Orthogonal Rational Functions with real coefficients and semiseparable matrices

A. Bultheel [a,1,2] P. Van gucht [b] M. Van Barel [a,1]

[a] *Department of Computer Science, K.U.Leuven, Celestijnenlaan 200 A, 3001 Heverlee, Belgium.*
[b] *Fluxys, Belgium.*

Abstract

When one wants to use Orthogonal Rational Functions (ORFs) in system identification or control theory, it is important to be able to avoid complex calculations. In this paper we study ORFs whose numerator and denominator polynomial have real coefficients. These ORFs with real coefficients (RORFs) appear when the poles and the interpolation points appear in complex conjugate pairs, which is a natural condition. Further we deduce that there is a strong connection between RORFs and semiseparable matrices.

Key words: orthogonal rational functions, semiseparable matrix, inverse eigenvalue problem, system identification
1991 MSC: 65F30, 65F18, 93B30

To Bill Gragg on the occasion of his 70th birthday

[1] This work is partially supported by the Research Foundation Flanders (FWO), project "RAM: Rational modelling: optimal conditioning and stable algorithms", grant #G.0423.05, the K.U.Leuven Research Fund projects "Large rank structured matrix computations", grant OT/05/16, and "OPTEC: Center of Excellence: Optimisation in engineering", and the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, project "DYSCO: Dynamical Systems and Control". The scientific responsibility rests with the authors.
[2] Email of the corresponding author: `Adhemar.Bultheel@cs.kuleuven.be`

1 Introduction

In system identification rational functions have been widely used and have proven to be useful (see e.g. [8,9,3]). For numerical stability it is better to use orthogonal rational functions, as proposed in [6]. In [2] we introduced a new point of view by looking at orthogonal rational functions as solutions of an inverse eigenvalue problem of a semiseparable plus diagonal matrix, which was first shown in [5]. In this approach the matrices involved are in general complex and this is not interesting in system identification and control theory. We show in this paper that it is possible to do all the calculations with real numbers, only by imposing some natural conditions on the given data. These conditions come from the fact that we want to approximate the transfer function $G(z)$ by a linear combination of ORFs. If the transfer function comes from a *real* problem, then $G(e^{\mathbf{i}\theta}) = G^*(e^{-\mathbf{i}\theta})$, which will lead to ORFs with real coefficients. Section 2 deals with the problem of existence of these RORFs. In Section 3 the relation with semiseparable matrices is fully investigated and in Sections 4 and 5 we look in detail at the structure of the matrices involved. For related work see [4].

We will use the notation $\mathbb{T}$ for the unit circle, $\mathbb{D}$ for the open unit disc and $\mathbb{E}$ for the exterior of the unit disc. $\mathbb{P}_k$ will denote the space of polynomials of degree at most k. By $\mathbf{i}$ we denote the imaginary unit and by $\mathbf{I}_k$ the $k \times k$ unit matrix. The transpose of a matrix A is denoted by A^T. The complex conjugate (i.e., Hermitian) transpose of a matrix A is denoted by A^* and hence it is just the complex conjugate when A is a scalar. For a function, the star means $f^*(z) = [f(z)]^*$. A function f of a complex variable z is said to be real if $f(z) + f(z^*)$ is real for all z, i.e., $f(z^*) = f^*(z)$.

2 Orthogonal Rational Functions with Real Coefficients

In this section we shall show that there exists a basis of ORFs with real numerator and denominator coefficients (RORFs) if we impose some natural restrictions on the poles and the interpolation points.

Let $\{z_k\}_{k=1}^N$ be a set of points on the upper half of $\mathbb{T}$, i.e. $z_k = e^{i\theta_k}$ with $0 < \theta_k < \pi$ (we exclude the points ± 1 for simplicity, but they can be included very easily). Furthermore, let $\{w_k\}_{k=1}^N$ be a set of strictly positive numbers. Consider a set of poles $\mathbb{A}_n = \{\alpha_k\}_{k=1}^N \subset \mathbb{D}$. Define

$$\ell_k(z) = (z - \alpha_k), \qquad \tilde{\ell}_k(z) = (z - \alpha_k^*);$$

and the nodal polynomials

$$\pi_1(z) = \ell_1(z), \text{ and } \tilde{\pi}_k(z) = \pi_k(z)\tilde{\ell}_k(z), \quad \pi_{k+1}(z) = \tilde{\pi}_k(z)\ell_{k+1}(z), \quad k = 1, 2, \dots$$

Note that $\tilde{\pi}_k(z) = \prod_{i=1}^{k}(z^2 - 2\Re(\alpha_i)z + |\alpha_i|^2)$ are real functions, collecting the pairs of complex conjugate poles.

Now consider for $n = 1, \dots, N$ the spaces

$$\begin{aligned}
\mathcal{L}_n &= \operatorname{span}\left\{\frac{1}{\pi_k(z)}, \frac{1}{\tilde{\pi}_l(z)} : k = 1, \dots, n; l = 1 \dots, n-1\right\} = \left\{\frac{q_{2n-2}}{\pi_n(z)} : q_{2n-2} \in \mathbb{P}_{2n-2}\right\} \\
\tilde{\mathcal{L}}_n &= \operatorname{span}\left\{\frac{1}{\pi_k(z)}, \frac{1}{\tilde{\pi}_k(z)} : k = 1, 2, \dots, n\right\} = \left\{\frac{q_{2n-1}}{\tilde{\pi}_n(z)} : q_{2n-1} \in \mathbb{P}_{2n-1}\right\},
\end{aligned}$$

which are equipped with the inner product

$$\langle f, g\rangle = \sum_{k=1}^{N} w_k^2\Big(f(z_k)g^*(z_k) + f(z_k^*)g^*(z_k^*)\Big). \tag{2.1}$$

Note that this inner product imposes the same weight on the points z_i and z_i^*. By the Gram-Schmidt orthogonalization procedure we can construct orthogonal basis functions $\phi_k(z)$, $\tilde{\phi}_k(z)$, $k = 1, 2, \dots, N$ such that

$$\begin{aligned}
\mathcal{L}_n &= \operatorname{span}\left\{\phi_k(z), \tilde{\phi}_l(z) : k = 1, 2, \dots, n; l = 1, \dots, n-1\right\} \\
\tilde{\mathcal{L}}_n &= \operatorname{span}\left\{\phi_k(z), \tilde{\phi}_k(z) : k = 1, 2, \dots, n\right\},
\end{aligned}$$

where $\phi_n(z) = p_{2n-2}(z)/\pi_n(z) \in \mathcal{L}_n \setminus \tilde{\mathcal{L}}_{n-1}$ and $\tilde{\phi}_n(z) = p_{2n-1}(z)/\tilde{\pi}_n(z) \in \tilde{\mathcal{L}}_n \setminus \mathcal{L}_n$. However, since $\{1/\pi_k(z), 1/\tilde{\pi}_k(z)\}$ and $\{1/\tilde{\pi}_k(z), z/\tilde{\pi}_k(z)\}$ span the same space, it follows that we may write

$$\tilde{\mathcal{L}}_n = \operatorname{span}\left\{\frac{1}{\tilde{\pi}_k(z)}, \frac{z}{\tilde{\pi}_k(z)} : k = 1, 2, \dots, n\right\}.$$

By applying the Gram-Schmidt procedure to this basis, we obtain another set of orthogonal basis functions $\psi_k(z)$, $\tilde{\psi}_k(z)$ such that

$$\tilde{\mathcal{L}}_n = \operatorname{span}\left\{\psi_k(z), \tilde{\psi}_k(z) : k = 1, 2, \dots, n\right\},$$

and ψ_k, $\tilde{\psi}_k \in \tilde{\mathcal{L}}_k \setminus \tilde{\mathcal{L}}_{k-1}$. It is crucial to note that this basis $\left\{\psi_k,\ \tilde{\psi}_k : k = 1, 2, \ldots, N\right\}$ consists of RORFs, i.e. the numerator and denominator of these ORFs have real coefficients since indeed both $1/\tilde{\pi}_k(z)$ and $z/\tilde{\pi}_k(z)$ are real functions and all the inner products in the Gram-Schmidt procedure give real numbers.

Further, by the equivalence of orthogonal bases there must exist unitary matrices $\mathbf{p}_k$ such that

$$\left[\psi_k(z)\ \tilde{\psi}_k(z)\right] = \left[\phi_k(z)\ \tilde{\phi}_k(z)\right] \mathbf{p}_k.$$

Thus we have proved

Theorem 2.1 *The functions* $\left\{\psi_k,\ \tilde{\psi}_k : k = 1, \ldots, n\right\}$ *constitute RORFs for* $\tilde{\mathcal{L}}_n$ *and they are related to the ORFs* $\left\{\phi_k,\ \tilde{\phi}_k : k = 1, \ldots, n\right\}$ *by relations of the form*

$$\left[\psi_k(z)\ \tilde{\psi}_k(z)\right] = \left[\phi_k(z)\ \tilde{\phi}_k(z)\right] \mathbf{p}_k, \qquad k = 1,\ \ldots,\ n,$$

where the $\mathbf{p}_k$ *are unitary matrices.*

Remark 2.2 Note that the functions ϕ_n and $\tilde{\phi}_n$ are unique up to a constant factor of modulus 1. The pair of functions $(\psi_n, \tilde{\psi}_n)$ is unique up to a (real) orthogonal transformation. Indeed, any (real) orthogonal basis of $\tilde{\mathcal{L}}_n \setminus \tilde{\mathcal{L}}_{n-1}$ will do.

3 Semiseparable matrices and an inverse eigenvalue problem

In this section we deduce a connection between RORFs and semiseparable matrices. We start with the following lemma.

Lemma 3.1 *Suppose that* $\phi_k(z) = p_{2k-2}(z)/\pi_k(z)$ *and* $\tilde{\phi}_k(z) = p_{2k-1}(z)/\tilde{\pi}_k(z)$ *are the orthonormal basis functions as above. Then*

$$\phi_k(z) \in \operatorname{span}\{(z - \alpha_i)\phi_i(z), (z - \alpha_i^*)\tilde{\phi}_i(z) : i = 1, \ldots, k\}$$
$$\tilde{\phi}_k(z) \in \operatorname{span}\{(z - \alpha_i)\phi_i(z), (z - \alpha_j^*)\tilde{\phi}_j(z) : i = 1, \ldots, k+1; j = 1, \ldots, k\}$$

This lemma is essentially the same as Lemma 3.3 in [2]. Note that in the original lemma, there was an extra condition that $p_{2k-2}(\alpha_{k-1}^*) \neq 0$ and $p_{2k-1}(\alpha_k) \neq 0$, but since all the zeros of the ORFs lie in $\mathbb{E}$ (see [1, Corollary 3.1.4]), this can never happen.

Along the lines of [2] adopted to our particular situation, we can reformulate the previous lemma in matrix form. The basic steps are as follows. Let $\boldsymbol{\phi}_i = [\phi_i, \tilde{\phi}_i]$ and $\boldsymbol{\varphi}_n = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_n]$ and $\mathbf{A}_n = \mathrm{diag}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n)$ with ith diagonal block $\boldsymbol{\alpha}_i = \mathrm{diag}(\alpha_i, \alpha_i^*)$. Then there must exist some (block) upper Hessenberg matrix $\mathbf{H}_n \in \mathbb{C}^{2n \times 2n}$ and some number $h_{2n+1,2n}$ such that

$$\boldsymbol{\varphi}_n(z) = [z\boldsymbol{\varphi}_n(z) - \boldsymbol{\varphi}_n(z)\mathbf{A}_n]\mathbf{H}_n + [0, \ldots, 0, \phi_{n+1}(z)(z - \alpha_{n+1})]h_{2n+1,2n}.$$

Moreover we know from Lemma 3.3 in [2] that the subdiagonal elements of $\mathbf{H}_n$ are nonzero.

Since there are only $2N$ points $\{z_i, z_i^*\}$ used in the inner products, we shall in the Gram-Schmidt procedure eventually arrive at $\|\phi_{N+1}\| = 0$ which means that ϕ_{N+1} vanishes in all the points $\{z_i, z_i^*\}_{i=1}^N$. Now we can write out the previous relation for all $z \in \{z_i, z_i^*\}_{i=1}^N$ which will lead to the following matrix relations.

Let us denote by $\boldsymbol{\Phi}$ the $2N \times 2N$ matrix whose (i,j)th 2×2-block element is

$$\begin{bmatrix} \phi_j(z_i) & \tilde{\phi}_j(z_i) \\ \phi_j(z_i^*) & \tilde{\phi}_j(z_i^*) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}_j(z_i) \\ \boldsymbol{\phi}_j(z_i^*) \end{bmatrix}.$$

Set $\mathbf{A} = \mathbf{A}_N$, $\mathbf{H} = \mathbf{H}_N$, and by $\mathbf{Z}$ we denote the block diagonal matrix $\mathrm{diag}(\mathbf{z}_1, \ldots, \mathbf{z}_N)$ with blocks $\mathbf{z}_i = \mathrm{diag}(z_i, z_i^*)$. We then have as in [2]:

$$\boldsymbol{\Phi} = (\mathbf{Z}\boldsymbol{\Phi} - \boldsymbol{\Phi}\mathbf{A})\mathbf{H}. \tag{3.1}$$

By defining $\mathbf{W} = \mathrm{diag}(\mathbf{w}_1, \ldots, \mathbf{w}_N)$ with $\mathbf{w}_i = w_i\mathbf{I}_2$ and $\mathbf{Q} = \mathbf{W}\boldsymbol{\Phi}$, and multiplying equation (3.1) on the left with $\mathbf{W}$, we get the relation

$$\mathbf{Q} = (\mathbf{Z}\mathbf{Q} - \mathbf{Q}\mathbf{A})\mathbf{H}. \tag{3.2}$$

The essence of this paper is devoted to transforming this relation involving complex-valued matrices into a relation involving only real-valued matrices.

Multiplication on the right by the block diagonal unitary matrix $\mathbf{P} = \mathrm{diag}(\mathbf{p}_1,\ \mathbf{p}_2,\ \ldots,\ \mathbf{p}_N)$, where $\mathbf{p}_k$ are the unitary matrices found in Theorem 2.1, gives

$$\mathbf{QP} = (\mathbf{ZQ} - \mathbf{QA})\mathbf{HP} = (\mathbf{ZQP} - \mathbf{QPP}^*\mathbf{AP})\mathbf{P}^*\mathbf{HP}. \tag{3.3}$$

The matrix $\mathbf{QP}$ consists of the blocks

$$\mathbf{w}_i \begin{bmatrix} \boldsymbol{\psi}_j(z_i) \\ \boldsymbol{\psi}_j(z_i^*) \end{bmatrix} = w_i \begin{bmatrix} \psi_j(z_i) & \tilde{\psi}_j(z_i) \\ \psi_j(z_i^*) & \tilde{\psi}_j(z_i^*) \end{bmatrix} = w_i \begin{bmatrix} \psi_j(z_i) & \tilde{\psi}_j(z_i) \\ \psi_j^*(z_i) & \tilde{\psi}_j^*(z_i) \end{bmatrix} = \mathbf{w}_i \begin{bmatrix} \boldsymbol{\psi}_j(z_i) \\ \boldsymbol{\psi}_j^*(z_i) \end{bmatrix}.$$

We can transform $\mathbf{QP}$ into a real-valued matrix if we multiply it on the left by a unitary block diagonal matrix $\mathbf{G}$ with blocks

$$\mathbf{g} = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{bmatrix}. \tag{3.4}$$

The equation (3.3) now becomes

$$\mathbf{GQP} = (\mathbf{GZQP} - \mathbf{GQPP}^*\mathbf{AP})\mathbf{P}^*\mathbf{HP} = (\mathbf{GZG}^*\mathbf{GQP} - \mathbf{GQPP}^*\mathbf{AP})\mathbf{P}^*\mathbf{HP}.$$

Setting $\mathbf{GQP} = \hat{\mathbf{Q}}$, $\mathbf{GZG}^* = \hat{\mathbf{Z}}$, $\mathbf{P}^*\mathbf{AP} = \hat{\mathbf{A}}$, and $\mathbf{P}^*\mathbf{HP} = \hat{\mathbf{H}}$, we get

$$\hat{\mathbf{Q}} = (\hat{\mathbf{Z}}\hat{\mathbf{Q}} - \hat{\mathbf{Q}}\hat{\mathbf{A}})\hat{\mathbf{H}}, \tag{3.5}$$

which resembles (3.2). Note that in this equation the matrices $\hat{\mathbf{Q}}$ and $\hat{\mathbf{Z}}$ are real.

The inverse of the upper Hessenberg matrix $\mathbf{H}$ is a lower semiseparable matrix $\mathbf{S}$ of rank 1 (see [2]), that is all its submatrices whose upper right corner is on the main diagonal have at most rank 1. In the rest of this paper we call this a semiseparable matrix for short. Since the matrix $\mathbf{P}$ is unitary, the inverse of the matrix $\hat{\mathbf{H}}$ is $\mathbf{P}^*\mathbf{SP}$. Multiplying equation (3.5) on the right with this inverse gives $\hat{\mathbf{Q}}\mathbf{P}^*\mathbf{SP} = \hat{\mathbf{Z}}\hat{\mathbf{Q}} - \hat{\mathbf{Q}}\hat{\mathbf{A}}$ or $\hat{\mathbf{Q}}(\mathbf{P}^*\mathbf{SP} + \hat{\mathbf{A}}) = \hat{\mathbf{Z}}\hat{\mathbf{Q}}$. Introducing the notation $\tilde{\mathbf{S}} = \mathbf{P}^*\mathbf{SP}$ we find

$$\tilde{\mathbf{S}} + \hat{\mathbf{A}} = \hat{\mathbf{Q}}^*\hat{\mathbf{Z}}\hat{\mathbf{Q}}. \tag{3.6}$$

Remark 3.2 The above relation is equivalent with $(\mathbf{S} + \mathbf{A})\mathbf{Q}^* = \mathbf{Q}^*\mathbf{Z}$, which illustrates that the columns of $\mathbf{Q}^*$ are eigenvectors of $\mathbf{S} + \mathbf{A}$ whose eigenvalues are the diagonal elements of $\mathbf{Z}$.

So the problem is an inverse eigenvalue problem: given the eigenvalues in $\mathbf{Z}$ of a matrix, find its orthonormal eigenvectors such that the matrix is of the form diagonal-plus-semiseparable of rank 1, with prescribed diagonal $\mathbf{A}$. To fix a solution, one needs some extra conditions which fix the first column of $\mathbf{Q}$ as proportional to the vector $[w_1, w_1, w_2, w_2, \ldots, w_N, w_N]^T$ (see [2]).

Remark 3.3 Note that the right hand side of equation (3.6) is a real matrix, so the left hand side should be real as well. Therefore the off-diagonal blocks of $\tilde{\mathbf{S}}$ are all real. Since $\hat{\mathbf{A}}$ is complex in general, also the diagonal blocks of $\tilde{\mathbf{S}}$ will be complex. Thus our next objective is to rewrite $\tilde{\mathbf{S}}+\hat{\mathbf{A}}$ as a *real* block semiseparable plus a *real* block diagonal. Therefore we split $\tilde{\mathbf{S}}$ as $\tilde{\mathbf{S}} = \hat{\mathbf{S}}+\mathbf{D}$ with $\hat{\mathbf{S}}$ a real block semiseparable matrix and $\mathbf{D}$ block diagonal such that $\hat{\mathbf{B}} = \mathbf{D} + \hat{\mathbf{A}}$ is real.

4 The existence of a real formulation

As we said in the beginning, we shall restrict ourselves in this paper to the case where the z_k are on the unit circle. This implies that not only the block lower triangular part of the semiseparable matrix has the rank one property, but also all the submatrices of the complementary upper block triangular will have rank one if none of the poles is zero (which, for simplicity, will also be assumed in the rest of this paper). See [2, Corallary 3.7].

In this section we will rewrite the left hand side of (3.6) i.e. $\tilde{\mathbf{S}} + \hat{\mathbf{A}}$, as $\hat{\mathbf{S}} + \hat{\mathbf{B}}$ with $\hat{\mathbf{S}}$ a block semiseparable and $\hat{\mathbf{B}}$ a block diagonal which are both real.

The matrices $\mathbf{p}_i$, found in Theorem 2.1 are unitary 2×2 matrices and so they have the following form with $\nu_i, \mu_i, \theta_i \in \mathbb{R}$ (see appendix)

$$\mathbf{p}_i = e^{\mathbf{i}\theta_i} \begin{bmatrix} \cos(t_i) & -\sin(t_i)e^{\mathbf{i}\gamma_i} \\ \sin(t_i)e^{\mathbf{i}\mu_i} & \cos(t_i)e^{\mathbf{i}\nu_i} \end{bmatrix}, \quad \gamma_i = \nu_i - \mu_i, \quad t_i \in \left[0, \frac{\pi}{2}\right]. \tag{4.1}$$

Since the matrix $\mathbf{S}$ is the inverse of a Hessenberg matrix with nonzero subdiagonals, it is semiseparable of rank 1 and it is generator-representable. This means that there exist vectors

$\mathbf{u}$, $\mathbf{v}$, $\mathbf{x}$ and $\mathbf{y}$, such that

$$\mathbf{S} = \begin{bmatrix} u_1 v_1^* & x_1\tilde{y}_1^* & \cdots & x_1 y_N^* & x_1\tilde{y}_N^* \\ \tilde{u}_1 v_1^* & \tilde{u}_1\tilde{v}_1^* & \cdots & \tilde{x}_1 y_N^* & \tilde{x}_1\tilde{y}_N^* \\ u_2 v_1^* & u_2\tilde{v}_1^* & \cdots & x_2 y_N^* & x_2\tilde{y}_N^* \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ u_N v_1^* & u_N\tilde{v}_1^* & \cdots & u_N v_N^* & x_N\tilde{y}_N^* \\ \tilde{u}_N v_1^* & \tilde{u}_N\tilde{v}_1^* & \cdots & \tilde{u}_N v_N^* & \tilde{u}_N\tilde{v}_N^* \end{bmatrix},$$

where $\mathbf{u} = [u_1,\ \tilde{u}_1,\ u_2,\ \tilde{u}_2,\ \ldots,\ u_N,\ \tilde{u}_N]^T$ and similarly for $\mathbf{v}$, $\mathbf{x}$ and $\mathbf{y}$.

Let $\mathbf{s}_{ij}$ be the (i,j)th 2×2 block of $\mathbf{S}$, then

$$\begin{aligned} \mathbf{s}_{ij} &= \mathbf{u}_i \mathbf{v}_j^*, \quad \mathbf{u}_i = \begin{bmatrix} u_i \\ \tilde{u}_i \end{bmatrix}, \quad \mathbf{v}_i = \begin{bmatrix} v_i \\ \tilde{v}_i \end{bmatrix}, \quad \text{if } i > j \\ &= \mathbf{x}_i \mathbf{y}_j^*, \quad \mathbf{x}_i = \begin{bmatrix} x_i \\ \tilde{x}_i \end{bmatrix}, \quad \mathbf{y}_i = \begin{bmatrix} y_i \\ \tilde{y}_i \end{bmatrix}, \quad \text{if } i < j \\ \mathbf{S}_{ii} &= \mathbf{u}_i \mathbf{v}_i^* + \begin{bmatrix} 0 & \Delta_i \\ 0 & 0 \end{bmatrix}, \quad \text{with } \Delta_i = x_i\tilde{y}_i^* - u_i\tilde{v}_i^* \text{ and } \mathbf{u}_i \text{ and } \mathbf{v}_i \text{ as above.} \end{aligned}$$

If we put

$$\begin{bmatrix} \mathbf{U}_i & \mathbf{V}_i & \mathbf{X}_i & \mathbf{Y}_i \end{bmatrix} = \mathbf{p}_i^* \begin{bmatrix} \mathbf{u}_i & \mathbf{v}_i & \mathbf{x}_i & \mathbf{y}_i, \end{bmatrix} \tag{4.2}$$

for $i = 1, \ldots, N$, then we find that $\tilde{\mathbf{S}} = \hat{\mathbf{S}} + \hat{\mathbf{D}}$ with

$$\hat{\mathbf{S}} = \begin{bmatrix} \mathbf{U}_1\mathbf{V}_1^* & \mathbf{X}_1\mathbf{Y}_2^* & \cdots & \mathbf{X}_1\mathbf{Y}_N^* \\ \mathbf{U}_2\mathbf{V}_1^* & \mathbf{U}_2\mathbf{V}_2^* & \cdots & \mathbf{X}_2\mathbf{Y}_N^* \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{U}_N\mathbf{V}_1^* & \mathbf{U}_N\mathbf{V}_2^* & \cdots & \mathbf{U}_N\mathbf{V}_N^* \end{bmatrix} = \left[\begin{array}{cc|c|cc} U_1V_1^* & U_1\tilde{V}_1^* & \cdots & X_1Y_N^* & X_1\tilde{Y}_N^* \\ \tilde{U}_1V_1^* & \tilde{U}_1\tilde{V}_1^* & \cdots & \tilde{X}_1Y_N^* & \tilde{X}_1\tilde{Y}_N^* \\ \hline U_2V_1^* & U_2\tilde{V}_1^* & \cdots & X_2Y_N^* & X_2\tilde{Y}_N^* \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline U_NV_1^* & U_N\tilde{V}_1^* & \cdots & U_NV_N^* & U_N\tilde{V}_N^* \\ \tilde{U}_NV_1^* & \tilde{U}_N\tilde{V}_1^* & \cdots & \tilde{U}_NV_N^* & \tilde{U}_N\tilde{V}_N^* \end{array}\right] \tag{4.3}$$

and $\hat{\mathbf{D}}$ is a block diagonal matrix with 2×2 blocks

$$\hat{\mathbf{d}}_i = \mathbf{p}_i^* \begin{bmatrix} 0 & \Delta_i \\ 0 & 0 \end{bmatrix} \mathbf{p}_i, \quad \Delta_i = x_i\tilde{y}_i^* - u_i\tilde{v}_i^*.$$

Note that $\hat{\mathbf{S}}$ is a block semiseparable matrix of rank one. If we denote $\Delta_i = d_i e^{\mathbf{i}\delta_i}$, it is easy to see that the diagonal blocks of the matrix $\hat{\mathbf{D}}$ are equal to

$$\hat{\mathbf{d}}_i = d_i e^{\mathbf{i}\delta_i} \begin{bmatrix} \sin(t_i)\cos(t_i)e^{\mathbf{i}\mu_i} & \cos^2(t_i)e^{\mathbf{i}\nu_i} \\ -\sin^2(t_i)e^{\mathbf{i}(2\mu_i-\nu_i)} & -\sin(t_i)\cos(t_i)e^{\mathbf{i}\mu_i} \end{bmatrix}. \tag{4.4}$$

Since $\mathbf{S}$ is invertible, the number $\tilde{u}_N$ is different from zero. By their definition (4.2), it is easy to see that at least one of the numbers U_N or $\tilde{U}_N$ is different from zero. Suppose that U_N is different from zero (the reasoning is the same for $\tilde{U}_N$ different from zero). Since $\mathbf{u}$ is unique up to a (complex) nonzero constant factor c due to the definition of semiseparable matrices (we can multiply $\mathbf{u}$ by c and divide $\mathbf{v}$ by c), we can choose this constant c (for example $c \in \mathbb{T}$) such that U_N is real. Now, if U_N is real, then the numbers $V_1, \ldots, \tilde{V}_{N-1}$ have to be real, since $\hat{\mathbf{S}}$ is real on the off diagonal blocks (due to Remark 3.3 at the end of the previous section). Furthermore, if $V_1, \ldots, \tilde{V}_{N-1}$ are real, then $U_2, \ldots, \tilde{U}_N$ are also real for the same reason. Analogously, if x_1 or $\tilde{x}_1$ is different from zero, we can use the same idea to make X_1 real and conclude that also $X_1, \ldots, \tilde{X}_{N-1}$ and $Y_2, \ldots, \tilde{Y}_N$ will have to be real. If x_1 and $\tilde{x}_1$ are zero, we look at x_2 and $\tilde{x}_2$ and so on.

So we have shown that the vectors $\mathbf{U}_2, \ldots, \mathbf{U}_N$, $\mathbf{V}_1, \ldots, \mathbf{V}_{N-1}$, $\mathbf{X}_1, \ldots, \mathbf{X}_{N-1}$, and $\mathbf{Y}_2, \ldots, \mathbf{Y}_N$

can be made real. We now prove that $\mathbf{U}_1$ will also be real. Suppose the first element U_1 is complex. According to the Remark 2.2 after Theorem 2.1, we have the freedom of choosing a real orthogonal transformation to select the basis functions of $\boldsymbol{\psi}_1$, and since such a transformation on $\boldsymbol{\psi}_1$ implies an orthogonal transformation on $\mathbf{U}_1$, we can choose it such that U_1 is real. Since the imaginary part of the diagonal elements of the blocks from $\hat{\mathbf{D}}$ have opposite signs and the same holds for the diagonal elements of the blocks from $\hat{\mathbf{A}}$ (see (5.1)), also the second element $\tilde{U}_1$ of $\mathbf{U}_1$ will be real. Since this real $\mathbf{U}_1$ was obtained after a real orthogonal transformation, this means that the original $\mathbf{U}_1$ was real from the beginning. The same idea can be used to prove that also $\mathbf{V}_N$ will be real. This means that the matrix $\hat{\mathbf{S}}$ of (4.3) is real.

We define $\hat{\mathbf{B}} = \hat{\mathbf{D}} + \hat{\mathbf{A}}$, so that equation (3.6) can be written as

$$\hat{\mathbf{Q}}^* \hat{\mathbf{Z}} \hat{\mathbf{Q}} = \hat{\mathbf{S}} + \hat{\mathbf{B}}. \tag{4.5}$$

In this equation the left hand side and the matrix $\hat{\mathbf{S}}$ are real. Therefore also $\hat{\mathbf{B}}$ is real.

5 An explicit expression for the matrix $\hat{\mathbf{B}}$

In this section we derive an explicit form for the block diagonal matrix $\hat{\mathbf{B}}$. This explicit form will lead to conditions defining the transformations $\mathbf{p}_i$ which are needed to design a practical algorithm to be described in the next section. First we consider $\hat{\mathbf{A}}$. For notational simplicity, we shall drop the index i on all quantities referring to the ith block in the rest of this section. A general diagonal block has the form (recall that $\gamma = \nu - \mu$ and $t, \mu, \nu, \gamma \in \mathbb{R}$)

$$\begin{aligned}
\hat{\mathbf{a}} = \mathbf{p}^* \boldsymbol{\alpha} \mathbf{p} &= \begin{bmatrix} \cos(t) & \sin(t)e^{-\mathbf{i}\mu} \\ -\sin(t)e^{-\mathbf{i}\gamma} & \cos(t)e^{-\mathbf{i}\nu} \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & \alpha^* \end{bmatrix} \begin{bmatrix} \cos(t) & -\sin(t)e^{\mathbf{i}\gamma} \\ \sin(t)e^{\mathbf{i}\mu} & \cos(t)e^{\mathbf{i}\nu} \end{bmatrix} \\
&= \begin{bmatrix} \cos^2(t)\alpha + \sin^2(t)\alpha^* & -\sin(t)\cos(t)(\alpha - \alpha^*)e^{\mathbf{i}\gamma} \\ -\cos(t)\sin(t)(\alpha - \alpha^*)e^{-\mathbf{i}\gamma} & \sin^2(t)\alpha + \cos^2(t)\alpha^* \end{bmatrix} = \begin{bmatrix} e & f \\ -f^* & e^* \end{bmatrix},
\end{aligned} \tag{5.1}$$

where

$$e = \Re(\alpha) + \mathbf{i}\Im(\alpha)\cos(2t) \tag{5.2}$$
$$f = \sin(2t)\Im(\alpha)(\sin(\gamma) - \mathbf{i}\cos(\gamma)). \tag{5.3}$$

Since we noted at the end of the previous section that $\hat{\mathbf{B}} = \hat{\mathbf{D}} + \hat{\mathbf{A}}$ is real, we can combine (4.4), (5.2) and (5.3) to see that the blocks of $\Im(\hat{\mathbf{B}})$ equal

$$\begin{bmatrix} \frac{d}{2}\sin(2t)\sin(\rho) + \Im(\alpha)\cos(2t) & d\cos^2(t)\sin(\rho+\gamma) - \Im(\alpha)\sin(2t)\cos(\gamma) \\ -d\sin^2(t)\sin(\rho-\gamma) - \Im(\alpha)\sin(2t)\cos(\gamma) & -\frac{d}{2}\sin(2t)\sin(\rho) - \Im(\alpha)\cos(2t) \end{bmatrix}$$

where besides $\gamma = \nu - \mu$ we have used $\rho = \delta + \mu$. This matrix needs to be zero, thus

$$\begin{cases} \frac{d}{2}\sin(2t)\sin(\rho) = -\Im(\alpha)\cos(2t) \\ d\cos^2(t)\sin(\rho+\gamma) = \Im(\alpha)\sin(2t)\cos(\gamma) \\ -d\sin^2(t)\sin(\rho-\gamma) = \Im(\alpha)\sin(2t)\cos(\gamma). \end{cases} \tag{5.4}$$

We will show that the third relation follows from the first and the second. In fact we show in Lemma B.1 of the appendix that if $\Im(\alpha) \neq 0$, then the previous system is equivalent with

$$\begin{cases} d\sin(\rho) = -2\Im(\alpha)\cot(2t) \\ d\sin(\rho)\cos(\gamma) = -\sin(\gamma)\cos(\rho)\cos(2t). \end{cases}$$

Note that we may assume here that $\Im(\alpha) \neq 0$, otherwise the poles of the ORFs are real and then we can take $\mathbf{p} = \mathbf{I}_2$, the identity matrix of order 2, because then the vector $\boldsymbol{\phi} = [\phi,\ \tilde{\phi}]^T$ is real already. This also implies that it is possible to add two distinct real poles instead of a pair of complex conjugate poles.

When we combine (4.1) and Theorem 2.1, we can write

$$\begin{aligned} \begin{bmatrix} \psi(z) \\ \tilde{\psi}(z) \end{bmatrix} &= \begin{bmatrix} e^{\mathbf{i}\theta}(\cos(t)\phi(z) + \sin(t)e^{\mathbf{i}\mu}\tilde{\phi}(z)) \\ e^{\mathbf{i}(\theta+\gamma)}(-\sin(t)\phi(z) + \cos(t)e^{\mathbf{i}\mu}\tilde{\phi}(z)) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & e^{\mathbf{i}\gamma} \end{bmatrix} \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix} \begin{bmatrix} e^{\mathbf{i}\theta}\phi(z) \\ e^{\mathbf{i}(\theta+\mu)}\tilde{\phi}(z) \end{bmatrix}. \end{aligned}$$

It is obvious that the 2 rotations θ and $\theta + \mu$ can be incorporated in ϕ and $\tilde{\phi}$ since these are only defined up to a unimodular constant factor. They could be used for example to fix the

leading coefficient in the numerator of ϕ and $\tilde{\phi}$ to be positive or whatever normalization that is considered to be suitable. Thus only t and γ need to be determined. The two relations we derived in Lemma B.1 completely define the matrix $\mathbf{p}$ except for the two rotations θ and $\theta + \mu$. But the latter two rotations are not necessary for deriving the algorithm.

It now follows from (4.4), (5.2) and (5.3) that the matrix $\hat{\mathbf{B}}$ has blocks of the form

$$\begin{bmatrix} \Re(\alpha) + \frac{d}{2}\sin(2t)\cos(\rho) & d\cos^2(t)\cos(\rho+\gamma) + \Im(\alpha)\sin(2t)\sin(\gamma) \\ -(d\sin^2(t)\cos(\rho-\gamma) + \Im(\alpha)\sin(2t)\sin(\gamma)) & \Re(\alpha) - \frac{d}{2}\sin(2t)\cos(\rho) \end{bmatrix}.$$

By Lemma B.2 (see appendix), we find that

$$\begin{aligned} \frac{d}{2}\sin(2t)\cos(\rho) &= \Im(\alpha)\cot(\gamma) \\ d\cos^2(t)\cos(\rho+\gamma) + \Im(\alpha)\sin(2t)\sin(\gamma) &= \Im(\alpha)\frac{\cot(t)}{\sin(\gamma)} \\ d\sin^2(t)\cos(\rho-\gamma) + \Im(\alpha)\sin(2t)\sin(\gamma) &= \Im(\alpha)\frac{\tan(t)}{\sin(\gamma)}. \end{aligned}$$

Note that these relations imply that the blocks of $\hat{\mathbf{B}}$ equal

$$\hat{\mathbf{b}} = \Re(\alpha)\mathbf{I}_2 + \Im(\alpha)\begin{bmatrix} \cot(\gamma) & \frac{\cot(t)}{\sin(\gamma)} \\ -\frac{\tan(t)}{\sin(\gamma)} & -\cot(\gamma) \end{bmatrix} = \Re(\alpha)\mathbf{I}_2 + \Im(\alpha)\begin{bmatrix} \cot(\gamma) & \frac{\lambda}{\sin(\gamma)} \\ -\frac{1}{\lambda\sin(\gamma)} & -\cot(\gamma) \end{bmatrix} \tag{5.5}$$

where $\lambda = \cot(t)$.

Remark 5.1 It is clear that by construction the eigenvalues of $\hat{\mathbf{a}}$ are α and α^*. Adding the matrix $\hat{\mathbf{D}}$ does not change the situation. It is indeed a simple exercise to verify that the eigenvalues of $\hat{\mathbf{b}}$ are still α and α^*. See lemma B.3 below.

6 The algorithm

Because of space limitations it is impossible to describe here the details of the algorithm. This will be done in a separate paper. It is different from, but similar to, the more general block

algorithm for vector orthogonal rational functions considered in [4]. We only give a brief outline of a general step.

Suppose we have already solved the problem for $i-1$ pairs of datapoints and that we want to add the ith pair. The first step is then a normalization.

Lemma 6.1 *Denote by* $\eta_1(z) = 1/\tilde{\pi}_1(z)$ *and* $\tilde{\eta}_1(z) = z/\tilde{\pi}_1(z)$, *with* $\tilde{\pi}_1(z) = z^2 - 2\Re(\alpha_1)z + |\alpha_1|^2$ *as in Section 2, the functions spanning* $\tilde{\mathcal{L}}_1$, *and define the block column matrix* $\mathbf{N}$ *whose* i*th block is*

$$\mathbf{n}_i = w_i \begin{bmatrix} \Re\left(\eta_1(z_i)\right) & \Re\left(\tilde{\eta}_1(z_i)\right) \\ \Im\left(\eta_1(z_i)\right) & \Im\left(\tilde{\eta}_1(z_i)\right) \end{bmatrix}, \quad i = 1, \ldots, N, \tag{6.1}$$

and with $\hat{\mathbf{Q}}$ *as above, we have that* $\hat{\mathbf{Q}}^T\mathbf{N} = \mathbf{\Omega}$ *where* $\mathbf{\Omega}$ *has all its blocks, except for the first one, equal to zero, which is a nonsingular upper triangular* 2×2 *matrix.*

PROOF. This follows from the fact that the matrix $\hat{\mathbf{Q}}$ contains the orthogonal rational functions ψ_i and $\tilde{\psi}_i$ and the way these functions are built (see the proof of Theorem 2.1).

Indeed, the elements of $\hat{\mathbf{Q}}^T\mathbf{N}$ correspond to inner products so that the ith block in the right hand side is

$$\begin{bmatrix} \langle \psi_i, \eta_1 \rangle & \langle \psi_i, \tilde{\eta}_1 \rangle \\ \left\langle \tilde{\psi}_i, \eta_1 \right\rangle & \left\langle \tilde{\psi}_i, \tilde{\eta}_1 \right\rangle \end{bmatrix}.$$

Because ψ_i and $\tilde{\psi}_i$, being in the space $\tilde{\mathcal{L}}_i \setminus \tilde{\mathcal{L}}_{i-1}$, are orthogonal to $\tilde{\mathcal{L}}_1$ spanned by η_1 and $\tilde{\eta}_1$, if $i > 1$, all blocks are zero except the first one. The (2,1) element of this first block is zero because it equals $\left\langle \tilde{\psi}_1, \eta_1 \right\rangle$ which is zero because $\tilde{\psi}_1$ is orthogonal to ψ_1 which is a normalized version of η_1.

The algorithm is recursive, thus the block column $\mathbf{N}$ has been reduced to its first block, making all the other blocks vanish. So when in the ith step of the algorithm new data are added (the pair of points (z_i, z_i^*) on $\mathbb{T}$ and the pair of poles (α_i, α_i^*) in $\mathbb{D}$) and if these are appended at the bottom of the matrices obtained already, then the normalisation will only require an orthogonal transformation on the first and the last block row of the matrices involved.

However, since we need *similarity* transformations, these transformations are not only applied to the first and last block of the matrices $\hat{\mathbf{S}}_{i-1}+\hat{\mathbf{B}}_{i-1}$ to which a new diagonal block $\hat{\mathbf{z}}_i$ is appended but also to their first and last block column. This destroys their semiseparable structure. Thus successive orthogonal (block Givens) similarity transformations are needed on 4×4 blocks to restore this structure. These are applied to the last block row and column and subsequently the first, second, third, etc. Using the explicit form of the matrices obtained in the previous section, it is possible to find these transformation matrices and the resulting parameterisation of the restored structured matrices by solving small 4×4 eigenvalue problems. All these operations should be applied to the generators, used in the previous sections, representing the semiseparable matrix. However, for numerical stability, it is much better to update a Givens-vector representation instead [7].

7 Conclusion

It is known that the construction of rational functions with prescribed poles that are orthogonal with respect to a discrete inner product on the unit circle can be constructed by solving an inverse eigenvalue problem. One has to find a semiseparable matrix of rank one plus a diagonal matrix (containing the poles) that is unitarily similar to a prescribed diagonal matrix (containing the data points).

On condition that the poles are introduced in complex conjugate pairs and if also the data points on the unit circle are prescribed in complex conjugate pairs, we can reformulate this problem in a real block form. The existence of such a formulation is proved and explicit expressions are given for the blocks on the diagonal (each block has a complex conjugate pair of poles). These explicit expressions are essential for the design of a practical algorithm. An outline of a recursive algorithm is given, but details will be described in another publication.

Appendix

A The most general unitary 2×2 matrix

Here we prove that a unitary 2×2 matrix has the form we used in Section 3.

Lemma A.1 *Any unitary matrix* $\mathbf{p} \in \mathbb{C}^{2\times 2}$ *can be written as*

$$\mathbf{p} = e^{\mathbf{i}\theta} \begin{bmatrix} \cos(t) & -\sin(t)e^{\mathbf{i}(\nu-\mu)} \\ \sin(t)e^{\mathbf{i}\mu} & \cos(t)e^{\mathbf{i}\nu} \end{bmatrix} = e^{\mathbf{i}\theta} \begin{bmatrix} 1 & 0 \\ 0 & e^{\mathbf{i}\mu} \end{bmatrix} \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{\mathbf{i}\gamma} \end{bmatrix}$$

where μ, ν *and* $\theta \in [0, 2\pi)$ *and* $t \in [0, \frac{\pi}{2}]$, $\gamma = \nu - \mu$.

PROOF. A matrix

$$\mathbf{p} = \begin{bmatrix} ae^{\mathbf{i}\alpha} & be^{\mathbf{i}\beta} \\ ce^{\mathbf{i}\gamma} & de^{\mathbf{i}\delta} \end{bmatrix},$$

where $a,\ b,\ c,\ d \in \mathbb{R}^+$ and $\alpha,\ \beta,\ \gamma,\ \delta \in [0, 2\pi)$ is unitary iff $\mathbf{p}\mathbf{p}^* = \mathbf{I} = \mathbf{p}^*\mathbf{p}$. This equality yields the following 4 equations:

$$\begin{cases} a = d \\ b = c \\ b = \sqrt{1-a^2} \\ \cos(\alpha - \beta - \gamma + \delta) = -1 \end{cases}.$$

The last equation can be solved for β:

$$\beta = \alpha - \gamma + \delta - (2k+1)\pi.$$

If we put

$$\begin{cases} a = \cos(t),\ t \in [0, \frac{\pi}{2}] \\ \alpha = \theta \\ \gamma - \alpha = \mu \\ \delta - \alpha = \nu \end{cases},$$

we find the form of the unitary matrix $\mathbf{p}$ that we proposed.

B Proof of auxiliary results

Lemma B.1 *If $\Im(\alpha) \neq 0$, then in the system (5.4), i.e.,*

$$\begin{cases} \frac{d}{2}\sin(2t)\sin(\rho) = -\Im(\alpha)\cos(2t) \\ d\cos^2(t)\sin(\rho+\gamma) = \Im(\alpha)\sin(2t)\cos(\gamma) \\ -d\sin^2(t)\sin(\rho-\gamma) = \Im(\alpha)\sin(2t)\cos(\gamma). \end{cases} \tag{B.1}$$

the third equation is implied by the first two and (B.1) is equivalent with the system

$$\begin{cases} d\sin(\rho) = -2\Im(\alpha)\cot(2t) \\ \sin(\rho)\cos(\gamma) = -\sin(\gamma)\cos(\rho)\cos(2t). \end{cases} \tag{B.2}$$

PROOF. The case $d = 0$ is trivial, so we assume $d \neq 0$ and set $\beta = \Im(\alpha)/d$.

Then, substituting the second, starting from the second and third equation of (B.1) we get

$$\begin{aligned} \sin^2(t)\sin(\rho-\gamma) + \beta\sin(2t)\cos(\gamma) &= \sin^2(t)\sin(\rho-\gamma) + \cos^2(t)\sin(\rho+\gamma) \\ &= \sin(\rho)\cos(\gamma) + \cos(2t)\sin(\gamma)\cos(\rho). \end{aligned} \tag{B.3}$$

The first part of the lemma is proved when this expression equals zero. Dividing the first two relations gives

$$\frac{\sin(2t)\sin(\rho)}{2\cos^2(t)\sin(\rho+\gamma)} = \frac{-\cos(2t)}{\sin(2t)\cos(\gamma)}.$$

Thus

$$\begin{aligned} 0 &= \sin^2(2t)\sin(\rho)\cos(\gamma) + 2\cos(2t)\cos^2(t)\sin(\rho+\gamma) \\ &= 2\cos^2(t)\left(2\sin^2(t)\sin(\rho)\cos(\gamma) + \cos(2t)\sin(\rho+\gamma)\right). \end{aligned}$$

Since we supposed that $\Im(\alpha) \neq 0$, this implies that $t \neq k\pi/2$ for integer k, from which $\cos(t) \neq 0$, so that the last relation yields

$$\begin{aligned} 0 &= 2\sin^2(t)\sin(\rho)\cos(\gamma) + \cos(2t)\sin(\rho+\gamma) \\ &= (2\sin^2(t) + \cos(2t))\sin(\rho)\cos(\gamma) + \cos(2t)\sin(\gamma)\cos(\rho) \\ &= \sin(\rho)\cos(\gamma) + \cos(2t)\sin(\gamma)\cos(\rho). \end{aligned}$$

And this is nothing but the right hand side of (B.3), which proves the lemma.

For the equivalence of the systems, note that the first equations in both systems are the same.

Take the two remaining equations in (B.1) and note that the left hand sides are equal because the right hand sides are. Writing out $\sin(\rho+\gamma)$ and $\sin(\rho-\gamma)$ leads after simplification to the second equation of (B.2).

In the other direction, we may rewrite the second equation of (B.2) as

$$\sin(\rho)\cos(\gamma) = \sin(\gamma)\cos(\rho)[2\sin^2(t) - 1]$$

or

$$\sin(\rho+\gamma) = 2\sin(\gamma)\cos(\rho)\sin^2(t). \tag{B.4}$$

Also by taking the ratio of the first and the second equation of (B.2) we get

$$\frac{d}{2}\sin(\gamma)\cos(\rho)\sin(2t) = \Im(\alpha)\cos(\gamma),$$

and after multiplying both sides with $\sin(2t)$

$$\begin{aligned}\Im(\alpha)\cos(\gamma)\sin(2t) &= \frac{d}{2}\sin(\gamma)\cos(\rho)\sin^2(2t)\\ &= d\cos^2(t)[2\sin(\gamma)\cos(\rho)\sin^2(t)]\\ &= d\cos^2(t)\sin(\rho+\gamma)\end{aligned}$$

where the last line is by (B.4). This proves the equivalence of the two systems.

Here we prove the relations that we need to find the simple form of the blocks of $\hat{\mathbf{B}}$ in (5.5).

Lemma B.2 *If $\Im(\alpha) \neq 0$ and system (B.2) holds, i.e.,*

$$\begin{cases} d\sin(\rho) &= -2\Im(\alpha)\cot(2t)\\ \sin(\rho)\cos(\gamma) &= -\sin(\gamma)\cos(\rho)\cos(2t)\end{cases}$$

then

$$\begin{aligned}\frac{d}{2}\sin(2t)\cos(\rho) &= \Im(\alpha)\cot(\gamma)\\ d\cos^2(t)\cos(\rho+\gamma)+\Im(\alpha)\sin(2t)\sin(\gamma) &= \Im(\alpha)\frac{\cot(t)}{\sin(\gamma)}\\ d\sin^2(t)\cos(\rho-\gamma)+\Im(\alpha)\sin(2t)\sin(\gamma) &= \Im(\alpha)\frac{\tan(t)}{\sin(\gamma)}.\end{aligned}$$

PROOF. By dividing the second relation by the first we find

$$\cos(\gamma) = \frac{d}{2\Im(\alpha)}\sin(\gamma)\cos(\rho)\sin(2t), \tag{B.5}$$

which can easily be rewritten as the first equation we need to prove.

Now we prove the second relation

$$
\begin{aligned}
d\cos^2(t)\cos(\rho+\gamma) &= d\cos^2(t)\cos(\rho)\cos(\gamma) - d\cos^2(t)\sin(\rho)\sin(\gamma) \\
&= d\cos^2(t)\cos(\rho)\cos(\gamma) + 2\Im(\alpha)\cos^2(t)\sin(\gamma)\cot(2t) \\
&= d\cos^2(t)\cos(\rho)\cos(\gamma) + \Im(\alpha)\cot(t)\sin(\gamma) - \Im(\alpha)\sin(2t)\sin(\gamma)
\end{aligned}
$$

where in the last step we used $\cot(2t) = \frac{1-2\sin^2(t)}{2\sin(t)\cos(t)}$. To prove the second equation, it remains to show that

$$
d\cos^2(t)\cos(\rho)\cos(\gamma) + \Im(\alpha)\cot(t)\sin(\gamma) = \Im(\alpha)\frac{\cot(t)}{\sin(\gamma)}.
$$

This is easy by using (B.5). The left hand side of the equation above equals

$$
\begin{aligned}
\Im(\alpha)\frac{\cot(t)}{\sin(\gamma)}\left(\frac{d\cos(t)\sin(t)\cos(\rho)\sin(\gamma)}{\Im(\alpha)}\cos(\gamma) + \sin^2(\gamma)\right) &= \Im(\alpha)\frac{\cot(t)}{\sin(\gamma)}(\cos^2(\gamma)+\sin^2(\gamma)) \\
&= \Im(\alpha)\frac{\cot(t)}{\sin(\gamma)}.
\end{aligned}
$$

The last equation can be proved analogously. This proves the lemma.

Lemma B.3 *The matrix*

$$
\hat{\mathbf{b}} = \begin{bmatrix} \Re(\alpha) + \Im(\alpha)\cot(\gamma) & \Im(\alpha)\frac{\lambda}{\sin(\gamma)} \\ -\Im(\alpha)\frac{1}{\lambda\sin(\gamma)} & \Re(\alpha) - \Im(\alpha)\cot(\gamma) \end{bmatrix}
$$

with $\Im(\alpha) \neq 0$ *has the following eigenvalue decomposition*

$$
\hat{\mathbf{b}}\begin{bmatrix} a\lambda & b\lambda \\ -ae^{-\mathbf{i}\gamma} & -be^{\mathbf{i}\gamma} \end{bmatrix} = \begin{bmatrix} a\lambda & b\lambda \\ -ae^{-\mathbf{i}\gamma} & -be^{\mathbf{i}\gamma} \end{bmatrix}\begin{bmatrix} \alpha & 0 \\ 0 & \alpha^* \end{bmatrix},
$$

with $a,\, b \in \mathbb{C} \setminus \{0\}$.

PROOF. This is just a simple calculation.

Note that α defines the eigenvalues and λ and γ define the eigenvectors up to a constant multiple (as long as $\Im(\alpha) \neq 0$).

References

[1] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. *Orthogonal rational functions*, volume 5 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 1999.

[2] A. Bultheel, M. Van Barel, and P. Van gucht. Orthogonal bases in discrete least squares rational approximation. *J. Comput. Appl. Math.*, 164-165:175–194, 2004.

[3] D.K. de Vries and P. van den Hof. Frequency domain identification with generalized orthonormal basis functions. *IEEE Trans. Automat. Control*, AC-43(5):656–669, 1998.

[4] S. Delvaux and M. Van Barel. Orthonormal rational function vectors. *Numer. Math.*, 100(3):409 – 440, 2005.

[5] M. Van Barel, D. Fasino, L. Gemignani, and N. Mastronardi. Orthogonal rational functions and structured matrices. *SIAM J. Matrix Anal. Appl.*, 26(3):810–829, 2005.

[6] P. Van gucht and A. Bultheel. Using orthogonal rational functions for system identification: numerical aspects. *IEEE Trans. Automat. Control*, 48(4):705–709, 2003.

[7] R. Vandebril, M. Van Barel, and N. Mastronardi. A note on the representation and definition of semiseparable matrices. *Numer. Linear Algebra Appl.*, 12:839–858, 2005.

[8] B. Wahlberg. System identification using Laguerre models. *IEEE Trans. Automat. Control*, AC-36(5):551–562, 1991.

[9] B. Wahlberg. System identification using Kautz models. *IEEE Trans. Automat. Control*, AC-39(6):1276–1282, 1994.