the right by $S = H^{-1}$, we obtain the formula (2.5). To determine the form of $\alpha_{n+1}(z)$ we look at the definition (2.6). It follows that $\alpha_{n+1}(z)$ is a rational function having degree of numerator at most one more than the degree of the denominator and having possible poles in $y_1, y_2, \ldots, y_n$. Recalling from Theorem 2.5 the notation $\vec{\alpha}_j = [\alpha_j(z_0), \ldots, \alpha_j(z_n)]^T$ and the equation $Q = D_w[\vec{\alpha}_0, \ldots, \vec{\alpha}_n]$, we can evaluate the previous equation in the points z_i and obtain:

$$D_z\,[\vec{\alpha}_0, \ldots, \vec{\alpha}_n]\,H - [\vec{\alpha}_0, \ldots, \vec{\alpha}_n]\,D_yH = [\vec{\alpha}_0, \ldots, \vec{\alpha}_n] + \vec{\alpha}_{n+1}\vec{e}_n^T.$$

Since $D_wD_z = D_zD_w$, multiplying to the left by D_w we obtain

$$D_zQH - QD_yH = Q + D_w\vec{\alpha}_{n+1}\vec{e}_{n+1}^T.$$

From equation (2.4) we obtain that $D_w\vec{\alpha}_{n+1}\vec{e}_{n+1}^T$ is a zero matrix; hence, it follows that $\alpha_{n+1}(z_i) = 0$, for $i = 0, 1, \ldots, n$, and this proves the theorem. □

Note that $\alpha_{n+1}(z)$ is orthogonal to all $\alpha_i(z)$, $i = 0, 1, \ldots, n$, since $\alpha_{n+1}(z) \notin \mathcal{R}_n$ and its norm is

$$\|\alpha_{n+1}\|^2 = \sum_{i=0}^{n} |w_i \alpha_{n+1}(z_i)|^2 = 0.$$

3. Solving the inverse eigenvalue problem. In this section we devise an efficient recursive procedure for the construction of the matrices Q and S solving Problem 2 (DS-IEP). Our procedure is recursive. The case $n = 0$ is trivial: It is sufficient to set $Q = z_0/|z_0|$ and $S = z_0 - y_0$. Let us assume we have already constructed a unitary matrix Q_k and a matrix S_k for the first $k+1$ points $z_0, z_1, \ldots, z_k$ with the corresponding weights $w_0, w_1, \ldots, w_k$. That is, (Q_k, S_k) satisfies

$$Q_k^H \vec{w}_k = \|\vec{w}_k\| \vec{e}_1$$
$$Q_k^H D_{z,k} Q_k = S_k + D_{y,k},$$

where $\vec{w}_k = [w_0, \ldots, w_k]^T$, $S_k \in \mathcal{S}_{k+1}$, $D_{z,k} = \mathrm{diag}[z_0, \ldots, z_k]$ and, similarly, $D_{y,k} = \mathrm{diag}[y_0, \ldots, y_k]$. The idea is now to add a new point z_{k+1} with corresponding weight w_{k+1} and construct the corresponding matrices Q_{k+1} and S_{k+1}.

Hence, we start with the following relations:

$$\begin{bmatrix} 1 & 0 \\ 0 & Q_k^H \end{bmatrix} \begin{bmatrix} w_{k+1} \\ \vec{w}_k \end{bmatrix} = \begin{bmatrix} w_{k+1} \\ \|\vec{w}_k\|\vec{e}_1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ 0 & Q_k^H \end{bmatrix} \begin{bmatrix} z_{k+1} & 0 \\ 0 & D_{z,k} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q_k \end{bmatrix} = \begin{bmatrix} z_{k+1} & 0 \\ 0 & S_k + D_{y,k} \end{bmatrix}.$$

Then, we find complex Givens rotations $G_i = I_{i-1} \oplus G_{i,i+1} \oplus I_{k-i+1}$,

$$G_{i,i+1} =: \begin{bmatrix} c & s \\ -\bar{s} & c \end{bmatrix}, \quad G_{i,i+1}^H G_{i,i+1} = I_2, \tag{3.1}$$

such that

$$G_k^H \cdots G_1^H \begin{bmatrix} 1 & 0 \\ 0 & Q_k^H \end{bmatrix} \begin{bmatrix} w_{k+1} \\ \vec{w}_k \end{bmatrix} = \begin{bmatrix} \|\vec{w}_{k+1}\| \\ 0 \end{bmatrix},$$

and, moreover,

$$G_k^H \cdots G_1^H \begin{bmatrix} 1 & 0 \\ 0 & Q_k^H \end{bmatrix} \begin{bmatrix} z_{k+1} & 0 \\ 0 & D_{z,k} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q_k \end{bmatrix} G_1 \cdots G_k \in \mathcal{S}_{k+2}.$$

Finally, we set

$$Q_{k+1} = \begin{bmatrix} 1 & 0 \\ 0 & Q_k \end{bmatrix} G_1 \cdots G_k,$$

and

$$S_{k+1} = G_k^H \cdots G_1^H \begin{bmatrix} z_{k+1} & 0 \\ 0 & S_k + D_{y,k} \end{bmatrix} G_1 \cdots G_k.$$

With the notation

$$SS\left(\begin{array}{c}[u_0,u_1,\dots,u_k]\\ [v_0,v_1,\dots,v_k]\end{array}\right)$$

we denote the lower triangular matrix whose nonzero part equals the lower triangular part of the rank 1 matrix $[u_{i-1}v_{j-1}]_{i=0,\dots,k}^{j=0,\dots,k}$. Moreover, with the notation

$$RR\left(\begin{array}{c}[\eta_0,\eta_1,\dots,\eta_{k-1}]\\ [\vec{r}_0,\vec{r}_1,\dots,\vec{r}_{k-2}]\end{array}\right)$$

we denote the strictly upper triangular matrix whose $(i+1)$-st row, $0\le i\le k-2$, is equal to $[\vec{0},\eta_i,\vec{r}_i^T]$.

Let us describe now in what way Givens rotations are selected in order to perform the updating of Q_k and S_k. In the first step we construct a Givens rotation working on the new weight. Let $G_{1,2}$ be a Givens rotation as in (3.1), such that

$$G_{1,2}^H\left[\begin{array}{c}w_{k+1}\\ \|\vec{w}_k\|\end{array}\right]=\left[\begin{array}{c}\|\vec{w}_{k+1}\|\\ 0\end{array}\right].$$

The matrix S_k is updated as follows: We know that

$$S_k=SS\left(\begin{array}{c}[u_0,u_1,\dots,u_k]\\ [v_0,v_1,\dots,v_k]\end{array}\right)+RR\left(\begin{array}{c}[\eta_0,\eta_1,\dots,\eta_{k-1}]\\ [\vec{r}_0,\vec{r}_1,\dots,\vec{r}_{k-2}]\end{array}\right).$$

Let

$$S_{k+1,1}+D_{y,k+1,1}:=\left[\begin{array}{cc}G_{1,2}^H&0\\0&I_k\end{array}\right]\left[\begin{array}{cc}z_{k+1}&0\\0&S_k+D_{y,k}\end{array}\right]\left[\begin{array}{cc}G_{1,2}&0\\0&I_k\end{array}\right],$$

where $S_{k+1,1}$ and $D_{y,k+1,1}$ are defined as follows:

$$S_{k+1,1}=SS\left(\begin{array}{c}[\hat{u}_0,\tilde{u}_1,u_1,u_2,\dots,u_k]\\ [\hat{v}_0,\tilde{v}_1,v_1,v_2,\dots,v_k]\end{array}\right)+RR\left(\begin{array}{c}[\hat{\eta}_0,\tilde{\eta}_1,\eta_1,\dots,\eta_{k-1}]\\ \left[\vec{\hat{r}}_0,\vec{\tilde{r}}_1,\vec{r}_1\dots,\vec{r}_{k-2}\right]\end{array}\right)$$

and

$$D_{y,k+1,1}=\operatorname{diag}(y_0,\tilde{y}_1,y_1,y_2,\dots,y_k),$$

with

$$\left[\begin{array}{cc}\alpha&\delta\\ \gamma&\beta\end{array}\right]:=G_{1,2}^H\left[\begin{array}{cc}z_{k+1}&0\\0&y_0+u_0v_0\end{array}\right]G_{1,2}$$

and

$$\begin{array}{lll}\hat{v}_0=-\overline{s}v_0 & \hat{u}_0=(\alpha-y_0)/\hat{v}_0 & \hat{\eta}_0=\delta\\ \tilde{v}_1=cv_0 & \tilde{y}_1=\beta-\tilde{u}_1\tilde{v}_1 & \tilde{u}_1=\gamma/\hat{v}_0\\ \tilde{\eta}_1=c\eta_0 & \vec{\hat{r}}_0=\left[-s\eta_0,-s\vec{r}_0^T\right]^T & \vec{\tilde{r}}_1=c\vec{r}_0.\end{array}$$

In the next steps, we are transforming $D_{y,k+1,1}$ into $D_{y,k+1}$. The first of these steps is as follows. If $v_1\tilde{u}_1-\tilde{\eta}_1\neq 0$, we choose t such that

$$\overline{t}=\frac{y_1-\tilde{y}_1}{v_1\tilde{u}_1-\tilde{\eta}_1},$$

and define the Givens rotation working on the 2-nd and 3-rd row and column as

$$G_{2,3} = \begin{bmatrix} 1 & t \\ -\bar{t} & 1 \end{bmatrix} / \sqrt{1+|t|^2}.$$

Otherwise, if $v_1\tilde{u}_1 - \tilde{\eta}_1 = 0$, we set

$$G_{2,3} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

It turns out that the associated similarity transforms $S_{k+1,1}$ and $D_{y,k+1,1}$ into $S_{k+1,2}$ and $D_{y,k+1,2}$ given by

$$S_{k+1,2} = SS\left(\begin{matrix} [\hat{u}_0, \hat{u}_1, \tilde{u}_2, u_2, \ldots, u_k] \\ [\hat{v}_0, \hat{v}_1, \tilde{v}_2, v_2, \ldots, v_k] \end{matrix} \right) + RR\left(\begin{matrix} [\hat{\eta}_0, \hat{\eta}_1, \tilde{\eta}_2, \eta_2, \ldots, \eta_{k-1}] \\ \left[\vec{\hat{r}}_0, \vec{\hat{r}}_1, \vec{\tilde{r}}_2, \vec{r}_2 \ldots, \vec{r}_{k-2}\right] \end{matrix} \right),$$

$$D_{y,k+1,2} = \mathrm{diag}(y_0, y_1, \tilde{y}_2, y_2, y_3, \ldots, y_k),$$

with

$$G_{2,3}^H \begin{bmatrix} \tilde{u}_1 \\ u_1 \end{bmatrix} = \begin{bmatrix} \hat{u}_1 \\ \tilde{u}_2 \end{bmatrix}, \quad [\tilde{v}_1, v_1]\, G_{2,3} = [\hat{v}_1, \tilde{v}_2]\,.$$

Moreover, $\tilde{y}_2 = \tilde{y}_1$, $\hat{\eta}_1$ is the $(1,2)$-entry of

$$G_{2,3}^H \begin{bmatrix} \tilde{u}_1\tilde{v}_1 + \tilde{y}_1 & \tilde{\eta}_1 \\ u_1\tilde{v}_1 & u_1v_1 + y_1 \end{bmatrix} G_{2,3}$$

and

$$G_{2,3}^H \begin{bmatrix} \vec{\tilde{r}}_1 \\ [\eta_1, \vec{r}_1] \end{bmatrix} = \begin{bmatrix} \vec{\hat{r}}_1 \\ \left[\tilde{\eta}_2, \vec{\tilde{r}}_2\right] \end{bmatrix}. \tag{3.2}$$

At the very end, after k steps, we obtain

$$S_{k+1,k} = SS\left(\begin{matrix} [\hat{u}_0, \hat{u}_1, \ldots, \hat{u}_k, \tilde{u}_{k+1}] \\ [\hat{v}_0, \hat{v}_1, \ldots, \hat{v}_k, \tilde{v}_{k+1}] \end{matrix} \right) + RR\left(\begin{matrix} [\hat{\eta}_0, \hat{\eta}_1, \hat{\eta}_2, \ldots, \hat{\eta}_k] \\ \left[\vec{\hat{r}}_0, \vec{\hat{r}}_1, \vec{\hat{r}}_2, \ldots, \vec{\hat{r}}_{k-1}\right] \end{matrix} \right)$$

and

$$D_{y,k+1,k} = \mathrm{diag}(y_0, y_1, \ldots, y_k, \tilde{y}_{k+1}).$$

The last step will transform $\tilde{y}_{k+1}$ into y_{k+1} by applying the transformation

$$\begin{aligned} \hat{u}_{k+1} &\leftarrow \tilde{u}_{k+1} \\ \hat{v}_{k+1} &\leftarrow (\tilde{y}_{k+1} - y_{k+1} - \tilde{u}_{k+1}\tilde{v}_{k+1})/\tilde{u}_{k+1}. \end{aligned}$$

The computational complexity of the algorithm is dominated by the cost of performing the multiplications (3.2). In general, adding new data $(w_{k+1}, z_{k+1}, y_{k+1})$ requires $\mathcal{O}(k^2)$ ops and hence, computing $S_n = S$ requires $\mathcal{O}(n^3)$ ops. In the next section we will show that these estimates reduce by an order of magnitude in the case where some special distributions of the points z_i are considered which lead to a matrix S with a structured upper triangular part. We stress the fact that, in the light of Theorem 2.2, the above procedure to solve DS-IEP, can also be seen as a method to compute the orthogonal factor in a QR factorization of a suitable Cauchy-like matrix.

4. Special configurations of points z_i. In this section we specialize our algorithm for the solution of DS-IEP to cover with the important case where the points z_i are assumed to lie on the real axis or on the unit circle in the complex plane. Under this assumption on the distribution of the points z_i, it will be shown that the resulting matrix S also possesses a semiseparable structure. The exploitation of this property allows us to overcome the multiplication (3.2) and to construct the matrix $S_n = S$ by means of a simpler parametrization, using $\mathcal{O}(n)$ ops per point, so that the overall cost of forming S reduces to $\mathcal{O}(n^2)$ ops.

4.1. Special case: all points z_i are real. When all the points z_i are real, we have that

$$S + D_y = Q^H D_z Q = (Q^H D_z Q)^H = (S + D_y)^H.$$

Hence, the matrix $S + D_y$ can be written as

$$S + D_y = \mathrm{tril}(\vec{u}\vec{v}^T, 0) + D_y + \mathrm{triu}(\bar{\vec{v}}\vec{u}^H, 1), \tag{4.1}$$

with $\bar{\vec{v}}$ the complex conjugate of the vector $\vec{v}$. Here we adopt the Matlab[1] notation $\mathrm{triu}(B, p)$ for the upper triangular portion of a square matrix B, where all entries below the p-th diagonal are set to zero ($p = 0$ is the main diagonal, $p > 0$ is above the main diagonal, and $p < 0$ is below the main diagonal). Analogously, the matrix $\mathrm{tril}(B, p)$ is formed by the lower triangular portion of B by setting to zero all its entries above the p-th diagonal. In particular, the matrix S is a Hermitian semiseparable matrix, and its computation requires only $\mathcal{O}(n)$ ops per point, since its upper triangular part needs not to be computed via (3.2). Moreover, its inverse matrix H is tridiagonal, hence the vectors $\vec{h}_j$ occurring in Definition 2.4 have only one nonzero entry.

When also all the poles y_i (and the weights w_i) are real, all computations can be performed using real arithmetic instead of doing operations on complex numbers. When all the poles are real or come in complex conjugate pairs, also all computations can be done using only real arithmetic. However, the algorithm works then with a block diagonal D_y instead of a diagonal matrix. The details of this algorithm are rather elaborate. So, we will not go into the details here.

4.2. Special case: all points z_i lie on the unit circle. The case of points z_i located on the unit circle $\mathbb{T} = \{z \in \mathbb{C} \ : \ |z| = 1\}$ in the complex plane can be reduced to the real case treated in the preceding subsection by using the concept of rational bilinear (Moebius) transformation [22]. To be specific, a function $\mathcal{M} : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ is a Moebius transformation if

$$\mathcal{M}(z) = \frac{\alpha z + \beta}{\gamma z + \delta}, \quad \alpha\delta - \beta\gamma \neq 0, \qquad \alpha, \beta, \gamma, \delta \in \mathbb{C}.$$

Interesting properties concerning Moebius transformations are collected in [22]. In particular, a Moebius transformation defines a one-to-one mapping of the extended complex plane into itself and, moreover, the inverse of a Moebius transformation is still a Moebius transformation given by

$$\mathcal{M}^{-1}(z) = \frac{\delta z - \beta}{-\gamma z + \alpha}. \tag{4.2}$$

[1] Matlab is a registered trademark of The MathWorks.

The Moebius transformation $\mathcal{M}(S)$ of a matrix S is defined as

$$\mathcal{M}(S) = (\alpha S + \beta I)(\gamma S + \delta I)^{-1}$$

if the matrix $\gamma S + \delta I$ is nonsingular. The basic fact relating semiseparable matrices with Moebius transformations is that in a certain sense the semiseparable structure is maintained under a Moebius transformation of the matrix. More precisely, we have that:

THEOREM 4.1. *Let* $S \in \mathcal{S}_{n+1}$ *with* $S \equiv (s_{i,j})$, $s_{i,j} = u_{i-1}v_{j-1}$ *for* $1 \leq j \leq i \leq n+1$, *and* $v_0 \neq 0$. *Moreover, let* $D_y = \mathrm{diag}[y_0, \ldots, y_n]$ *and assume that* $\mathcal{M}$ *maps the eigenvalues of both* $S + D_y$ *and* D_y *into points of the ordinary complex plane, i.e.,* $-\delta/\gamma$ *is different from all the points* y_i, z_i. *Then, we find that*

$$\mathcal{M}(S + D_y) - \mathcal{M}(D_y) \in \mathcal{S}_{n+1}.$$

Proof. Observe that $S \in \mathcal{S}_{n+1}$ implies that $RSU \in \mathcal{S}_{n+1}$ for R and U upper triangular matrices. Hence, if we define $R = I - \vec{e}_1[0, v_1/v_0, \ldots, v_n/v_0]$, the theorem is proven by showing that

$$R^{-1}(\mathcal{M}(S + D_y) - \mathcal{M}(D_y))R \in \mathcal{S}_{n+1},$$

which is equivalent to

$$R^{-1}\mathcal{M}(S + D_y)R - \mathcal{M}(D_y) \in \mathcal{S}_{n+1}.$$

One immediately finds that

$$R^{-1}\mathcal{M}(S + D_y)R = ((\gamma(S + D_y) + \delta I)R)^{-1}(\alpha(S + D_y) + \beta I)R,$$

from which it follows

$$R^{-1}\mathcal{M}(S + D_y)R = (\gamma v_0 \vec{u}\vec{e}_1^T + R_1)^{-1}(\alpha v_0 \vec{u}\vec{e}_1^T + R_2),$$

where R_1 and R_2 are upper triangular matrices with diagonal entries $\gamma y_i + \delta$ and $\alpha y_i + \beta$, respectively. In particular, R_1 is invertible and, by applying the Sherman-Morrison formula we obtain

$$R^{-1}\mathcal{M}(S + D_y)R = (I - \sigma R_1^{-1}\vec{u}\vec{e}_1^T)(\alpha v_0 R_1^{-1}\vec{u}\vec{e}_1^T + R_1^{-1}R_2),$$

for a suitable σ. The thesis is now established by observing that the diagonal entries of $R_1^{-1}R_2$ coincides with the ones of $\mathcal{M}(D_y)$ and, moreover, from the previous relation one gets

$$R^{-1}\mathcal{M}(S + D_y)R - R_1^{-1}R_2 \in \mathcal{S}_{n+1},$$

and the proof is complete. □

This theorem has several interesting consequences since it is well known that we can determine Moebius transformations mapping the unit circle $\mathbb{T}$ except for one point onto the real axis in the complex plane. To see this, let us first consider Moebius transformations of the form

$$\mathcal{M}_1(z) = \frac{z + \bar{\alpha}}{z + \alpha}, \quad \alpha \in \mathbb{C} \setminus \mathbb{R}.$$

It is immediately found that $\mathcal{M}_1(z)$ is invertible and, moreover, $\mathcal{M}_1(z) \in \mathbb{T}$ whenever $z \in \mathbb{R}$. For the sake of generality, we also introduce Moebius transformations of the form

$$\mathcal{M}_2(z) = \frac{z-\beta}{1-\bar{\beta}z}, \quad |\beta| \neq 1,$$

which are invertible and map the unit circle $\mathbb{T}$ into itself. Then, by composition of $\mathcal{M}_2(z)$ with $\mathcal{M}_1(z)$ we find a fairly general bilinear transformation $\mathcal{M}(z)$ mapping the real axis into the unit circle:

$$\mathcal{M}(z) = \mathcal{M}_2(\mathcal{M}_1(z)) = \frac{(1-\beta)z + (\bar{\alpha} - \beta\alpha)}{(1-\bar{\beta})z + (\alpha - \bar{\alpha}\bar{\beta})}. \tag{4.3}$$

Hence, the inverse transformation $\mathcal{M}^{-1}(z) = \mathcal{M}_1^{-1}(\mathcal{M}_2^{-1}(z))$, where

$$\mathcal{M}_1^{-1}(z) = \frac{\alpha z - \bar{\alpha}}{-z+1}, \quad \mathcal{M}_2^{-1}(z) = \frac{z+\beta}{\bar{\beta}z+1},$$

is the desired invertible transformation which maps the unit circle (except for one point) into the real axis.

By combining these properties with Theorem 4.1, we obtain efficient procedures for the solution of Problem 2 in the case where all the points z_i belong to the unit circle $\mathbb{T}$.

Let $D_y = \mathrm{diag}[y_0, \ldots, y_n]$ and $D_z = \mathrm{diag}[z_0, \ldots, z_n]$ with $|z_i| = 1$. Moreover, let $\mathcal{M}(z)$ be as in (4.3), such that $\mathcal{M}^{-1}(z_i)$ and $\mathcal{M}^{-1}(y_i)$ are finite, i.e., $z_i, y_i \neq (1-\beta)/(1-\bar{\beta}) = \mathcal{M}_2(1)$, $0 \leq i \leq n$. The solution (Q, S) of Problem 2 with input data $\vec{w}$, $\{\mathcal{M}^{-1}(z_i)\}$ and $\{\mathcal{M}^{-1}(y_i)\}$ is such that

$$Q^H \,\mathrm{diag}[\mathcal{M}^{-1}(z_0), \ldots, \mathcal{M}^{-1}(z_n)]Q = S + \mathrm{diag}[\mathcal{M}^{-1}(y_0), \ldots, \mathcal{M}^{-1}(y_n)],$$

from which it follows that

$$\mathcal{M}(Q^H \,\mathrm{diag}[\mathcal{M}^{-1}(z_0), \ldots, \mathcal{M}^{-1}(z_n)]Q) = \mathcal{M}(S + \mathrm{diag}[\mathcal{M}^{-1}(y_0), \ldots, \mathcal{M}^{-1}(y_n)]).$$

By invoking Theorem 4.1, this relation gives

$$\mathcal{M}(Q^H \,\mathrm{diag}[\mathcal{M}^{-1}(z_0), \ldots, \mathcal{M}^{-1}(z_n)]Q) = Q^H D_z Q = \widehat{S} + D_y, \quad \widehat{S} \in \mathcal{S}_{n+1},$$

and, therefore, a solution of the original inverse eigenvalue problem with points $z_i \in \mathbb{T}$ is $(\widehat{Q}, \widehat{S})$ where $\widehat{Q} = Q$ and $\widehat{S}$ is such that

$$\widehat{S} + D_y = \mathcal{M}(S + \mathrm{diag}[\mathcal{M}^{-1}(y_0), \ldots, \mathcal{M}^{-1}(y_n)]). \tag{4.4}$$

Having shown in (4.1) that the matrix S satisfies

$$S = \mathrm{tril}(\vec{u}\vec{v}^T, 0) + \mathrm{triu}(\bar{\vec{v}}\vec{u}^H, 1),$$

for suitable vectors $\vec{u}$ and $\vec{v}$, we can use (4.4) to further investigate the structure of $\widehat{S}$. From (4.4) we deduce that

$$\widehat{S}^H + D_y^H = \tilde{\mathcal{M}}(S^H + \mathrm{diag}[\mathcal{M}^{-1}(y_0), \ldots, \mathcal{M}^{-1}(y_n)]^H).$$

The Moebius transformation $\tilde{\mathcal{M}}$ of a matrix S is defined as

$$\tilde{\mathcal{M}} = (\bar{\gamma}S + \bar{\delta}I)^{-1}(\bar{\alpha}S + \bar{\beta}I)$$

when $\mathcal{M} = (\alpha z + \beta)/(\gamma z + \delta)$. By applying again Theorem 4.1, assuming that all y_i are different from zero, this implies that

$$\widehat{S}^H + D \in \mathcal{S}_{n+1},$$

for a certain diagonal matrix D. Summing up, we obtain that

$$\widehat{S} = \mathrm{tril}(\vec{u}\vec{v}^T, 0) + \mathrm{triu}(\vec{p}\vec{q}^T, 1), \tag{4.5}$$

for suitable vectors $\vec{u}$, $\vec{v}$, $\vec{p}$ and $\vec{q}$. In case, one or more of the y_i are equal to zero, it can be shown that $\widehat{S}$ is block lower triangular where each of the diagonal blocks has the desired structure. The proof is rather technical. Therefore, we omit it here.

From a computational viewpoint, these results can be used to devise several different procedures for solving Problem 2 in the case of points z_i lying on the unit circle at the cost of $O(n^2)$ ops. By taking into account the semiseparable structure of $\widehat{S}$ (4.5) we can simply modify the algorithm stated in the previous section in such a way to compute its upper triangular part without performing multiplications (3.2). A different approach is outlined in the next subsection.

4.3. Special case: all points z_i lie on a generic circle. Another approach to deal with the preceding special case, that generalizes immediately to the case where the nodes z_i belong to a given circle in the complex plane, $\{z \in \mathbb{C} : |z-p| = r\}$, exploits an invariance property of Cauchy-like matrices under a Moebius transformation of the nodes. Such property is presented in the next lemma for the case of classical Cauchy matrices; the Cauchy-like case can be dealt with by introducing suitable diagonal scalings. With minor changes, all forthcoming arguments also apply to the case where all abscissas lie on a generic line in the complex plane, since the image of $\mathbb{R}$ under a Moebius transformation is either a circle or a line.

LEMMA 4.2. *Let z_i, y_j, for $1 \le i, j \le n$, be pairwise distinct complex numbers, let*

$$\mathcal{M}(z) = \frac{\alpha z + \beta}{\gamma z + \delta}, \quad \alpha\delta - \beta\gamma \neq 0,$$

be a Moebius transformation and let $C_{\mathcal{M}} \equiv (1/(\mathcal{M}(z_i) - \mathcal{M}(y_j)))$. Then $C_{\mathcal{M}}$ is a Cauchy-like matrix with nodes z_i, y_j.

Proof. Using the notations above, we have

$$\frac{1}{\mathcal{M}(z_i) - \mathcal{M}(y_j)} = \frac{1}{\alpha\delta - \beta\gamma} \frac{(\gamma z_i + \delta)(\gamma y_j + \delta)}{z_i - y_j}.$$

Hence $C_{\mathcal{M}}$ has the form $C_{\mathcal{M}} \equiv (u_i v_j/(z_i - y_j))$. □

In the next theorem, we show how to construct a Moebius transformation mapping $\mathbb{R}$ onto a prescribed circle without one point, thus generalizing formula (4.3). Together with the preceding lemma, it will allow us to translate Problem 2 with nodes on a circle into a corresponding problem with real nodes. The latter can be solved with the technique laid down in Subsection 4.1.

THEOREM 4.3. *Let the center of the circle $p \in \mathbb{C}$ and its radius $r > 0$ be given. Consider the following algorithm:*

1. *Choose arbitrary nonzero complex numbers* $\gamma = |\gamma|\mathrm{e}^{\mathrm{i}\theta_\gamma}$ *and* $\delta = |\delta|\mathrm{e}^{\mathrm{i}\theta_\delta}$ *such that* $\mathrm{e}^{2\mathrm{i}(\theta_\gamma-\theta_\delta)} \neq 1$*; moreover, choose* $\tilde\theta \in [0, 2\pi]$.
2. *Set* $\alpha = p\gamma + r|\gamma|\mathrm{e}^{\mathrm{i}\tilde\theta}$.
3. *Set* $\hat\theta = \tilde\theta + \theta_\gamma - \theta_\delta$.
4. *Set* $\beta = p\delta + r|\delta|\mathrm{e}^{\mathrm{i}\hat\theta}$.

Then the function $\mathcal{M}(z) = (\alpha z + \beta)/(\gamma z + \delta)$ *is a Moebius transformation mapping the real line onto the circle* $\{z \in \mathbb{C} : |z - p| = r\}$ *without the point* $\hat z = \alpha/\gamma$.

Proof. After simple manipulations, the equation

$$\left|\frac{\alpha z + \beta}{\gamma z + \delta} - p\right|^2 = r^2$$

leads to the equation

$$z^2|\alpha - p\gamma|^2 + 2z\Re((\alpha - p\gamma)\overline{(\beta - p\delta)}) + |\beta - p\delta|^2 = \tag{4.6}$$

$$= z^2r^2|\gamma|^2 + 2zr^2\Re(\gamma\bar\delta) + r^2|\delta|^2.$$

Here and in the following, $\Re(z)$ denotes the real part of $z \in \mathbb{C}$. By construction, we have $|\alpha - p\gamma| = r|\gamma|$ and $|\beta - p\delta| = r|\delta|$. Moreover,

$$\begin{aligned}\Re((\alpha - p\gamma)\overline{(\beta - p\delta)}) &= r^2|\gamma\delta|\Re(\mathrm{e}^{\mathrm{i}(\tilde\theta-\hat\theta)}) \\ &= r^2|\gamma\delta|\Re(\mathrm{e}^{\mathrm{i}(\theta_\delta-\theta_\gamma)}) \\ &= r^2\Re(\gamma\bar\delta).\end{aligned}$$

Hence equation (4.6) is fulfilled for any real z. The missing point is given by

$$\hat z = \lim_{z\to\infty}\frac{\alpha z + \beta}{\gamma z + \delta} = \frac{\alpha}{\gamma}.$$

It remains to prove that $\alpha\delta - \beta\gamma \neq 0$. Indeed, we have

$$\begin{aligned}\alpha\delta - \beta\gamma &= (p\gamma + r|\gamma|\mathrm{e}^{\mathrm{i}\tilde\theta})\delta - (p\delta + r|\delta|\mathrm{e}^{\mathrm{i}\hat\theta})\gamma \\ &= r|\gamma|\delta\mathrm{e}^{\mathrm{i}\tilde\theta} - r\gamma|\delta|\mathrm{e}^{\mathrm{i}\hat\theta} \\ &= r|\gamma\delta|(\mathrm{e}^{\mathrm{i}(\tilde\theta+\theta_\delta)} - \mathrm{e}^{\mathrm{i}(\hat\theta+\theta_\gamma)}) \\ &= r|\gamma\delta|\mathrm{e}^{\mathrm{i}(\tilde\theta+\theta_\delta)}(1 - \mathrm{e}^{2\mathrm{i}(\theta_\gamma-\theta_\delta)}).\end{aligned}$$

Since $\mathrm{e}^{2\mathrm{i}(\theta_\gamma+\theta_\delta)} \neq 1$ we obtain $\alpha\delta - \beta\gamma \neq 0$. □

Suppose we want to solve Problem 2 with data w_i, z_i, y_i, where $|z_i - p| = r$. As seen from the proof of Theorem 2.2, if we let $C \equiv (w_{i-1}(z_{i-1} - y_0)/(z_{i-1} - y_{j-1}))$ and $C = QR$, then a solution is (Q, S), with $S = Q^H D_z Q - D_y$. Let $\mathcal{M}(z) = (\alpha z + \beta)/(\gamma z + \delta)$ be a Moebius transformation built from Theorem 4.3. Recalling the inversion formula (4.2), let $\tilde z_i = \mathcal{M}^{-1}(z_i)$, $\tilde y_i = \mathcal{M}^{-1}(y_i)$, $v_i = \gamma\tilde y_i + \delta$, and

$$\tilde w_i = w_i\frac{z_i - y_0}{\tilde z_i - \tilde y_0}\frac{\gamma\tilde z_i + \delta}{\alpha\delta - \beta\gamma}, \qquad 0 \le i \le n.$$

Note that $\tilde z_i \in \mathbb{R}$, by construction. From Lemma 4.2, we also have

$$C \equiv \left(\frac{\tilde w_{i-1}(\tilde z_{i-1} - \tilde y_0)v_{j-1}}{\tilde z_{i-1} - \tilde y_{j-1}}\right).$$

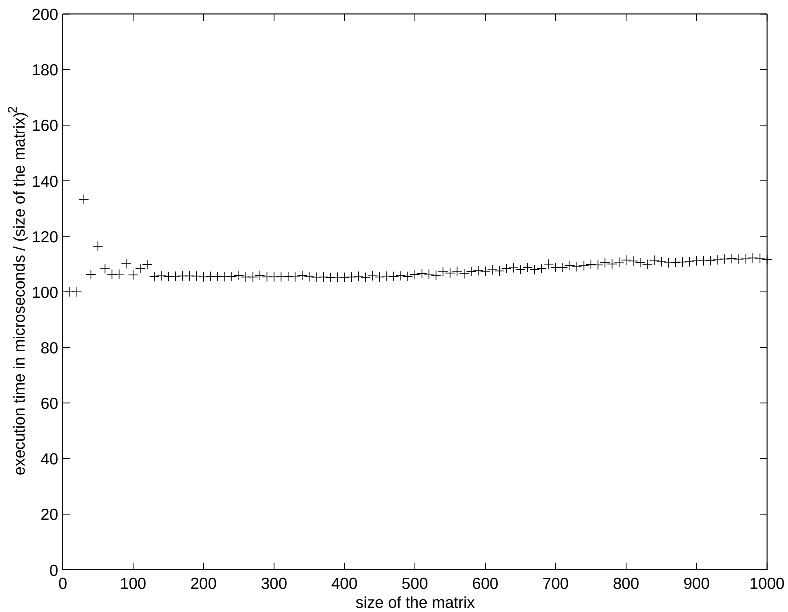

FIG. 5.1. *Computational complexity*

Again from Theorem 2.2, we see that the solution of Problem 2 with data $\tilde{w}_i, \tilde{z}_i, \tilde{y}_i$ is $(Q, \widetilde{S})$ where

$$\widetilde{S} = Q^H \mathcal{M}^{-1}(D_z) Q - \mathcal{M}^{-1}(D_y).$$

Let $\widehat{S} = \widetilde{S} + \mathcal{M}^{-1}(D_y)$. Observe that $\widehat{S}$ is a *diagonal-plus-semiseparable matrix* [10, 12, 15]. After simple passages, we have

$$S = \mathcal{M}(\widehat{S}) - D_y = [\alpha \widehat{S} + \beta I][\gamma \widehat{S} + \delta I]^{-1} - D_y.$$

Hence S can be recovered from $\widetilde{S}$ by determining the entries in its first and last rows and columns. This latter task can be carried out at a linear cost by means of several different algorithms for the solution of diagonal-plus-semiseparable linear systems. See, e.g., [10, 12, 23, 35].

5. Numerical experiments. In this section, we show the numerical behaviour of the solution of the inverse eigenvalue problem for some real points z_i and real poles y_i, $i = 0, 1, \ldots, n$. The points are $z_i = i + n$, for $i = 0, 1, 2, \ldots, n$, with corresponding weights $w_i = 1$. The poles are $y_i = i + n - \frac{1}{2}$. We implemented the $\mathcal{O}(n^2)$ algorithm in Matlab on a PC running at 833 MHz and having 512MB of RAM. To show that the algorithm is indeed $\mathcal{O}(n^2)$, we plot in Figure 5.1 the execution time divided by n^2 for the different sizes of the problem. Here we set $n = 10, 20, 30, \ldots, 1000$. Figure 5.2 gives the maximum relative error on the eigenvalues of the computed diagonal-plus-semiseparable matrix compared to the original points z_i for $n = 10, 20, 30, \ldots, 500$. In Figure 5.3, the same is done for the weights. Figures 5.2 and 5.3 show that the algorithm is accurate for this specific data set. We have tried other data sets resulting in less accurate results. It seems much depends on the lay-out of the poles y_i with respect to the points z_i. More research has to be done to develop a robust and accurate algorithm.

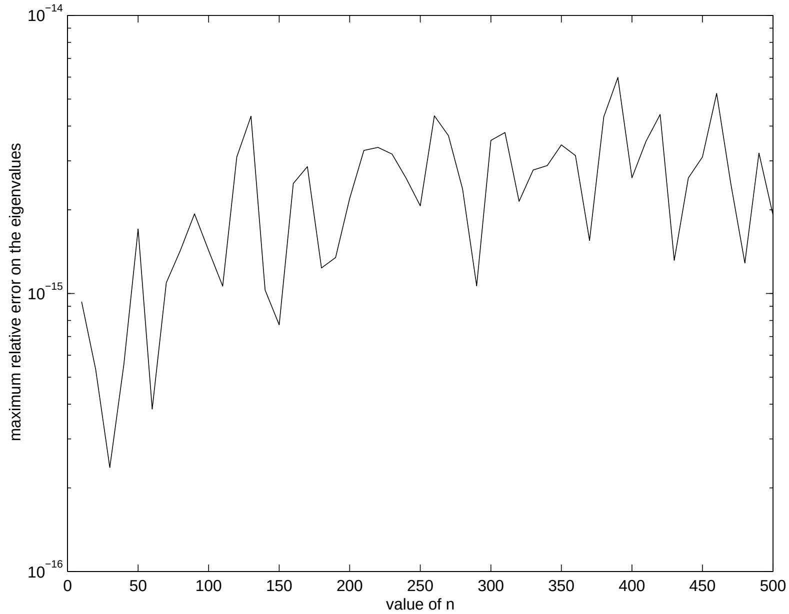

Fig. 5.2. *Relative accuracy of the eigenvalues*

6. Conclusions and further developments. In this paper, we have shown that solving a certain inverse eigenvalue problem gives all the information necessary to construct orthogonal rational functions in an efficient way. The algorithm which we developed here, gives accurate results for a lot of data sets but we found examples for which the algorithm does not perform so well. Further research is necessary to identify if for these data sets the problem is ill-conditioned or the algorithm is numerically not stable.

REFERENCES

[1] G. Ammar and W. Gragg, *$O(n^2)$ reduction algorithms for the construction of a band matrix from spectral data*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 426–431.

[2] G. Ammar, W. Gragg, and L. Reichel, *Constructing a unitary Hessenberg matrix from spectral data*, in Numerical linear algebra, digital signal processing and parallel algorithms, G. Golub and P. Van Dooren, eds., vol. 70 of NATO-ASI Series, F: Computer and Systems Sciences, Berlin, 1991, Springer-Verlag, pp. 385–395.

[3] D. Boley and G. Golub, *A survey of matrix inverse eigenvalue problems*, Inverse Problems, 3 (1987), pp. 595–622.

[4] A. Bultheel and M. V. Barel, *Padé techniques for model reduction in linear system theory*, J. Comput. Appl. Math., 14 (1986), pp. 401–438.

[5] A. Bultheel and B. De Moor, *Rational approximation in linear systems and control*, J. Comput. Appl. Math., 121 (2000), pp. 355–378.

[6] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad, *Orthogonal rational functions with poles on the unit circle*, J. Math. Anal. Appl., 182 (1994), pp. 221–243.

[7] ———, *Orthogonal rational functions*, vol. 5 of Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 1999.

[8] ———, *Quadrature and orthogonal rational functions*, J. Comput. Appl. Math., 127 (2001), pp. 67–91.

[9] A. Bultheel and M. Van Barel, *Vector orthogonal polynomials and least squares approximation*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 863–885.

[10] S. Chandrasekaran and M. Gu, *A fast and stable solver for recursively semi-separable sys-*

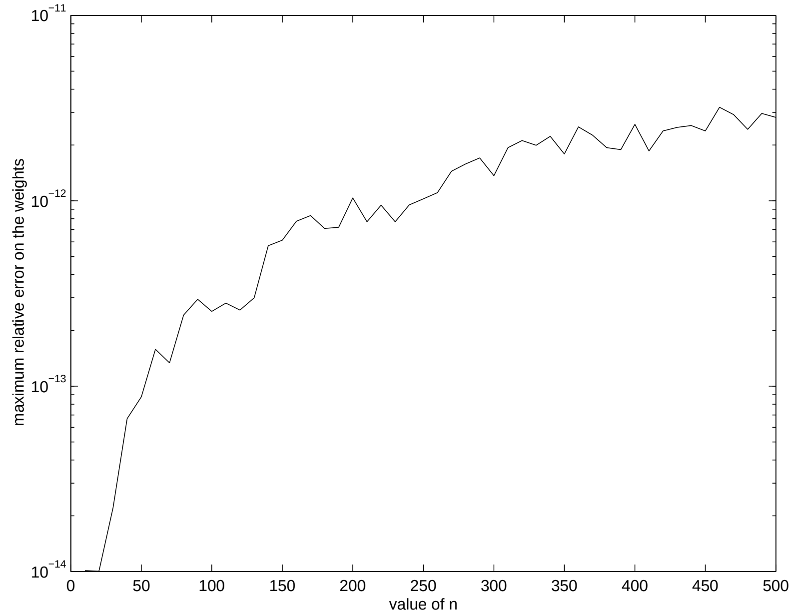

Fig. 5.3. *Relative accuracy of the weights*

tems of linear equations, in Structured matrices in mathematics, computer science, and engineering, II (Boulder, CO, 1999), vol. 281 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2001, pp. 39–53.

[11] P. Delsarte, Y. Genin, and Y. Kamp, *On the role of the Nevanlinna-Pick problem in circuit and system theory*, Internat. J. Circuit Theory Appl., 9 (1981), pp. 177–187.

[12] Y. Eidelman and I. Gohberg, *A look-ahead block Schur algorithm for diagonal plus semiseparable matrices*, Comput. Math. Appl., 35 (1998), pp. 25–34.

[13] S. Elhay, G. Golub, and J. Kautsky, *Updating and downdating of orthogonal polynomials with data fitting applications*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 327–353.

[14] D. Fasino and L. Gemignani, *A Lanczos-type algorithm for the QR factorization of Cauchy-like matrices*. To appear on Contemporary Mathematics.

[15] ———, *Direct and inverse eigenvalue problems for diagonal-plus-semiseparable matrices*. Submitted, 2001.

[16] ———, *A Lanczos-type algorithm for the QR factorization of regular Cauchy matrices*, Numer. Linear Algebra Appl., 9 (2002), pp. 305–319.

[17] B. Fischer and G. Golub, *How to generate unknown orthogonal polynomials out of known orthogonal polynomials*, J. Comput. Appl. Math., 43 (1992), pp. 99–115.

[18] F. R. Gantmacher and M. G. Kreĭn, *Oszillationsmatrizen, Oszillationskerne und kleine Schwingungen mechanischer Systeme*, Akademie-Verlag, Berlin, 1960.

[19] W. Gautschi, *The use of rational functions in numerical quadrature*, J. Comput. Appl. Math., 133 (2001), pp. 111–126.

[20] W. Gragg and A. Lindquist, *On the partial realization problem*, Linear Algebra Appl., 50 (1983), pp. 277–319.

[21] W. B. Gragg and W. J. Harrod, *The numerically stable reconstruction of Jacobi matrices from spectral data*, Numer. Math., 44 (1984), pp. 317–335.

[22] P. Henrici, *Applied and Computational Complex Analysis*, vol. 1, Wiley, 1974.

[23] I. Koltracht, *Linear complexity algorithm for semiseparable matrices*, Integral Equations Operator Theory, 29 (1997), pp. 313–319.

[24] B. Ninness and F. Gustafsson, *A unifying construction of orthonormal bases for system identification*, IEEE Trans. Automat. Control, 42 (1997), pp. 515–521.

[25] V. Olshevsky and V. Pan, *Polynomial and rational evaluation and interpolation (with structured matrices)*, in Automata, languages and programming, vol. 1644 of Lecture Notes in Comput. Sci., Springer, Berlin, 1999, pp. 585–594.

[26] L. Reichel, *Fast QR decomposition of Vandermonde-like matrices and polynomial least squares approximation*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 552–564.

[27] ———, *Construction of polynomials that are orthogonal with respect to a discrete bilinear form*, Adv. Comput. Math., 1 (1993), pp. 241–258.

[28] L. Reichel, G. Ammar, and W. Gragg, *Discrete least squares approximation by trigonometric polynomials*, Math. Comp., 57 (1991), pp. 273–289.

[29] H. Rutishauser, *On Jacobi rotation patterns*, in Proceedings of Symposia in Applied Mathematics, vol. 15, Experimental Arithmetic, High Speed Computing and Mathematics, Providence, 1963, Amer. Math. Society, pp. 219–239.

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

[31] ———, *A parallel algorithm for discrete least squares rational approximation*, Numer. Math., 63 (1992), pp. 99–121.

[32] M. Van Barel and A. Bultheel, *Discrete linearized least-squares rational approximation on the unit circle*, J. Comput. Appl. Math., 50 (1994), pp. 545–563.

[33] M. Van Barel and A. Bultheel, *Discrete linearized least-squares rational approximation on the unit circle*, J. Comput. Appl. Math., 50 (1994), pp. 545–563.

[34] ———, *Orthonormal polynomial vectors and least squares approximation for a discrete inner product*, Electronic Transactions on Numerical Analysis, 3 (1995), pp. 1–23.

[35] E. Van Camp, N. Mastronardi, and M. Van Barel, *Two fast algorithms for solving diagonal-plus-semiseparable linear systems*, Tech. Rep. 17/2002, Instituto per le Applicazioni del Calcolo "M. Picone", Sez. Bari (Consiglio Nazionale delle Ricerche), Italy, Aug. 2002.