

Darboux transformations for CMV matrices

M. J. Cantero¹, F. Marcellán², L. Moral¹, L. Velázquez¹ *

¹ Departamento de Matemática Aplicada and Instituto Universitario de Matemáticas y Aplicaciones (IUMA), Universidad de Zaragoza, Spain

² Instituto de Ciencias Matemáticas (ICMAT) and Departamento de Matemáticas, Universidad Carlos III de Madrid, Spain

Abstract

We develop a theory of Darboux transformations for CMV matrices, canonical representations of the unitary operators. In perfect analogy with their self-adjoint version – the Darboux transformations of Jacobi matrices – they are equivalent to Laurent polynomial modifications of the underlying measures. We address other questions which emphasize the similarities between Darboux transformations for Jacobi and CMV matrices, like their (almost) isospectrality or the relation that they establish between the corresponding orthogonal polynomials, showing also that both transformations are connected by the Szegő mapping.

Nevertheless, we uncover some features of the Darboux transformations for CMV matrices which are in striking contrast with those of the Jacobi case. In particular, when applied to CMV matrices, the matrix realization of the inverse Darboux transformations – what we call ‘Darboux transformations with parameters’ – leads to spurious solutions whose interpretation deserves future research. Such spurious solutions are neither unitary nor band matrices, so Darboux transformations for CMV matrices are much more subject to the subtleties of the algebra of infinite matrices than their Jacobi counterparts.

A key role in our theory is played by the Cholesky factorizations of infinite matrices. Actually, the Darboux transformations introduced in this paper are based on the Cholesky factorizations of degree one Hermitian Laurent polynomials evaluated on CMV matrices. We also show how these transformations can be generalized to higher degree Laurent polynomials, as well as to the extension of CMV matrices to quasi-definite functionals – what we call ‘quasi-CMV’ matrices.

Keywords and phrases: Darboux transformations, CMV matrices, Cholesky factorizations, almost isospectrality, Schur parameters, orthogonal Laurent polynomials, measures on the unit circle, Hermitian linear functionals, linear spectral transformations, Christoffel transformations, Geronimus transformations.

*The work of the second author has been partially supported by the research project MTM2012-36732-C03-01 from Dirección General de Investigación Científica y Técnica, Ministerio de Economía y Competitividad of Spain. The work of the rest of the authors has been supported in part by the research project MTM2011-28952-C02-01 from the Ministry of Science and Innovation of Spain and the European Regional Development Fund (ERDF), and by Project E-64 of Diputación General de Aragón (Spain).

Contents

1	Introduction	2
2	Hessenberg type matrices, zig-zag bases and Darboux factorizations	8
2.1	Matrix representations of the multiplication operator	13
2.2	Change of basis and Darboux factorization	16
3	Darboux for CMV and Christoffel transformations	20
4	Inverse Darboux for CMV and Geronimus transformations	29
5	Jacobi versus CMV	40
5.1	The Szegő connection	44
5.2	The DVZ connection	49
6	Darboux versus QR	51
7	Further extensions	55
7.1	Higher degree transformations	55
7.2	Quasi-CMV matrices	58
8	Conclusions and Outlook	67
9	Appendix: Cholesky factorizations	69

1 Introduction

Darboux transformations go back to the end of the 19th century, with a result of Moutard [60] which influenced subsequent works of Darboux on transformations of second order linear differential operators [15, 16]. The result was largely ignored till the first half of the 20th century, when the needs of the recently born quantum mechanics rescued it, although with a different look. A factorization trick introduced by Dirac to solve the Schrödinger equation for the harmonic oscillator [23] was used later by Schrödinger himself [64, 65] and others [44, 45, 46] to deal with less trivial instances of such an equation. This ended up with a factorization machinery to generate hierarchies of solvable systems, known nowadays as Darboux transformations because they constitute a reformulation of the results from Darboux.

The quantum renaissance of Darboux ideas apparently occurred with no knowledge of Darboux antecedents. It was only later recognized that the factorization trick was closely related to Darboux work, who on the other hand never mentioned such a factorization interpretation of his result. Actually, the pressure from the new physics

was not the only reason for the revival of Darboux work, but its new factorization look played a central role in this respect. Darboux factorizations turned out to be an efficient tool in areas far beyond their initial scope, like soliton theory, integrable systems, special functions, bispectral problems, signal theory, general relativity or cosmology (see [38, 42, 57, 59, 71] and references therein).

Darboux factorizations, initially emerged as a tool to find solvable systems, have acquired a conceptual meaning which has placed them at the very heart of the foundations of modern physics. The ladder operators appearing in the factorization of the harmonic oscillator were the origin of the creation-annihilation operator formalism, cornerstone of the quantum theory of radiation and the treatment of quantum many body systems, which led to the current formulation of the fundamental laws in quantum physics, i.e. the quantum field theory. The factorization and intertwining of operators behind Darboux transformations are also the core of supersymmetry, a promising alternative to overcome the drawbacks in unifying quantum theory and gravitation. A modern and friendly review of these ideas can be found in [59].

In its simplest setting, a Darboux transformation starts with the factorization of a non-negative self-adjoint operator into a product of another operator and its adjoint. This is followed by a commutation of both factors to give a new non-negative self-adjoint operator. If the initial self-adjoint operator is not non-negative but has a spectrum which is bounded from below (as it is the case of a typical quantum Hamiltonian), we start instead with a non-negative operator obtained by shifting the spectrum, adding a suitable multiple of the identity which is subtracted after the intertwining. This procedure which connects pairs of self-adjoint operators is called a Darboux transformation. Of course, the referred factorization of a non-negative operator is not unique, so the interest of a Darboux transformation highly depends on the features of the corresponding Darboux factors. In particular, a useful Darboux transformation should preserve the class of operators under study (e.g., Schrödinger operators in quantum mechanics), permitting also its iterative implementation.

Concerning spectral issues, when the Darboux factors are bounded with a trivial kernel, the Darboux transformation is isospectral and the Darboux factors connect eigenvectors corresponding to the same eigenvalues. This was the original reason for the use of Darboux transformations in quantum mechanics, since they allowed to reduce the spectral problem of a Schrödinger operator to a simpler one. However, even if a self-adjoint operator is bounded (so that every Darboux factor is also bounded), a Darboux factorization may involve factors with a non-trivial kernel, a fact that could break isospectrality. Nevertheless, in that case, the Darboux transformation remains almost isospectral since it preserves the spectrum with the exception of, at most, a single point (the opposite of the shift used in the transformation). This situation remains useful for the reduction of spectral problems, but can be also used to engineer operators by deleting or inserting eigenvalues on demand. For instance, one can insert into a gap of a lattice Hamiltonian discrete energy levels representing bound states created by an impurity [26].

Since their origin, tied to Schrödinger equations, Darboux transformations have

been extended to many other contexts because self-adjoint operators are ubiquitous in mathematics and physics. For instance, the fact that the 1D Schrödinger operators generate the Lax pair of the KdV equation made possible to export Darboux transformations to KdV and its generalizations [57]. The discrete version of these connections established a similar link between Darboux transformations for Jacobi matrices (discrete 1D Schrödinger) and the Toda lattice (discrete KdV) [71].

Jacobi matrices also connect with the theory of orthogonal polynomials on the real line, where the use of Darboux techniques has led to the discovery of new bispectral instances [39, 40, 41] (see also [38, 42] and references therein). From the orthogonal polynomial point of view, it is known that Darboux transformations of Jacobi matrices are equivalent to degree one polynomial modifications of the underlying orthogonality measure, a.k.a. Christoffel transformations. Therefore, inverse Darboux corresponds to the so called Geronimus transformations which, apart from dividing the measure by a polynomial of degree one, may insert a mass point at the zero of the polynomial (in agreement with the fact that Darboux transformations may change, at most, a single eigenvalue). This links Darboux transformations of Jacobi matrices to the so called linear spectral transformations on the real line [78].

The connection with orthogonal polynomials on the real line is also the root of the understanding of Jacobi matrices as the canonical representations of self-adjoint operators in Hilbert spaces, a result which goes back to the classical work of Stone [70]. This result is the deep reason for the ubiquity of Jacobi matrices in applications, since it means that tridiagonal matrices give the simplest non-trivial discrete models described by self-adjoint operators. This explains the success of the particularization of Darboux transformations to Jacobi matrices.

Darboux transformations of Jacobi matrices also had, and continue to have, a remarkable impact in numerical analysis, constituting the center of a rich interplay between the theory of orthogonal polynomials on the real line and numerical linear algebra. This symbiosis is evidenced by a large list of references combining these subjects, among them those which first established the link of Darboux and similar factorization methods, like the QR algorithm, with the Christoffel and Geronimus transformations [5, 6, 7, 24, 27, 31, 32, 35, 49, 74, 77].

The works in numerical linear algebra also promoted the extension of this type of intertwining methods to the unitary case, in particular to unitary (more generally, isometric) Hessenberg matrices, in close connection with the theory of orthogonal polynomials on the unit circle [17, 28, 30, 36, 43, 73]. However, the study of linear spectral transformations on the unit circle via Darboux transformations of Hessenberg matrices finds two important drawbacks which reflect that these matrices are not exactly the unitary counterpart of Jacobi: The lack of a band structure and the eventual lack of unitarity since they are in general only isometric. This is related to the fact that the related orthogonal polynomials on the unit circle do not always provide a complete orthogonal system of the corresponding Hilbert space of square-integrable functions.

A basis of such a Hilbert space is always given by the orthogonal Laurent poly-

nomials on the unit circle, which allow to overcome the above difficulties. They lead to the true unitary analogue of Jacobi matrices, a kind of five-diagonal unitary matrices nowadays known as CMV matrices. CMV matrices date back to works on the unitary eigenproblem [4, 8, 74]. Indeed, the interpretation of the QR algorithm for CMV matrices in terms of transformations of measures already appeared in the same publication [74] which first introduced CMV matrices systematically. Yet, Darboux methods have not been applied to CMV matrices so far, despite their rediscovery a decade after [74] in the context of orthogonal polynomials on the unit circle [11] and the numerous applications in which CMV matrices have been involved since then (the monograph [67] serves as an illustration of some of these applications). A quick introduction to CMV matrices can be found in [68].

The unitary version of Darboux (and in particular its extension to CMV matrices) is of great interest because, together with self-adjointness, unitarity is the most common operator symmetry in pure and applied sciences. For instance, while the energy levels of an electron in a quantum system are given by the spectrum of a self-adjoint stationary Schrödinger operator, its evolution is governed by a unitary operator determined by the time-dependent Schrödinger equation. The relativistic version of this unitary evolution is given by the Dirac equation, whose discretization in 1D leads to the so called quantum walks [1, 58] (the quantum version of random walks), characterized by an evolution given by a CMV matrix [9].

CMV matrices play the role of Lax operators for a classical integrable system equivalent to relativistic Toda, known under the name of Ablowitz-Ladik [51, 53, 62, 63]. They provide also the Lax pair of the so called Schur flows (discrete mKdV) [34], an integrable system which was known to be intimately connected with the orthogonal polynomials on the unit circle [3, 25, 61] (see also [2] for a CMV approach to this kind of integrable systems via LU factorizations).

The ubiquity of CMV matrices in applications involving unitaries is due to similar reasons as those given for Jacobi matrices: CMV matrices provide the canonical representations of the unitary operators in Hilbert spaces [12]. Analogously to the Jacobi case, the proof of this relies on the relation between CMV matrices and orthogonal polynomials on the unit circle, to which they were linked from the very beginning [11, 74]. In this context, CMV matrices provide the five term recurrence relation of the orthogonal Laurent polynomials on the unit circle, closely related to the orthogonal polynomials themselves. Curiously enough, these orthogonal polynomials can be considered as the discrete version of the so called Krein systems, which are known to be equivalent to the 1D Dirac equation [19].

This closes a circle of relations which reveal the strong parallelism between Jacobi and CMV matrices, showing that Darboux transformations for CMV matrices should be as interesting as Darboux transformations for Jacobi matrices. Apart from the possible applications to quantum physics, integrable systems or numerical analysis, Darboux transformations of CMV matrices may open new ways in the search for bispectral situations on the unit circle.

The aim of the present work is to develop a theory of Darboux transformations

for CMV matrices which could be useful for applications which parallel those of the Jacobi case. The strategy is to transform a CMV matrix into a self-adjoint band matrix to which we apply standard Darboux transformations based on triangular band factorizations. Darboux for Jacobi matrices is usually implemented via LU factorizations of non-symmetric tridiagonal matrices obtained by a diagonal conjugation of the Jacobi one. However, it is possible to rewrite these transformations by performing Cholesky factorizations directly onto the Jacobi matrix. This avoids losing the hermiticity in the process, which fits with the general setting described previously for Darboux transformations of self-adjoint operators. We will use the Cholesky approach for the CMV version of Darboux since it requires no modification of the initial CMV matrix, permitting a closer control of unitarity, which is a more intricate property than hermiticity.

The control of the unitarity becomes especially important when discussing the ‘Darboux transformations with parameters’, the matrix realization of inverse Darboux. These transformations start with a ‘reversed’ Cholesky factorization depending on a number of parameters which reflect the lack of uniqueness of such a factorization. Then, the preservation of the unitarity is not automatically guaranteed by the transformation due to the abundance of degrees of freedom for the factorization if compared with the Jacobi case. Nevertheless, due to the advantages of the Cholesky approach, the CMV solutions can be characterized by simple constraints on the degrees of freedom of the factorization (see Theorem 4.3). The origin of this difference with respect to the Jacobi case is that transforming a unitary matrix into a Hermitian one requires a Laurent polynomial transformation instead of a polynomial one. The appearance of spurious solutions has to do with the fact that the minimal choice for a non-trivial Laurent polynomial mapping unitary matrices into Hermitian ones involves Laurent polynomials with two zeros. This resembles the situation in the case of higher order Darboux transformations of Jacobi matrices, based on transforming the Jacobi matrix by a polynomial of degree greater than one.

The existence of the above spurious solutions, which are neither unitary nor band matrices, requires a careful manipulation of infinite matrices for the Darboux transformations of CMV matrices. Properties like the associativity of matrix multiplication or the uniqueness of the inverse matrix may fail [14]. These troubles are not so evident for Jacobi matrices due to the absence of spurious solutions. The situation in the CMV case can be successfully handled thanks to the Hessenberg type structure of the spurious solutions. These reasons make it advisable to specify from the very beginning the kind of matrix operations that will be admissible and their properties. This will be necessary to legitimate future matrix manipulations, but will also serve to shed light on some aspects of Darboux transformations which, even in the Jacobi case, are not usually explicitly addressed. An example of this is the typical proof of uniqueness of Cholesky factorizations based on regrouping lower and upper triangular factors, concluding that these regroupings must be diagonal (see Appendix 9). At least formally, this proof also works for reversed Cholesky factorizations, in contradiction with the true result. The clarification of the rules for

the operations with infinite matrices, as requested by Darboux for CMV, will help to unravel the flaws in this kind of reasonings.

Despite the above differences, we will find strong similarities between Darboux for Jacobi and CMV. Both of them are almost isospectral transformations, a result that follows easily from their interpretation in terms of polynomial and Laurent polynomial modifications of measures, which in turn connects with the so called linear spectral transformations on the real line and the unit circle [29], respectively. The only quirk of Darboux for CMV is that these transformations may change two eigenvalues instead of only one, due to the use of Laurent polynomials with two zeros. The Darboux triangular factor is a band matrix for both, Jacobi and CMV. In the (CMV) Jacobi case it is (3-band) 2-band and plays the role of change of basis between the orthogonal (Laurent) polynomials of the matrices connected by the transformation. Actually, this kind of triangular band relation between two sequences of orthogonal (Laurent) polynomials characterizes the Darboux transformations (see Theorem 3.2). The analogy between Darboux for Jacobi and CMV can be tested by using the Szegő connection between orthogonal polynomials on the real line and the unit circle. This connection shows clearly that the Darboux transformations for CMV matrices are the unit circle counterpart of the Darboux transformations for Jacobi matrices. Actually, the corresponding Darboux factors are closely related.

Darboux transformations for CMV matrices can be compared with previous matrix transformations based on factorizations and commutation of the factors. A precedent of this is the QR algorithm for CMV matrices already mentioned, which is equivalent to a special type of Laurent polynomial modifications of the orthogonality measure [74]. Nevertheless, as we will see, the QR algorithm has a more limited applicability than the Darboux transformations, which can be considered as a way to extend the QR algorithm to general Laurent polynomial modifications of measures. Actually, the Darboux factor provides the triangular factor of the QR algorithm, but Darboux makes sense in cases where QR not.

Darboux transformations can be generalized to higher order Laurent polynomial transformations of CMV matrices, as well as to situations related to orthogonal polynomials with respect to quasi-definite functionals. This last item requires the use of generalized Cholesky factorizations of non-positive definite matrices and the extension of CMV matrices to quasi-definite functionals [11].

The above ideas are developed in the paper according to the following schedule: Section 2 introduces the basic setting in which the rest of the paper will be conducted. It covers the explicit description of the algebra of infinite matrices that will be used, the properties of the matrix representations of the multiplication operator with respect to ordered basis of Laurent polynomials ('zig-zag' bases) and general Darboux factorizations for Hermitian Laurent polynomials evaluated on such matrix representations. Section 3 deals with the Darboux transformations for CMV matrices. The main result is Theorem 3.2, which gives several characterizations of such transformations. Inverse Darboux transformations and their matrix realization, the Darboux transformations with parameters, are accounted for in Section 4, including

the discussion about the spurious solutions and the characterization of the CMV ones. Section 5 establishes the link between Darboux for Jacobi and CMV via the Szegő mapping, showing also their relation with a new connection between the real line and the unit circle recently obtained [22]. In Section 6 we find a comparison between Darboux transformations and the QR algorithm for CMV matrices, which also serves to rewrite the former ones in operator language. Higher degree Darboux transformations and the extension to quasi-definite functionals appear in Section 7. Section 8 summarizes the conclusions, pointing out some open problems suggested by the previous results. Finally, Appendix 9 deals with the existence and uniqueness of Cholesky factorizations for infinite matrices, which are crucial issues in the development of the paper. Some illustrative examples of Darboux transformations for CMV matrices can be found at the end of Sections 4, 5 and 7.2.

2 Hessenberg type matrices, zig-zag bases and Darboux factorizations

The Darboux transformation of a Jacobi matrix \mathcal{J} follows from a factorization of the symmetric matrix polynomial $\wp(\mathcal{J}) = \mathcal{J} + \beta I$ for some parameter $\beta \in \mathbb{R}$ (see Section 5). The translation of this idea to the case of a CMV matrix \mathcal{C} requires the substitution of the polynomial $\wp(x) = x + \beta$ by a Hermitian Laurent polynomial $\ell(z) = \alpha z + \beta + \bar{\alpha}z^{-1}$, $\alpha \in \mathbb{C} \setminus \{0\}$, $\beta \in \mathbb{R}$, so that the matrix Laurent polynomial $\ell(\mathcal{C}) = \alpha\mathcal{C} + \beta I + \bar{\alpha}\mathcal{C}^{-1}$ becomes self-adjoint too due to the unitarity of \mathcal{C} . A suitable factorization of $\ell(\mathcal{C})$ should provide the CMV version of the Darboux transformation.

A central role in Darboux for Jacobi matrices is played by the orthogonal polynomials on the real line since their recurrence relation is codified by a Jacobi matrix. The corresponding CMV analogue are the orthogonal Laurent polynomials on the unit circle, which are expected to be an essential ingredient in Darboux for CMV. Among other things, the orthogonal polynomials give to Jacobi matrices the meaning of a matrix representation of a symmetric multiplication operator. This is also true for orthogonal Laurent polynomials and CMV matrices, but in this case the multiplication operator is unitary instead of symmetric.

A special feature of Darboux for CMV matrices is the apparition of spurious transformations involving non-CMV (actually, even non-unitary) matrices with no band structure, which are related to non-orthogonal Laurent polynomials. Therefore, we will need to deal with general ordered sequences of Laurent polynomials (assuming no orthogonality property) and related matrix representations of a multiplication operator, which forces us to work in a general setting, avoiding any *a priori* assumption of a Hilbert space structure. This means that the space of functions that we will consider is simply the complex vector space of Laurent polynomials $\Lambda = \text{span}\{z^n\}_{n \in \mathbb{Z}} = \mathbb{C}[z, z^{-1}]$. Besides, we will work with general infinite matrices not necessarily attached to operators on Hilbert spaces, a fact that requires a special care with matrix manipulations because the algebra of general infinite matrices suf-

fers from certain pathological problems [14]. We will briefly describe some of these problems below with the aim of stating a simple setting where such pathologies can be well controlled by a few easy rules. This will be enough for our purposes, but the reader can find more comprehensive treatments of infinite matrices in classical references like [14, 54] or the recent review [66].

The first difficulty of dealing with infinite matrices is that matrix products can be ill defined, as it is the case of $S^t S$ where

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & 0 & \dots \\ 1 & 1 & 1 & 0 & 0 & \dots \\ 1 & 1 & 1 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \quad (1)$$

To avoid this problem we will restrict matrix products to those whose coefficients can be computed with a finite number of algebraic operations on the matrix coefficients of the factors.

Definition 2.1. We will say that the product AB of two matrices A, B is *admissible* if any matrix entry $(AB)_{i,k} = \sum_j A_{i,j} B_{j,k}$ involves only a finite (i, k -dependent) number of non-null summands.

The above definition applies not only to square matrices, but also to rectangular ones, so we can talk about admissible products between matrices and vectors, or between vectors. In what follows we will consider only admissible matrix products.

The following special type of infinite matrices, which will arise in the Darboux transformations for CMV matrices, provide admissible matrix products.

Definition 2.2. We will say that a matrix A is lower (upper) *Hessenberg type* if $A_{i,j}$ can be non-null only for $j - i \leq N$ ($i - j \leq N$) for some $N \in \mathbb{Z}$.

Graphically, lower Hessenberg type matrices are characterized by the shape

$$\begin{pmatrix} * & * & \dots & * & 0 & 0 & 0 & 0 & \dots \\ * & * & \dots & * & * & 0 & 0 & 0 & \dots \\ * & * & \dots & * & * & * & 0 & 0 & \dots \\ * & * & \dots & * & * & * & * & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

while upper Hessenberg type matrices correspond to the transpose of this structure.

Hessenberg type matrices are not necessarily square matrices. In particular, any column (row) vector is trivially a lower (upper) Hessenberg type matrix, and if it has finitely many non-null components then it is also an upper (lower) Hessenberg type matrix.

Any product AB with A lower Hessenberg type or B upper Hessenberg type is admissible. Instances of this are $ST = I = TS$ and $S^t T = T - I$, where I is the infinite identity matrix, S is given in (1) and

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ -1 & 1 & 0 & 0 & 0 & \dots \\ 0 & -1 & 1 & 0 & 0 & \dots \\ 0 & 0 & -1 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \quad (2)$$

Indeed, the set of lower (upper) Hessenberg type matrices is closed under multiplication. Band matrices are precisely those which are simultaneously upper and lower Hessenberg type, thus they are closed under multiplication too, providing always admissible matrix products regardless whether they act as left or right factors.

Some of the properties of the multiplication of finite matrices are also valid for admissible products of infinite matrices. For instance, the distributive property $A(B + C) = AB + AC$ holds as long as the products AB and AC are admissible, and analogously for $(B + C)A = BA + CA$. Besides, if AB is admissible, then B^+A^+ is admissible too and $(AB)^+ = B^+A^+$, where A^+ denotes the adjoint (transpose conjugate) of A .

However, even if all the involved matrix products are admissible, the associativity can fail, as it is shown by the example

$$S^t(TS) = S^tI = S^t, \quad (S^tT)S = (T - I)S = I - S.$$

The root of this problem is the non-associative character of oscillatory series which appear in the multiplication process. To see this clearly, let us fix our attention on the $(0,0)$ -entry of the above products. Denoting $X = (1, 1, 1, \dots)$,

$$(S^t(TS))_{0,0} = X(TX^t) = \sum_{j \geq 0} X_j \left(\sum_{k \geq 0} T_{j,k} X_k \right) = X_0^2 + \sum_{j \geq 1} X_j (-X_{j-1} + X_j) = 1,$$

$$((S^tT)S)_{0,0} = (XT)X^t = \sum_{k \geq 0} \left(\sum_{j \geq 0} X_j T_{j,k} \right) X_k = \sum_{k \geq 0} (X_k - X_{k+1}) X_k = 0,$$

thus the different value of these entries is a consequence of the non-associativity of the oscillatory series $1 - 1 + 1 - 1 + \dots$.

We can guarantee the associativity of matrix products requiring the absence of infinite series in the multiplication process. More precisely, assuming that the products $(AB)C$ and $A(BC)$ are admissible, a sufficient condition for the associativity $(AB)C = A(BC)$ is the existence for any indices i, l of only a finite (i, l -dependent) number of non-null terms $A_{i,j}B_{j,k}C_{k,l}$ since in that case

$$\sum_k \left(\sum_j A_{i,j} B_{j,k} \right) C_{k,l} = \sum_{j,k} A_{i,j} B_{j,k} C_{k,l} = \sum_j A_{i,j} \left(\sum_k B_{j,k} C_{k,l} \right)$$

due to the presence of a finite number of non-null summands in all the sums.

The presence of Hessenberg type matrix factors which provide admissible matrix products also guarantee their associativity according to the following simple rules.

Proposition 2.3. *The associative property $(AB)C = A(BC)$ of a matrix product is valid in any of the following cases:*

1. *A and B are lower Hessenberg type.*

2. B and C are upper Hessenberg type.

3. A is lower Hessenberg type and C upper Hessenberg type.

Proof. Suppose for instance that A and B are lower Hessenberg type. Then, AB is admissible and lower Hessenberg type too, thus $(AB)C$ is admissible. On the other hand, BC is admissible since B is lower Hessenberg type, thus $A(BC)$ is admissible because A is also lower Hessenberg type. Besides, due the lower Hessenberg type structure of A and B ,

$$\left. \begin{array}{l} A_{i,j} \text{ can be non-null only for } j - i \leq N \\ B_{j,k} \text{ can be non-null only for } k - j \leq M \end{array} \right\} \Rightarrow \begin{array}{l} A_{i,j}B_{j,k}C_{k,l} \text{ can be non-null only for} \\ j \leq i + N \text{ and } k \leq i + N + M. \end{array}$$

This ensures the associativity according to the sufficient condition given above. A similar argument proves the associativity in the two remaining cases. \square

Another pathology of admissible products of infinite matrices is that, in contrast to the case of finite matrices, the existence of inverse can be consistent with a non-trivial left or right kernel. This is illustrated by the matrix T given in (2) which, despite having the matrix (1) as an inverse, satisfies

$$(1, 1, 1, \dots)T = 0.$$

In what follows we will use the following notation for the left and right kernel of a matrix $A \in \mathbb{C}^{M \times N}$, $M, N \in \mathbb{N} \cup \{\infty\}$,

$$\ker_L(A) = \{X \in \mathbb{C}^{1 \times M} : XA = 0\}, \quad \ker_R(A) = \{X \in \mathbb{C}^{N \times 1} : AX = 0\},$$

where we understand that the vectors X are such that the products are admissible. So, in the case of the matrix T introduced in (2), $\ker_L(T) = \text{span}\{(1, 1, 1, \dots)\}$ and $\ker_R(T) = \{0\}$.

The above fact also reveals possible uniqueness problems for the inverse of a general infinite matrix because such an inverse could be modified by adding a matrix with rows and columns lying on the left and right kernel respectively. An example of this is given by

$$J = \begin{pmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 1 & 2 & 3 & \dots \\ 1 & 1 & 2 & 3 & \dots \\ 2 & 2 & 2 & 3 & \dots \\ 3 & 3 & 3 & 3 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} + c \begin{pmatrix} 1 & 1 & 1 & 1 & \dots \\ 1 & 1 & 1 & 1 & \dots \\ 1 & 1 & 1 & 1 & \dots \\ 1 & 1 & 1 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

which satisfy $JK = I = KJ$ for any $c \in \mathbb{C}$ because $\ker_L(J) = \text{span}\{(1, 1, 1, \dots)\}$ and $\ker_R(J) = \text{span}\{(1, 1, 1, \dots)^t\}$.

The previous example not only shows that a Hessenberg type matrix can have multiple inverses, but also that these inverses can be non-Hessenberg type. Nevertheless, if a lower (upper) Hessenberg type matrix A has a lower (upper) Hessenberg type inverse B , no other inverse exists. For instance, if A and B are

lower Hessenberg type, due to Proposition 2.3.1, any other inverse C should satisfy $C = (BA)C = B(AC) = B$. In the upper Hessenberg case the associativity requirements force us to modify the uniqueness proof as $C = C(AB) = (CA)B = B$ according to Proposition 2.3.2. This result is the origin of the following definition.

Definition 2.4. An infinite lower (upper) Hessenberg type matrix A with a lower (upper) Hessenberg type inverse will be called a lower (upper) *H-matrix*. Since the inverse is unique in this case, it will be denoted by A^{-1} .

Examples of H-matrices are S and T given in (1) and (2), with $S = T^{-1}$. Although H-matrices could seem quite scarce at first sight, we will see that they are precisely the kind of matrices arising in the CMV version of Darboux transformation. Actually, CMV matrices themselves are examples of matrices which are simultaneously lower and upper H-matrices.

A triangular infinite matrix with non-zero diagonal entries is a special case of H-matrix because it always has a triangular inverse. The condition on the diagonal guarantees for the leading submatrices the existence of triangular inverses obtained by enlarging the smaller ones. An induction process on the size of the leading submatrices generates the triangular inverse of the infinite triangular matrix.

In general, we will denote by A^{-1} the inverse of a matrix A whenever this inverse is unique. Some of the rules for the manipulation of inverses of finite matrices also hold for infinite ones as long as they have a unique inverse. For instance, taking adjoints we see that $AB = I = BA$ is equivalent to $B^+A^+ = I = A^+B^+$, thus A and A^+ have a unique inverse simultaneously, and in this case $(A^+)^{-1} = (A^{-1})^+$. In particular, this equality is valid for any H-matrix.

H-matrices are essential in the Darboux transformation for CMV matrices since they allow us to perform standard algebraic manipulations which are forbidden for general infinite matrices. For instance, solving a simple matrix equation in the following way,

$$AB = C \Rightarrow B = (A^{-1}A)B = A^{-1}(AB) = A^{-1}C, \quad (3)$$

is possible if A is a lower H-matrix due to Proposition 2.3.1, but it is not possible for general infinite matrices. As an application of these ideas, we have the following result.

Proposition 2.5. *If A is a lower (upper) H-matrix, then $\ker_R(A) = \{0\}$ ($\ker_L(A) = \{0\}$) and $\ker_L(A) \setminus \{0\}$ ($\ker_R(A) \setminus \{0\}$) does not contain finitely non-null vectors, i.e. vectors with finitely many non-null entries.*

Proof. Suppose that A is a lower H-matrix. Using (3) we find that $AX = 0 \Rightarrow X = 0$. On the other hand, if X is a finitely non-null row vector, it can be considered as a lower Hessenberg type matrix, thus $XA = 0 \Rightarrow X = X(AA^{-1}) = (XA)A^{-1} = 0$ due to Proposition 2.3.1. The result for upper H-matrices follows analogously from Proposition 2.3.2. \square

The set of lower (upper) H-matrices is closed under multiplication and inversion, and contains the identity matrix. Hence, this set is a multiplicative group since associativity holds for lower (upper) Hessenberg type matrices due to Proposition 2.3.1 (Proposition 2.3.2). Actually, the previous results show that the set of infinite lower (upper) Hessenberg type matrices with the sum and multiplication is a ring whose group of units is the set of lower (upper) H-matrices. Interesting subgroups of this multiplicative group are the sets of infinite lower (upper) triangular matrices with positive or non-null diagonal entries.

Following the general idea of the Darboux transformation, and mimicking the Jacobi case according to previous comments, Darboux for CMV should deal with a factorization $\ell(\mathcal{C}) = AB$ for a Hermitian Laurent polynomial $\ell(z) = \alpha z + \beta + \bar{\alpha}z^{-1}$ ($\alpha \in \mathbb{C}$, $\beta \in \mathbb{R}$), and a matrix \mathcal{C} (eventually a CMV matrix) representing in some basis of Λ the *multiplication operator*

$$\begin{aligned} z: \Lambda &\longrightarrow \Lambda \\ f(z) &\mapsto zf(z) \end{aligned} \tag{4}$$

Different matrix representations \mathcal{C} of this multiplication operator are related by changes of basis and give rise to different factors A, B . Besides, we will see that the factors A, B themselves are closely related to certain changes of basis in Λ . Thus, changes of basis will play a central role in the factorizations related to the Darboux transformation for CMV matrices.

In the following section we will analyze the matrix representations of the multiplication operator (4) for bases of Λ with a structure similar to that one of a CMV basis, but assuming no orthogonality requirement. We will see that these representations are H-matrices, a key result to avoid the pathologies of general infinite matrices.

2.1 Matrix representations of the multiplication operator

A CMV matrix can be understood as the matrix representation of the multiplication operator (4) with respect to a basis $\chi = (\chi_0, \chi_1, \chi_2, \dots)^t$ of orthonormal Laurent polynomials. This basis is obtained by applying to the canonical one $\eta = (1, z, z^{-1}, z^2, z^{-2}, \dots)^t$ the Gram-Schmidt orthonormalization with respect to a positive Borel measure supported on an infinite subset of the unit circle (hereinafter ‘measure on the unit circle’). Thus, the orthonormal basis χ satisfies

$$\chi_n \in \mathbb{L}_n := \begin{cases} \text{span}\{z^{-k}, \dots, z^k\} & (\text{coefficient of } z^{-k} > 0), & n = 2k, \\ \text{span}\{z^{-k}, \dots, z^{k+1}\} & (\text{coefficient of } z^{k+1} > 0), & n = 2k + 1. \end{cases} \tag{5}$$

Due to the needs of Darboux for CMV pointed out previously, we will consider in this section bases χ of Λ satisfying simply (5), regardless of their orthogonality properties.

Definition 2.6. A basis χ of Λ satisfying (5) will be called a *zig-zag basis*.

Due to the nested structure of the subspaces \mathbb{L}_n , for any zig-zag basis χ ,

$$z\chi_0 \in \mathbb{L}_1 \setminus \mathbb{L}_0, \quad z\chi_n \in \begin{cases} \mathbb{L}_{n+1}, & \text{even } n, \\ \mathbb{L}_{n+2} \setminus \mathbb{L}_{n+1}, & \text{odd } n. \end{cases}$$

This can be rewritten in matrix form as

$$z\chi = \mathcal{C}\chi, \quad \mathcal{C} = \begin{pmatrix} * & + & & & & & & & \\ * & * & * & + & & & & & \\ * & * & * & * & & & & & \\ * & * & * & * & * & + & & & \\ * & * & * & * & * & * & & & \\ * & * & * & * & * & * & * & + & \\ * & * & * & * & * & * & * & * & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}, \quad (6)$$

where the coefficients denoted by $+$ are positive and the omitted ones are zero. The matrix representation \mathcal{C} of the multiplication operator (4) in a zig-zag basis χ is therefore lower Hessenberg type.

The inverse z^{-1} of the multiplication operator is represented in a zig-zag basis χ by a lower Hessenberg type matrix too. Indeed,

$$z^{-1}\chi_0 \in \mathbb{L}_2 \setminus \mathbb{L}_1, \quad z^{-1}\chi_n \in \begin{cases} \mathbb{L}_{n+2} \setminus \mathbb{L}_{n+1}, & \text{even } n, \\ \mathbb{L}_{n+1}, & \text{odd } n, \end{cases}$$

or equivalently

$$z^{-1}\chi = \tilde{\mathcal{C}}\chi, \quad \tilde{\mathcal{C}} = \begin{pmatrix} * & * & + & & & & & & \\ * & * & * & & & & & & \\ * & * & * & * & + & & & & \\ * & * & * & * & * & & & & \\ * & * & * & * & * & * & + & & \\ * & * & * & * & * & * & * & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \quad (7)$$

On the other hand, from Proposition 2.3.1 we find that $(\tilde{\mathcal{C}}\mathcal{C} - I)\chi = (z^{-1}z - 1)\chi = 0$, thus $\tilde{\mathcal{C}}\mathcal{C} = I$ by the linear independence of χ . A similar proof yields $\mathcal{C}\tilde{\mathcal{C}} = I$. Hence, $\tilde{\mathcal{C}}$ is a lower Hessenberg type inverse of \mathcal{C} , which thus has a unique inverse $\mathcal{C}^{-1} = \tilde{\mathcal{C}}$.

In consequence, the matrix representation \mathcal{C} of the multiplication operator in a zig-zag basis χ is always an H-matrix.

The simplest case corresponds to the canonical basis η , which leads to

$$z\eta = \mathcal{S}\eta, \quad \mathcal{S} = \begin{pmatrix} 0 & 1 & & & & & & & \\ 0 & 0 & 0 & 1 & & & & & \\ 1 & 0 & 0 & 0 & & & & & \\ 0 & 0 & 0 & 0 & 0 & 1 & & & \\ 0 & 0 & 1 & 0 & 0 & 0 & & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \quad (8)$$

The matrix \mathcal{S} is unitary, i.e. $\mathcal{S}^{-1} = \mathcal{S}^+$. We will refer to \mathcal{S} as the *shift matrix* since it represents the multiplication shift operator $z^n \mapsto z^{n+1}$ with respect to the basis η .

Any zig-zag basis χ is characterized up to a positive factor by the corresponding matrix representation \mathcal{C} of the multiplication operator. A way to understand this is to note that \mathcal{C} determines $\tilde{\mathcal{C}} = \mathcal{C}^{-1}$, and the first and even equations of (6) combined

with the odd equations of (7) give each Laurent polynomial χ_n as a linear combination of the previous ones $\chi_0, \dots, \chi_{n-1}$ and

$$\begin{aligned} z\chi_0 & n = 0, \\ z\chi_{n-2} & \text{odd } n, \\ z^{-1}\chi_{n-2} & \text{even } n > 0. \end{aligned}$$

Therefore, \mathcal{C} and χ_0 determine χ_n for $n \geq 1$.

The matrix representations of the multiplication operator in a zig-zag basis do not cover all the H-matrices with the shape given in (6). A simple counterexample is obtained by slightly perturbing the shift matrix changing 0 by 1 in the (1,2) and (2,3) coefficients,

$$H = \begin{pmatrix} 0 & 1 & & & & & & & & & \\ 0 & 0 & \textcircled{1} & 1 & & & & & & & \\ 1 & 0 & 0 & \textcircled{1} & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & & & & \\ 0 & 0 & 1 & 0 & 0 & 0 & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & & \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}, \quad H^{-1} = \begin{pmatrix} 0 & -1 & 1 & 0 & \textcircled{1} & & & & & & \\ 1 & 0 & 0 & & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 1 & & & & & \\ 0 & 1 & 0 & 0 & -1 & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & & & \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Thus, H is an H-matrix with the shape (6), but its inverse does not fit with the structure (7) due to the encircled coefficient 1 at site (0,4).

Changes of basis allow us to specify the set of matrix representations under study. By definition, zig-zag bases χ are those related to the canonical one η by

$$\chi = T\eta, \quad T \in \mathcal{T} = \begin{array}{l} \text{set of lower triangular infinite matrices} \\ \text{with positive diagonal entries.} \end{array}$$

Therefore, Proposition 2.3.1 leads to $(\mathcal{C}T - T\mathcal{S})\eta = \mathcal{C}\chi - z\chi = 0$, so the linear independence of η implies that $\mathcal{C}T = T\mathcal{S}$. Then, using again Proposition 2.3.1 yields $\mathcal{C} = T\mathcal{S}T^{-1}$, where parenthesis are omitted due to the associativity of the product.

Hence, the matrix representations of the multiplication operator in a zig-zag basis constitute the equivalence class

$$[\mathcal{S}] = \{T\mathcal{S}T^{-1} : T \in \mathcal{T}\}$$

of the shift matrix with respect to the relation ‘‘conjugation by the subgroup \mathcal{T} ’’, which is an equivalence relation in the multiplicative group of lower H-matrices due to the associativity properties of Proposition 2.3.

Another subgroup of H-matrices with special interest for us is the set of lower (upper) H-matrices A which are unitary, i.e. $A^{-1} = A^+$. Every such a matrix A is necessarily banded because $A = (A^{-1})^+$ is simultaneously lower and upper Hessenberg type. This ensures that both A^+A and AA^+ are admissible products for unitary H-matrices. However, analyzing the unitarity of an H-matrix whose band structure is not guaranteed *a priori* requires avoiding to deal with the product A^+A (AA^+) for a lower (upper) H-matrix A , since it could be non-admissible. Nevertheless, the following result shows that the unitarity of an H-matrix can be checked resorting only to admissible products.

Proposition 2.7. *A lower (upper) H-matrix A is unitary iff $AA^+ = I$ ($A^+A = I$).*

Proof. Suppose that A is a lower H-matrix, so that it has a unique inverse A^{-1} and this inverse is lower Hessenberg type. Then, $A^{-1} = A^+$ implies that $AA^+ = I$. Conversely, $AA^+ = I$ leads to $A^{-1} = A^{-1}(AA^+) = (A^{-1}A)A^+ = A^+$, where we have used Proposition 2.3.1. A similar proof runs for upper H-matrices. \square

The unitary representatives of the class $[\mathcal{S}]$ have a special meaning.

Proposition 2.8. *If $C \in [\mathcal{S}]$ is unitary, then it is a band matrix with zig-zag shape*

$$C = \begin{pmatrix} * + & & & & & & & & \\ * * * + & & & & & & & & \\ + * * * & & & & & & & & \\ & * * * + & & & & & & & \\ & + * * * & & & & & & & \\ & & * * * + & & & & & & \\ & & + * * * & & & & & & \\ & & & * * * + & & & & & \\ & & & + * * * & & & & & \\ & & & & \dots & \dots & \dots & & \end{pmatrix}. \quad (9)$$

Proof. Bearing in mind the shapes of C and $\tilde{C} = C^{-1}$ given in (6) and (7), respectively, the result follows directly from the equality $C^+ = C^{-1}$. \square

It was proved in [12, Theorem 3.2] that the unitary matrices with the zig-zag shape (9) are exactly the CMV matrices. Therefore, we get the following identification, which provides an alternative definition of CMV matrices.

Corollary 2.9. *C is a CMV matrix iff $C \in [\mathcal{S}]$ and is unitary.*

The previous result is an algebraic characterization of CMV matrices based on a previous characterization in terms of their shape. An explicit parametrization of CMV matrices is known, and we will introduce it later on.

Zig-zag bases which are orthonormal with respect to a measure on the unit circle yield the CMV matrix representations of the multiplication operator. Bearing in mind the one-to-one correspondence between elements of $[\mathcal{S}]$ and zig-zag bases of Λ , Corollary 2.9 implies that the unitarity of an element of $[\mathcal{S}]$ is equivalent to the orthonormality of the corresponding zig-zag basis with respect to a measure on the unit circle.

The class $[\mathcal{S}]$ constitutes the framework where we will develop the factorization leading to the Darboux transformation of CMV matrices. The study of this factorization is the objective of the following section.

2.2 Change of basis and Darboux factorization

Let us introduce the Darboux factorization for the class $[\mathcal{S}]$. For this purpose, consider a Hermitian Laurent polynomial $\ell \in \Lambda \setminus \{0\}$, i.e. such that $\ell_* = \ell$, where the substar operation in Λ is defined by

$$\ell_*(z) := \overline{\ell(1/\bar{z})}. \quad (10)$$

The general form of such an ℓ is

$$\ell(z) = \sum_{j=-d}^d \alpha_j z^j, \quad \alpha_j \in \mathbb{C}, \quad \alpha_{-j} = \bar{\alpha}_j, \quad \alpha_d \neq 0.$$

Although most of the discussions along the paper are for the simplest case $d = 1$, in this section we will not assume this restriction (more comments on this general case will appear in Section 7.1). For convenience, sometimes we will distinguish the ‘degree’ d of ℓ using the number $N_z = 2d$ of zeros of ℓ (counting multiplicity) rather than the value of d itself.

Due to the ring structure of the set of square infinite lower Hessenberg type matrices, $\ell(\mathcal{C})$ is of this type for any $\mathcal{C} \in [\mathcal{S}]$. The alluded factorization will have the form $\ell(\mathcal{C}) = AB$ for some lower Hessenberg type matrices A, B . The hermiticity of ℓ ensures that $\ell(\mathcal{C})$ is self-adjoint for a unitary \mathcal{C} , i.e. when \mathcal{C} is a CMV matrix. This situates the Darboux factorization of CMV matrices close to the standard setting of Darboux for self-adjoint operators.

Let χ be a zig-zag basis related to $\mathcal{C} \in [\mathcal{S}]$. We will see that any choice of a new zig-zag basis ω generates a factorization of $\ell(\mathcal{C})$. First, note that χ and ω are related by a triangular change of basis with positive diagonal entries, i.e.

$$\chi = A\omega, \quad A \in \mathcal{T}. \quad (11)$$

A new matrix B comes from expressing $\ell\omega$ in the basis χ ,

$$\ell\omega = B\chi. \quad (12)$$

While A is a lower H-matrix because it lies on \mathcal{T} , for the moment we only can assure that B is lower Hessenberg type. This last statement follows from the relation

$$\ell\omega_n \in \ell(\mathbb{L}_n \setminus \mathbb{L}_{n-1}) \subset \mathbb{L}_{N_z+n} \setminus \mathbb{L}_{N_z+n-1}, \quad N_z = \text{number of zeros of } \ell,$$

which shows that B has the shape

$$B = \begin{pmatrix} \overbrace{* \ * \ \dots \ *}^{N_z} \ \circledast \\ * \ * \ \dots \ * \ * \ \circledast \\ * \ * \ \dots \ * \ * \ * \ \circledast \\ * \ * \ \dots \ * \ * \ * \ * \ \circledast \\ \dots \ \dots \ \dots \ \dots \ \dots \ \dots \ \dots \end{pmatrix}, \quad (13)$$

where the symbol \circledast stands for a non-null coefficient.

The matrix B is never a lower H-matrix because it has a non-trivial right kernel. This follows from (12), which implies that $B\chi(\zeta) = 0$ for any zero ζ of ℓ . Actually, B may have no inverse at all. This occurs for instance if $\ell(1) = 0$, $\omega = \eta$ and $\chi = T\eta$ with T given by (2), because then $0 = B\chi(1) = B(1, 0, 0, \dots)^t$ so that the first column of B vanishes. Nevertheless, we will see later on that in cases of interest B becomes an upper H-matrix with a band structure.

The matrices A and B are the key ingredients of a factorization not only of $\ell(\mathcal{C})$, but also of $\ell(\mathcal{D})$, where $\mathcal{D} \in [\mathcal{S}]$ is the representation of the multiplication operator in the basis ω .

Proposition 2.10. *Let \mathcal{C}, \mathcal{D} be the matrix representations of the multiplication operator with respect to the zig-zag bases χ, ω respectively, and let ℓ be a Hermitian Laurent polynomial. Then, the matrices A, B defined by (11) and (12) satisfy the following identities:*

1. $\ell(\mathcal{C}) = AB$ and $\ell(\mathcal{D}) = BA$.
2. $\mathcal{C}A = A\mathcal{D}$ and $BC = \mathcal{D}B$.
3. $\mathcal{C} = ADA^{-1}$ and $\mathcal{D} = A^{-1}\mathcal{C}A$.

Proof. From (6), (7), the analogous relations for \mathcal{D} and \mathcal{D}^{-1} , (11), (12) and Proposition 2.3.1, we obtain

$$\begin{aligned} (\ell(\mathcal{C}) - AB)\chi &= \ell\chi - \ell A\omega = 0, \\ (\ell(\mathcal{D}) - BA)\omega &= \ell\omega - B\chi = 0, \\ (\mathcal{C}A - A\mathcal{D})\omega &= \mathcal{C}\chi - zA\omega = z\chi - zA\omega = 0, \\ (BC - \mathcal{D}B)\chi &= zB\chi - \ell\mathcal{D}\omega = zB\chi - z\ell\omega = 0. \end{aligned}$$

Relations 1 and 2 follow from the linear independence of χ and ω . The remaining equalities are a consequence of the first identity in 2 and Proposition 2.3.1. \square

Parenthesis are omitted in Proposition 2.10.3 due to the associativity of products of lower H-matrices. We will use this standard convention for associative products in what follows, making explicit the parenthesis only when possible non-associative products appear.

Proposition 2.10.1 is what we will call the Darboux factorization for the matrices in the class $[\mathcal{S}]$. It is worth remarking that this factorization is not defined for single elements of $[\mathcal{S}]$, but for every ordered pair of matrices in this class. Besides, there is a trivial dependence on the zig-zag basis chosen for each pair, since these basis are defined up to a positive factor. This degree of freedom amounts to the substitution $A \rightarrow cA$, $B \rightarrow c^{-1}B$, $c > 0$, in Proposition 2.10. In the next section we will study a particularization of these factorizations for CMV matrices.

Proposition 2.10.3 is a direct consequence of identifying the factor A as the change of basis between χ and ω . We cannot formulate the analogue of this statement for the second identity of Proposition 2.10.2 due to the invertibility problems of B .

The identities of Proposition 2.10.3 provide a way to obtain \mathcal{D} starting from \mathcal{C} and the factorization $\ell(\mathcal{C}) = AB$, as well as a way to obtain \mathcal{C} starting from \mathcal{D} and the factorization $\ell(\mathcal{D}) = BA$. Despite the apparent symmetry, there is a significant difference between these two identities when compared to the equality $\mathcal{C}A = A\mathcal{D}$ of Proposition 2.10.2 which originates them. As in (3), the equation $\mathcal{C}A = A\mathcal{Y}$ has a unique solution in the unknown matrix \mathcal{Y} because A is a lower H-matrix. However, the equation $\mathcal{X}A = A\mathcal{D}$ can have multiple solutions in the unknown matrix \mathcal{X} because $\ker_L(A)$ can be non-trivial. The general solution has the form $\mathcal{X} = \mathcal{C} + K_A$ where $\mathcal{C} = ADA^{-1}$ and K_A is any matrix whose rows belong to $\ker_L(A)$. Of course, due to

Proposition 2.3.1, we know that imposing a lower Hessenberg type structure on \mathcal{X} leaves only the solution \mathcal{C} . The conclusion is that no non-null lower Hessenberg type matrix K_A can be generated from $\ker_L(A)$. This is in agreement with Proposition 2.5.

We will finish this section with an exact description of $\ker_R(B)$ which will be of interest later on.

Proposition 2.11. *Let \mathcal{C} be the matrix representation of the multiplication operator with respect to the zig-zag basis χ and let ℓ be a Hermitian Laurent polynomial with N_z zeros counting multiplicity. If ω is any other zig-zag basis and B is the matrix given by (12), then $\ker_R(\ell(\mathcal{C})) = \ker_R(B)$ and a basis of this subspace is given by*

$$\bigcup_{\substack{\zeta \text{ zero of } \ell \\ m \text{ multiplicity of } \zeta}} \{\chi(\zeta), \chi'(\zeta), \dots, \chi^{(m-1)}(\zeta)\}.$$

Moreover, a matrix of $\mathcal{K}_B = \{K = K^+ : BK = 0\}$ vanishes iff its leading submatrix of order N_z is null.

Proof. First of all note that, due to the structure (13) of B , $\dim \ker_R(B)$ cannot be greater than the number N_z of zeros of ℓ .

Besides, using the fact that A is a lower H-matrix and B is lower Hessenberg type we find from Proposition 2.3.1 that $(AB)X = 0 \Leftrightarrow BX = 0$. Bearing in mind Proposition 2.10.1, this means that $\ker_R(\ell(\mathcal{C})) = \ker_R(B)$.

Taking derivatives in (12) we conclude that $\chi^{(j)}(\zeta) \in \ker_R(B)$ for any zero ζ of ℓ and any $j = 0, 1, \dots, m-1$, smaller than its multiplicity m . Since the set

$$\bigcup_{\substack{\zeta \text{ zero of } \ell \\ m \text{ multiplicity of } \zeta}} \{\chi(\zeta), \chi'(\zeta), \dots, \chi^{(m-1)}(\zeta)\}$$

has N_z vectors lying on $\ker_R(B)$ and $\dim \ker_R(B) \leq N_z$, to prove the first statement of the proposition it only remains to see that this set is linearly independent.

For this purpose it is enough to show that $\det \Delta_{N_z} \neq 0$ where, denoting by ζ_1, \dots, ζ_k the zeros of ℓ , Δ_{N_z} is the leading submatrix of order N_z of the matrix Δ of order $\infty \times N_z$ given by

$$\Delta = (\Delta(\zeta_1), \Delta(\zeta_2), \dots, \Delta(\zeta_k)), \quad \begin{aligned} \Delta(\zeta) &= (\chi(\zeta), \chi'(\zeta), \dots, \chi^{(m-1)}(\zeta)), \\ m &= \text{multiplicity of } \zeta \text{ as a zero of } \ell. \end{aligned}$$

If $\det \Delta_{N_z} = 0$ then there exists a row vector $X \neq 0$ such that $X\Delta_{N_z} = 0$. This is also equivalent to state that the Laurent polynomial $f = X(\chi_0, \chi_1, \dots, \chi_{N_z-1})^t$ satisfies $f^{(j)}(\zeta) = 0$ for any zero ζ of ℓ and any $j = 0, 1, \dots, m-1$, smaller than its multiplicity m . In other words, f should have at least N_z zeros, counting multiplicities. This is impossible because f is a linear combination of $\chi_0, \dots, \chi_{N_z-1}$, thus it has no more than $N_z - 1$ zeros due to the zig-zag structure of the basis χ .

Suppose now that an infinite square matrix K satisfies $BK = 0$. This is equivalent to state that the columns of K belong to $\ker_R(B)$, i.e. $K = \Delta V$ for a matrix V of size $N_z \times \infty$. Indicating by the subindex N_z the leading submatrix of order N_z , we have that $K_{N_z} = \Delta_{N_z} V_{N_z}$. Since $\det \Delta_{N_z} \neq 0$, the condition $K_{N_z} = 0$ implies that $V_{N_z} = 0$, so the first N_z columns of V and K are null. If, besides, $K = K^+$, then the first N_z rows of K vanish too, which can be expressed as $\Delta_{N_z} V = 0$. The condition $\det \Delta_{N_z} \neq 0$ implies that $V = 0$, thus $K = 0$. \square

We have described the Darboux factorization in the general framework of matrix representations of the multiplication operator with respect to zig-zag bases. In the next section we will explore some consequences of the previous results for the case of CMV matrix representations, i.e. when the zig-zag bases are orthonormal with respect to a measure on the unit circle. Our purpose is to show that suitable changes of basis connect standard factorizations (usual in the context of Darboux for self-adjoint operators) with well known transformations of measures on the unit circle.

3 Darboux for CMV and Christoffel transformations

In this section we will analyze an especially interesting Darboux factorization for CMV matrices which is behind the CMV version of Darboux transformations.

The zig-zag bases leading to CMV matrices are precisely those which are orthonormal with respect to a measure on the unit circle. This establishes a one-to-one correspondence between CMV matrices and such measures, up to normalization.

Apart from introducing Darboux transformations for CMV matrices, one of the purposes of the present section is to translate these transformations to the corresponding measures. However, in tackling the CMV version of Darboux transformations we will need to deal with zig-zag bases which are not necessarily orthogonal with respect to a measure on the unit circle. This forces us to work in the more general setting of linear functionals in Λ , a subset of which can be identified with the set of measures on the unit circle. Besides, this general setting will allow us to simplify the notation along the paper, providing at the same time a direct extension of Darboux transformations to quasi-CMV matrices related to quasi-definite functionals which are not necessarily associated with positive measures (see Section 7.2). Thus, we will first comment on the relation between measures on the unit circle and linear functionals in Λ .

Any measure μ on the unit circle generates a linear functional u in Λ defined by $u[f] = \int f d\mu$, so that no different measures give rise to the same functional [37]. For convenience, we will write $u \equiv d\mu$ to indicate the functional u generated by μ . Behind this notational convention lies the identification of a measure on the unit circle and the corresponding functional, much in the same way as in the case of functions and distributions.

The inner product in Λ associated with the measure μ can be rewritten in terms of its functional u as $\int f\bar{g} d\mu = u[fg_*]$, the substar operation in Λ being as in (10). This provides a one-to-one correspondence between measures on the unit circle and positive definite Hermitian linear functionals in Λ , i.e. those linear functionals $u: \Lambda \rightarrow \mathbb{C}$ satisfying

$$u = u_* \quad (\text{Hermitian}), \quad u[ff_*] > 0, \quad f \in \Lambda \setminus \{0\} \quad (\text{positive definite}),$$

where the substar operation is defined for linear functionals in Λ by

$$u_*[f] = \overline{u[f_*]}, \quad f \in \Lambda.$$

Equivalently, $u_*[z^n] = \overline{u[z^{-n}]}$ for every $n \in \mathbb{Z}$.

We will extend the substar operation to Laurent polynomial matrices $M(z)$ by $M_*(z) := \overline{M(1/\bar{z})}$, defining also the new operation $M^+(z) := M_*(z)^t = M(1/\bar{z})^+$. For convenience, in what follows M^+ will refer to $M^+(z)$, while the adjoint of $M(z)$ will be explicitly denoted by $M(z)^+$. Then, using the natural notation $u[M] = (u[M_{i,j}])$, the orthonormality of a zig-zag basis χ with respect to a measure μ on the unit circle can be compactly expressed as $u[\chi\chi^+] = I$ in terms of the functional $u \equiv d\mu$. This orthonormal basis χ is completely determined by the positive definite functional u .

If χ is the zig-zag basis which is orthonormal with respect to a measure μ , the functional $u \equiv d\mu$ is determined by the conditions $u[\chi_n\chi_{0*}] = \delta_{n,0}$. Actually, these conditions determine a linear functional u in Λ for any zig-zag basis χ (hereinafter ‘the functional of the zig-zag basis’). Since the zig-zag basis χ of a matrix $C \in [\mathcal{S}]$ is determined up to a positive factor, this associates a unique functional u , up to normalization, to any matrix $C \in [\mathcal{S}]$. When C is a CMV matrix, u is the orthogonality functional of χ , otherwise u can be even non-Hermitian.

The Gram matrix of a general linear functional u in Λ with respect to an arbitrary basis $l = (l_0, l_1, \dots)^t$ of Λ is defined as $u[ll^+]$, so that it coincides with the standard Gram matrix of an inner product when u is positive definite. The Hermitian functionals are precisely those with a Hermitian Gram matrix and, among them, the positive definite functionals are characterized by a positive definite Gram matrix, i.e. $Xu[ll^+]X^+ > 0$ for every finitely non-null row vector $X \neq 0$ (equivalently, the leading principal minors of the Gram matrix are all positive). Therefore, a basis l of Λ is orthonormal with respect to a measure on the unit circle iff there exists a linear functional u such that $u[ll^+] = I$ because the fact that this Gram matrix is trivially Hermitian and positive definite implies that u is Hermitian and positive definite.

Any transformation of a measure on the unit circle can be understood as a transformation of the corresponding functional. For instance, if a Laurent polynomial ℓ is non-negative in the support of a positive measure μ on the unit circle, then $\ell d\mu$ is again a positive measure with associated functional $u\ell[f] := u[\ell f]$. We use this identity to extend the meaning of $u\ell$ to any linear functional u in Λ and any $\ell \in \Lambda$, non of them necessarily Hermitian. Then, it easy to see that $(u\ell)_* = u_*\ell_*$.

Let us introduce now the specific factorization which will be the origin of the Darboux transformations for CMV matrices.

Given a CMV matrix \mathcal{C} , suppose that a Hermitian Laurent polynomial ℓ of degree one makes the self-adjoint matrix $\ell(\mathcal{C})$ positive definite. This is equivalent to state that there exists a Cholesky factorization

$$\ell(\mathcal{C}) = AA^+, \quad A \in \mathcal{T},$$

which is known to be unique (see Appendix 9).

Moreover, in view of the zig-zag shape (9) of \mathcal{C} , the matrix $\ell(\mathcal{C})$ has the five-diagonal structure

$$\ell(\mathcal{C}) = \begin{pmatrix} * & * & \otimes & & & & & & \\ * & * & * & \otimes & & & & & \\ \otimes & * & * & * & \otimes & & & & \\ & \otimes & * & * & * & \otimes & & & \\ & & \otimes & * & * & * & \otimes & & \\ & & & \otimes & * & * & * & \otimes & \\ & & & & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (14)$$

with non-null entries in the upper and lower diagonals. This implies that A is not only lower triangular with positive main diagonal, but also 3-band with non-null entries in the lower subdiagonal,

$$A = \begin{pmatrix} + & & & & & & & & \\ * & + & & & & & & & \\ \otimes & * & + & & & & & & \\ & \otimes & * & + & & & & & \\ & & \otimes & * & + & & & & \\ & & & \otimes & * & + & & & \\ & & & & \dots & \dots & \dots & & \end{pmatrix}. \quad (15)$$

The above Cholesky factorization is a particular case of the Darboux factorization analyzed in the previous section. To show this, consider a zig-zag basis χ related to \mathcal{C} and a new one ω given by $\chi = A\omega$. Then, Proposition 2.10.1 shows that $\ell(\mathcal{C}) = AB$, where B is the lower Hessenberg type matrix defined by $\ell\omega = B\chi$. Since $\ell(\mathcal{C}) = AA^+$ and A is a lower H-matrix we conclude by Proposition 2.3.1 that $B = A^+$. Hence, if \mathcal{D} is the matrix representation of the multiplication operator in the basis ω , Proposition 2.10 reads in this case as

$$\ell(\mathcal{C}) = AA^+, \quad \ell(\mathcal{D}) = A^+A, \quad (16)$$

$$\mathcal{C}A = AD, \quad \mathcal{C}^+A = AD^+, \quad (17)$$

$$\mathcal{C} = ADA^{-1}, \quad \mathcal{D} = A^{-1}\mathcal{C}A. \quad (18)$$

Thus, the triangular matrix A coming from the Cholesky factorization of $\ell(\mathcal{C})$ provides a ‘reversed’ Cholesky factorization of $\ell(\mathcal{D})$. Nevertheless, as we pointed out after Proposition 2.10, the matrix \mathcal{D} can be obtained directly via the second identity in (18). Note that the reversed factorization is admissible because A is a band matrix. The freedom of the zig-zag basis $\chi \rightarrow c\chi$ in a positive factor c changes $\omega \rightarrow c\omega$, so it does not alter the matrix \mathcal{D} neither the relations (16), (17), (18).

Summarizing, any Hermitian Laurent polynomial ℓ of degree one defines a mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ between the set

$$\mathcal{C}_\ell := \{\mathcal{C} \text{ CMV} : \ell(\mathcal{C}) \text{ is positive definite}\}$$

and the class $[\mathcal{S}]$. This mapping satisfies (16), (17), (18) with A as in (15), and, up to a positive rescaling, any zig-zag bases χ, ω associated with \mathcal{C}, \mathcal{D} , respectively, are related by

$$\chi = A\omega, \quad \ell\omega = A^+\chi. \quad (19)$$

The mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ can be defined using the first relation of (16) and the second one of (18) together with the condition $A \in \mathcal{T}$.

The following questions regarding the mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ appear naturally:

- (Q1) When is \mathcal{D} a CMV matrix? This is equivalent to ask about the unitarity of \mathcal{D} , or, alternatively, about the orthonormality of ω .
- (Q2) Is $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ an isospectral mapping? If not, what changes may it induce in the spectrum?
- (Q3) What is the relation between two linear functionals u, v associated with the matrices \mathcal{C}, \mathcal{D} ? In other words, what is the transformation $u \xrightarrow{\ell} v$ generated by the mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$?
- (Q4) Relations (19) yield $\chi_n \in \text{span}\{\omega_{n-2}, \omega_{n-1}, \omega_n\}$ and $\ell\omega_n \in \text{span}\{\chi_n, \chi_{n+1}, \chi_{n+2}\}$. Does any of these conditions characterize the relation between \mathcal{C} and \mathcal{D} given by the mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$?
- (Q5) It is known that CMV matrices are parametrized by sequences of complex numbers in the unit disk, the so-called Schur parameters. If \mathcal{D} is a CMV matrix, how is the relation between the Schur parameters of \mathcal{C} and \mathcal{D} encoded in the factor A ?

The aim of the rest of this section is to answer these questions.

The answer to (Q1) follows easily from the properties of the mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$, using Proposition 2.3 to take care of the associativity.

Proposition 3.1. *The mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ preserves the set \mathcal{C}_ℓ , i.e. it transforms every CMV matrix \mathcal{C} with $\ell(\mathcal{C})$ positive definite into a CMV matrix \mathcal{D} with $\ell(\mathcal{D})$ positive definite.*

Proof. Since \mathcal{C} is a CMV matrix, from the second relation in (17) and Proposition 2.3.1 we obtain $\mathcal{D}^+ = A^{-1}\mathcal{C}^+A$. Combining this identity with the second relation in (18) and bearing in mind that the product of lower Hessenberg type matrices is associative, yields $\mathcal{D}\mathcal{D}^+ = A^{-1}\mathcal{C}AA^{-1}\mathcal{C}^+A = I$. Thus, Proposition 2.7 implies that \mathcal{D} is unitary and hence CMV.

From the second identity in (16), and taking into account that any finitely non-null row vector X is a lower Hessenberg type matrix, using again Proposition 2.3.1 we find that $X\ell(\mathcal{D})X^+ = XA^+AX^+ = \|AX^+\|^2 \geq 0$, so $X\ell(\mathcal{D})X^+ = 0$ iff $AX^+ = 0$, which implies that $X = 0$ because A is an H-matrix. Therefore, $\ell(\mathcal{D})$ is positive definite. \square

We will refer to the mappings $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ as the *Darboux transformations for CMV matrices*. Both, the definition of these mappings (which uses the Cholesky factorization of $\ell(\mathcal{C})$) as well as their domains (\mathcal{C}_ℓ) depend on the choice of the Hermitian Laurent polynomial ℓ . Nevertheless, a rescaling $\ell \rightarrow c\ell$, $c > 0$, changes $A \rightarrow \sqrt{c}A$, so it yields the same transformation $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$.

An important consequence of Proposition 3.1 is that it guarantees that Darboux transformations of CMV matrices can be iterated because \mathcal{C}_ℓ constitutes the class of CMV matrices \mathcal{C} for which the Cholesky factorization of $\ell(\mathcal{C})$ exists (see Appendix 9).

As for (Q3) and (Q4), Proposition 3.1 states that the transformation $u \xrightarrow{\ell} v$ preserves the positive definiteness of the functionals, so when answering these questions we can suppose without loss of generality that u, v are positive definite and χ, ω the corresponding orthonormal zig-zag bases. The answer to (Q3) and (Q4) is given by the following theorem, which is the main result of this section.

Theorem 3.2. *Let u, v be positive definite functionals in Λ with orthonormal zig-zag bases χ, ω and CMV matrices \mathcal{C}, \mathcal{D} respectively. Then, the following statements are equivalent:*

- (i) $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ for some Hermitian Laurent polynomial ℓ of degree one, i.e. $\mathcal{C} \in \mathcal{C}_\ell$ and $\mathcal{D} = A^{-1}\mathcal{C}A$, where $A \in \mathcal{T}$ comes from the Cholesky factorization $\ell(\mathcal{C}) = AA^+$.
- (ii) $v = u\ell$ for some Hermitian Laurent polynomial ℓ of degree one.
- (iii) $\chi_n \in \text{span}\{\omega_{n-2}, \omega_{n-1}, \omega_n\}$ for $n \geq 2$ and $\chi_n \notin \text{span}\{\omega_n\}$ for some n .
- (iv) There exists a Hermitian Laurent polynomial ℓ of degree one such that $\ell\omega_n \in \text{span}\{\chi_n, \chi_{n+1}, \chi_{n+2}\}$ for $n \geq 0$ and $\ell \notin \text{span}\{\chi_1, \chi_2\}$.

The Laurent polynomials ℓ mentioned in (i) and (ii) coincide up to a constant positive factor, and they also coincide with that one in (iv) up to a constant real factor. Moreover, if ℓ coincides exactly in (i) and (ii) then $\chi = A\omega$ and $\ell\omega = A^+\chi$ without any rescaling.

Proof. The implications (i) \Rightarrow (iii) and (i) \Rightarrow (iv) are already proved because we know that, up to a positive rescaling, (i) gives the relation (19) between χ and ω , and A has the structure (15).

$$\boxed{(i) \Leftrightarrow (ii)}$$

Assuming (i), we can suppose without loss of generality that $\chi = A\omega$ and $\ell\omega = A^+\chi$ by rescaling ℓ with a positive factor if necessary. Then, bearing in mind that any column (row) vector is a lower (upper) Hessenberg type matrix and using Proposition 2.3, we find that $u\ell[\omega\omega^+] = u[\omega(\ell\omega)^+] = A^{-1}u[\chi\chi^+]A = I$. Due to the uniqueness of the orthonormality functional of ω , we conclude that $v = u\ell$.

Suppose now that $v = u\ell$. Then, (6) and Proposition 2.3 imply that $v[\chi\chi^+] = u[\ell\chi\chi^+] = \ell(\mathcal{C})u[\chi\chi^+] = \ell(\mathcal{C})$. This means that, assuming u positive definite, v

is positive definite too iff $\mathcal{C} \in \mathcal{C}_\ell$. Therefore, the hypothesis of the theorem guarantee the existence of the Cholesky factorization $\ell(\mathcal{C}) = AA^+$. Also, $\chi = A\omega$ because, using again Proposition 2.3, we get $v[(A^{-1}\chi)(A^{-1}\chi)^+] = A^{-1}v[\chi\chi^+](A^+)^{-1} = A^{-1}\ell(\mathcal{C})(A^+)^{-1} = I$. Then, the previous results prove that $\ell\omega = A^+\chi$ and $\mathcal{D} = A^{-1}\mathcal{C}A$.

$(iii) \Rightarrow (ii)$

Let us look for a solution $\ell(z) = \alpha z + \beta + \gamma z^{-1}$ of $u\ell = v$, once (iii) is assumed.

First, for any Laurent polynomial of degree not bigger than one, $(u\ell - v)[\chi_n] = 0$ when $n \geq 3$ because, from (6) and the five-diagonal structure (14) of $\ell(\mathcal{C})$,

$$\begin{aligned} \ell\chi_n &= (\ell(\mathcal{C})\chi)_n \in \text{span}\{\chi_{n-2}, \dots, \chi_{n+2}\} \Rightarrow u\ell[\chi_n] = u[\ell\chi_n] = 0, \quad n \geq 3, \\ \chi_n &\in \text{span}\{\omega_{n-2}, \omega_{n-1}, \omega_n\} \Rightarrow v[\chi_n] = 0, \quad n \geq 3. \end{aligned}$$

Hence, $u\ell - v$ vanishes on the orthogonal complement of \mathbb{L}_2 with respect to u .

Second, there is a choice of ℓ such that $u\ell - v$ vanishes on \mathbb{L}_2 . Using the canonical basis $\eta^{(2)} = (1, z, z^{-1})^t$ and expressing $\ell = \eta^{(2)+}(\beta, \gamma, \alpha)^t$, the condition $(u\ell - v)[\eta^{(2)}] = 0$ becomes the linear system $u[\eta^{(2)}\eta^{(2)+}](\beta, \gamma, \alpha)^t = v[\eta^{(2)}]$. Since u is positive definite, $\det(u[\eta^{(2)}\eta^{(2)+}]) > 0$ and this system has a unique solution (α, β, γ) .

We conclude that $v = u\ell$ for a unique Laurent polynomial ℓ of degree not bigger than one. Since u and v are Hermitian, taking adjoints we find that $v = u\ell_*$. Therefore, the uniqueness of ℓ implies that $\ell = \ell_*$, so that $\gamma = \bar{\alpha}$. If $\alpha = 0$, then v is proportional to u , thus χ is proportional to ω , which contradicts the hypothesis. Therefore, ℓ is a Hermitian polynomial of degree one.

$(iv) \Rightarrow (ii)$

The condition $\ell\omega_n \in \text{span}\{\chi_n, \chi_{n+1}, \chi_{n+2}\}$ implies that $u\ell[\omega_n] = u[\ell\omega_n] = 0$ for $n \geq 1$. Hence, $(u\ell - v)[\omega_n] = 0$ for $n \geq 1$ and any $c \in \mathbb{C}$. Since $\ell \in \mathbb{L}_2 \setminus \text{span}\{\chi_1, \chi_2\}$, necessarily $u[\ell] \neq 0$. Therefore, we can choose $c \in \mathbb{R}$ such that $u\ell = v$ because $(u\ell - v)[\omega_0] = 0$ iff $c = v[1]/u[\ell]$, which is real because u, v and ℓ are Hermitian. \square

This theorem states that the Darboux transformations of CMV matrices correspond to the Laurent polynomial modifications of degree one $d\mu \xrightarrow{\ell} \ell d\mu$ for measures on the unit circle. In the particular case $\ell(z) = (z - \zeta)(z^{-1} - \bar{\zeta})$, $|\zeta| < 1$, these are known as the Christoffel transformations on the unit circle [17, 55], which preserve the whole set of positive definite functionals (i.e. \mathcal{C}_ℓ is the whole set of CMV matrices in this case) because such an ℓ is non-negative on the unit circle. Nevertheless, a general Laurent polynomial modification $d\mu \xrightarrow{\ell} \ell d\mu$ preserves positive definiteness only when ℓ is non-negative on the support of μ . The proof of (i) \Leftrightarrow (ii) in Theorem 3.2 uncovers a matrix characterization of such measures, which we enunciate separately.

Corollary 3.3. *Given a Hermitian Laurent polynomial ℓ and a positive definite functional u with CMV matrix \mathcal{C} ,*

$$u\ell \text{ is positive definite} \Leftrightarrow \ell(\mathcal{C}) \text{ is positive definite.}$$

Although the proof of this result given in Theorem 3.2 was enunciated for the case of a degree one ℓ , the proof for an arbitrary degree remains unchanged. This characterization is important when an unknown measure is determined only by its Schur parameters, which provide directly the corresponding CMV matrix.

The equivalence with Laurent polynomial modifications of measures gives information about the spectral behaviour of Darboux transformations. Let \mathcal{C} be a CMV matrix with associated measure μ and consider the Hilbert space L_μ^2 of μ -square-integrable functions. Then, \mathcal{C} can be understood as the representation of the unitary multiplication operator

$$\begin{aligned} U_\mu: L_\mu^2 &\longrightarrow L_\mu^2 \\ f(z) &\mapsto zf(z) \end{aligned} \quad (20)$$

with respect to the orthonormal basis of L_μ^2 given by the zig-zag basis χ of \mathcal{C} . Therefore, the spectrum of \mathcal{C} coincides with that of U_μ , which is the support of μ . The mass points are the corresponding eigenvalues, which are simple because their eigenvectors are spanned by the characteristic functions of the mass points [12, 67]. On the other hand, every transformation $\mu \xrightarrow{\ell} \nu$ preserves the support of μ up to the mass points located at the zeros of ℓ , which do not appear in ν . Therefore, we have the following spectral consequence of the previous theorem, which answers (Q2).

Corollary 3.4. *Given a Hermitian Laurent polynomial ℓ of degree one, the mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ preserves the spectrum for every CMV matrix $\mathcal{C} \in \mathcal{C}_\ell$, with the exception of at most two points: The spectrum of \mathcal{D} is obtained from that of \mathcal{C} by excluding the eigenvalues which are zeros of ℓ .*

In particular, $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ is a isospectral transformation whenever ℓ has its zeros outside of the unit circle. Otherwise it is almost isospectral, in the sense that it preserves the spectrum up to finitely many points. More precisely, if ℓ has its zeros on the unit circle the only spectral changes that the transformation may produce is the elimination of one or two eigenvalues depending whether ℓ has one or two different zeros.

Remember that the factor B of general Darboux factorizations was only lower Hessenberg type and not necessarily invertible. However, in the case of the Darboux transformations for CMV, the factor $B = A^+$ is not only invertible but also 3-band and upper triangular, so it is an upper H-matrix. Therefore, in this case we can complete the relations (16), (17), (18) with the additional ones

$$\mathcal{C} = B^{-1}\mathcal{D}B, \quad \mathcal{D} = BCB^{-1}.$$

Moreover, the relation between the factors A and $B = A^+$ gives information about the left kernel of A (the right one is trivial): $\ker_L(A) = \ker_R(B)^+$, where $\ker_R(B)$ is given by Proposition 2.11.

Question (Q5) refers to the explicit parametrization of CMV matrices given by [11, 67, 74] (our notation is related to that of [67] by $a_n = -\bar{\alpha}_{n-1}$, while we take the

transposed of the CMV matrix primarily used in [67]),

$$\mathcal{C} = \begin{pmatrix} -\bar{a}_0 a_1 & \bar{a}_0 \rho_1 & & & & & & & & \\ -\rho_1 a_2 & -\bar{a}_1 a_2 & -\rho_2 a_3 & \rho_2 \rho_3 & & & & & & \\ \rho_1 \rho_2 & \bar{a}_1 \rho_2 & -\bar{a}_2 a_3 & \bar{a}_2 \rho_3 & & & & & & \\ & & -\rho_3 a_4 & -\bar{a}_3 a_4 & -\rho_4 a_5 & \rho_4 \rho_5 & & & & \\ & & \rho_3 \rho_4 & \bar{a}_3 \rho_4 & -\bar{a}_4 a_5 & \bar{a}_4 \rho_5 & & & & \\ & & & & \dots & \dots & \dots & & & \end{pmatrix}, \quad a_0 = 1. \quad (21)$$

where $a_n \in \mathbb{C}$ are the so-called Schur parameters (or Verblunsky coefficients), which satisfy $|a_n| < 1$ for $n \geq 1$, and $\rho_n = \sqrt{1 - |a_n|^2}$. Schur parameters establish a one-to-one correspondence between sequences in the open unit disk and CMV matrices, thus they parametrize also the measures on the unit circle up to normalization. The Schur parameters also determine the orthonormal polynomials φ_n with respect to the corresponding functional $u \equiv d\mu$ via the forward and backward recurrence relations

$$\begin{cases} \rho_n \varphi_n(z) = z \varphi_{n-1}(z) + a_n \varphi_{n-1}^*(z), \\ \rho_n z \varphi_{n-1}(z) = \varphi_n(z) - a_n \varphi_n^*(z), \end{cases} \quad \varphi_n^*(z) = z^n \varphi_{n*}(z), \quad (22)$$

where φ_n^* is known as the reversed polynomial of φ_n . As a consequence,

$$\varphi_n(z) = \kappa_n (z^n + \dots + a_n), \quad \kappa_{n-1} = \rho_n \kappa_n, \quad \kappa_0 = \frac{1}{\sqrt{u[1]}}. \quad (23)$$

The orthonormal polynomials (ONP) are connected to the orthonormal Laurent polynomials (ONLP) by the relations

$$\chi_{2n}(z) = z^{-n} \varphi_{2n}^*(z), \quad \chi_{2n+1}(z) = z^{-n} \varphi_{2n+1}(z), \quad (24)$$

which, combined with (22), lead to

$$\rho_{2n} \chi_{2n} = \chi_{2n-1*} + \bar{a}_{2n} \chi_{2n-1}, \quad \rho_{2n} \chi_{2n-1} = \chi_{2n*} - a_{2n} \chi_{2n}. \quad (25)$$

Before answering (Q5) let us fix a notation concerning the mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$, which will be used in the rest of the paper.

CMV	Functional	ONLP	ONP	Schur	
\mathcal{C}	$u \equiv d\mu$	χ_n	φ_n	a_n	$\begin{aligned} \rho_n &= \sqrt{1 - a_n ^2} \\ \kappa_n^{-1} &= \rho_n \cdots \rho_1 \sqrt{u[1]} \end{aligned}$
\mathcal{D}	$v \equiv d\nu$	ω_n	ψ_n	b_n	$\begin{aligned} \sigma_n &= \sqrt{1 - b_n ^2} \\ \lambda_n^{-1} &= \sigma_n \cdots \sigma_1 \sqrt{v[1]} \end{aligned}$

(26)

The coefficients of the factor A will be denoted as

$$A = \begin{pmatrix} r_0 & & & & & & & \\ s_0 & r_1 & & & & & & \\ t_0 & \bar{s}_1 & r_2 & & & & & \\ & \bar{t}_1 & s_2 & r_3 & & & & \\ & & t_2 & \bar{s}_3 & r_4 & & & \\ & & & \bar{t}_3 & s_4 & r_5 & & \\ & & & & & & \ddots & \ddots & \ddots \end{pmatrix}, \quad r_n > 0, \quad s_n \in \mathbb{C}, \quad t_n \in \mathbb{C} \setminus \{0\}. \quad (27)$$

From the explicit parametrization (21) of \mathcal{C} we see that, not only the upper and lower diagonals of the five-diagonal matrix $\ell(\mathcal{C})$ are non-null, but its $(1,0)$ and $(0,1)$ matrix coefficients cannot vanish either, so that the structure (14) becomes

$$\ell(\mathcal{C}) = \begin{pmatrix} * \otimes \otimes & & & & & & & \\ \otimes * * \otimes & & & & & & & \\ \otimes * * * \otimes & & & & & & & \\ & \otimes * * * \otimes & & & & & & \\ & & \otimes * * * \otimes & & & & & \\ & & & \otimes * * * \otimes & & & & \\ & & & & \dots & \dots & \dots & \dots \end{pmatrix}. \quad (28)$$

This implies that, not only $t_n \neq 0$ for all $n \geq 0$, but also $s_0 \neq 0$.

Using (24), the relations $\chi = A\omega$ and $\ell\omega = A^+\chi$ read as (taking the reverse of the equations for even n)

$$\begin{aligned} \varphi_n(z) &= \bar{t}_{n-2} z \psi_{n-2}(z) + s_{n-1} \psi_{n-1}^*(z) + r_n \psi_n(z), & n \geq 0, \quad t_k = s_k = 0 \text{ if } k < 0. \\ z \ell(z) \psi_n(z) &= r_n z \varphi_n(z) + s_n \varphi_{n+1}^*(z) + t_n \varphi_{n+2}(z), \end{aligned}$$

If $\ell(z) = \alpha z + \beta + \bar{\alpha} z^{-1}$, identifying in the above equalities the terms of highest and lowest degree and using (23) leads to

$$\begin{aligned} r_n &= \frac{\kappa_n}{\lambda_n}, \quad t_n = \alpha \frac{\lambda_n}{\kappa_{n+2}} = \alpha \rho_{n+2} \frac{\rho_{n+1}}{r_n}, \\ b_n &= a_n - s_{n-1} \frac{\lambda_{n-1}}{\kappa_n} = a_n - s_{n-1} \frac{\rho_n}{r_{n-1}}, \quad n \geq 0, \quad a_0 = b_0 = 1. \\ \bar{\alpha} b_n &= \alpha a_{n+2} + s_n \frac{\kappa_{n+1}}{\lambda_n} = \alpha a_{n+2} + s_n \frac{r_n}{\rho_{n+1}}, \end{aligned} \quad (29)$$

The equation of the second line is the translation to the Schur parameters of the Darboux transformations for CMV matrices. The last equation is an additional constraint that appears when fixing the Laurent polynomial ℓ . It is worth noting that the identities of the first line lead to the simple relations

$$\frac{r_{n-1}}{r_n} = \frac{\rho_n}{\sigma_n}, \quad \frac{t_{n-1}}{t_n} = \frac{\sigma_n}{\rho_{n+2}}. \quad (30)$$

The factorization $\ell(\mathcal{C}) = AA^+$, explicitly written in terms of (27), reads as

$$\begin{cases} |t_{n-2}|^2 + |s_{n-1}|^2 + r_n^2 = \beta - 2\operatorname{Re}(\alpha \bar{a}_n a_{n+1}), \\ s_{n-1} \bar{t}_{n-1} + r_n s_n = \rho_{n+1} (\bar{\alpha} a_n - \alpha a_{n+2}), \\ r_n t_n = \alpha \rho_{n+1} \rho_{n+2}, \end{cases} \quad n \geq 0, \quad (31)$$

which, starting from the initial conditions $t_{-2} = t_{-1} = s_{-1} = 0$, determines inductively the coefficients r_n, s_n, t_n for $n \geq 0$. This shows that, if a solution of (31) exists, it is unique. The existence of such a solution requires the positivity of $\beta - 2\operatorname{Re}(\alpha \bar{a}_n a_{n+1}) - |s_{n-1}|^2 - |t_{n-2}|^2$ at every induction step, and it is equivalent to the positive definiteness of $\ell(\mathcal{C})$ (see Appendix 9).

4 Inverse Darboux for CMV and Geronimus transformations

In this section we will study the ‘inverse’ of the Darboux transformations for CMV matrices introduced in the previous section. More precisely, given a Hermitian Laurent polynomial ℓ of degree one and a CMV matrix \mathcal{D} , we will search for the CMV matrices \mathcal{C} which are transformed into \mathcal{D} by the mapping $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$. In view of Theorem 3.2, this amounts to characterizing the CMV matrices \mathcal{C} of the positive measures μ satisfying $\ell d\mu = d\nu$, where ν is a measure associated with \mathcal{D} . Corollary 3.4 implies that the inverse Darboux transformations are almost isospectral, since they preserve the spectrum up to the addition of at most two eigenvalues.

In contrast to Darboux transformations, the inverse Darboux transformations can have no solution or multiple solutions. For instance, if $\ell(z) = (z - \zeta)(z^{-1} - \bar{\zeta})$, $|\zeta| = 1$, no measure μ on the unit circle solves the equation $\ell d\mu = d\nu$, where $d\nu(e^{i\theta}) = d\theta/2\pi$ is the Lebesgue measure. On the other hand, under the same choice of ℓ , if $\ell d\mu = d\nu$, then $\ell(d\mu + m\delta_\zeta) = d\nu$ for any positive mass m , where δ_ζ is the Dirac delta at ζ .

The parametric structure of the set of solutions for the inverse Darboux transformations depends on the localization of the zeros ζ_1, ζ_2 of ℓ with respect to the unit circle. Due to the hermiticity of ℓ , the transformation $\zeta \rightarrow \bar{\zeta}^{-1}$ leaves invariant the set of zeros of ℓ . Hence, given a measure ν on the unit circle and a Hermitian Laurent polynomial ℓ , we have the following possibilities for the solutions μ of $\ell d\mu = d\nu$:

- Zeros outside of the unit circle: $\zeta_1 = \zeta = \bar{\zeta}_2^{-1}$, $|\zeta| < 1$.

Then, $\ell(z) = c|z - \zeta|^2$, $c \in \mathbb{R} \setminus \{0\}$, for $|z| = 1$, so $\ell d\mu = d\nu$ is solved by a unique positive measure $d\mu = d\nu/\ell$ or by no one depending whether $c > 0$ or $c < 0$.

- Zeros on the unit circle: $\zeta_1 = e^{i\theta_1}$, $\zeta_2 = e^{i\theta_2}$.

Then, $\ell(z) = c \sin(\frac{\theta - \theta_1}{2}) \sin(\frac{\theta - \theta_2}{2})$, $c \in \mathbb{R} \setminus \{0\}$, for $z = e^{i\theta}$. Thus, depending on the integrability and positivity of $d\nu/\ell$, either no positive measure solves $\ell d\mu = d\nu$, or there are infinitely many solutions

$$\begin{aligned} d\mu &= \frac{d\nu}{\ell} + m\delta_\zeta, & m \geq 0, & \quad \text{if } \zeta_1 = \zeta = \zeta_2, \\ d\mu &= \frac{d\nu}{\ell} + m_1\delta_{\zeta_1} + m_2\delta_{\zeta_2}, & m_1, m_2 \geq 0, & \quad \text{if } \zeta_1 \neq \zeta_2, \end{aligned}$$

parametrized by one or two real parameters, namely, the masses of μ at ζ_1, ζ_2 .

The first of these cases is closely related to the Geronimus transformations on the unit circle [28, 30], defined by

$$d\mu = \frac{d\nu}{\ell} + m\delta_\zeta + \bar{m}\delta_{\bar{\zeta}^{-1}}, \quad \ell(z) = (z - \zeta)(z^{-1} - \bar{\zeta}), \quad |\zeta| < 1, \quad m \in \mathbb{C}.$$

Starting from a positive measure ν on the unit circle, the Geronimus transformations yield in general a complex measure μ which generates a Hermitian linear functional $u[f] = \int f d\mu$ in Λ . Since positive definite functionals are represented by measures supported on the unit circle, this functional is positive definite iff $m = 0$, which is the case covered by the inverse Darboux transformations. Geronimus transformations can provide quasi-definite functionals u for $m \neq 0$, but the study of these cases in the setting of Darboux transformations requires their generalization to quasi-CMV matrices, i.e. to the matrices playing the role of CMV in the case of quasi-definite functionals on the unit circle (see Section 7.2).

Summarizing, the inverse Darboux transformation of a CMV matrix \mathcal{D} leads to a set

$$\mathcal{C}_\ell(\mathcal{D}) := \{\mathcal{C} \text{ CMV} : \mathcal{C} \xrightarrow{\ell} \mathcal{D}\} \subset \mathcal{C}_\ell$$

which can be eventually empty, and whose elements can be naturally labelled by at most two real parameters. These elements \mathcal{C} must satisfy the identities (16), (17), (18), and their orthonormal zig-zag bases χ must be related to an orthonormal zig-zag basis ω of \mathcal{D} by (19).

According to the results of the previous section, a matrix procedure to obtain the inverse Darboux transforms $\mathcal{C}_\ell(\mathcal{D})$ of \mathcal{D} starts by performing a reversed Cholesky factorization $\ell(\mathcal{D}) = A^+A$, $A \in \mathcal{F}$, which implies that A must have the 3-band structure (15). It is possible to prove that, similarly to the standard Cholesky factorization, the reversed one is possible iff $\mathcal{D} \in \mathcal{C}_\ell$ while, in contrast to the case of finite matrices, reversed Cholesky factorizations of infinite matrices are not unique in general (see Appendix 9). This is in agreement with the fact that $\mathcal{C}_\ell(\mathcal{D})$ can have more than one element. For each of these reversed factorizations we can obtain a matrix $\mathcal{C} = ADA^{-1} \in [\mathcal{S}]$, and any solution of the inverse Darboux transformation should be obtained in this way. Since we expect in general a parametric solution for this matrix procedure, we will refer to it as the *Darboux transformation with parameters* corresponding to the Hermitian polynomial ℓ .

Let us analyze in detail the Darboux transformation with parameters. Suppose $\mathcal{D} \in \mathcal{C}_\ell$ with a zig-zag basis ω , and consider a reversed Cholesky factorization $\ell(\mathcal{D}) = A^+A$, $A \in \mathcal{F}$. Defining the zig-zag basis $\chi = A\omega$ and the matrix B by $\ell\omega = B\chi$ leads to $B\chi = \ell\omega = \ell(\mathcal{D})\omega = A^+A\omega = A^+\chi$, so that (19) holds and $B = A^+$ due to the linear independence of χ . Then, we know by Proposition 2.10 that (16), (17) and (18) are valid for the matrix $\mathcal{C} \in [\mathcal{S}]$ related to χ . Hence, $\mathcal{C} \in \mathcal{C}_\ell(\mathcal{D})$ as long as it is a CMV matrix. However, up to now there is no reason to assume that \mathcal{C} is CMV. Bearing in mind Corollary 2.9, to prove this we only need to show that \mathcal{C} is unitary, which in view of Proposition 2.7 is equivalent to the identity $\mathcal{C}\mathcal{C}^+ = I$.

Trying for \mathcal{C} a proof similar to that of the unitarity of \mathcal{D} for the direct Darboux transformations would start from the identities $\mathcal{C}^+A = A\mathcal{D}^+$ and $\mathcal{C} = ADA^{-1}$ in (17) and (18). However, the temptation to use the first of these identities to write $\mathcal{C}^+ = A\mathcal{D}^+A^{-1}$ fails because, using Proposition 2.3, we can only deduce that

$$(\mathcal{C}^+A)A^{-1} = (A\mathcal{D}^+)A^{-1} = ADA^{-1}, \quad (32)$$

since A and \mathcal{D}^+ are band matrices. The associativity of the first term is not guaranteed by Proposition 2.3 because A is band but A^{-1} is lower Hessenberg type and \mathcal{C}^+ is upper Hessenberg type (we do not know yet if \mathcal{C} is a CMV!).

An attempt to overcome this problem resorts to the direct use of the two identities in (17), together with Proposition 2.3, as follows

$$(\mathcal{C}\mathcal{C}^+)A = \mathcal{C}(\mathcal{C}^+A) = \mathcal{C}(A\mathcal{D}^+) = (\mathcal{C}A)\mathcal{D}^+ = (A\mathcal{D})\mathcal{D}^+ = A(\mathcal{D}\mathcal{D}^+) = A. \quad (33)$$

However, the same associativity problem found in (32) reappears now when trying to prove that $\mathcal{C}\mathcal{C}^+ = I$ using (33). The only conclusion from (33) is that the rows of $\mathcal{C}\mathcal{C}^+ - I$ belong to $\ker_L(A) = \ker_R(A^+)^+$, which is known to be non trivial by Proposition 2.11. Since $\mathcal{C}\mathcal{C}^+ - I$ is Hermitian, this is equivalent to state that their columns lie on $\ker_R(A^+)$, or using the notation of Proposition 2.11,

$$\mathcal{C}\mathcal{C}^+ = I + K, \quad K \in \mathcal{K}_{A^+} = \{K = K^+ : A^+K = 0\}. \quad (34)$$

There is reason for the failure of the previous attempts to prove the unitarity of \mathcal{C} : not all the reversed Cholesky factorizations $\ell(\mathcal{D}) = A^+A$, $A \in \mathcal{F}$, lead to a unitary $\mathcal{C} = ADA^{-1}$. In other words, the Darboux transformation with parameters not only yields the set $\mathcal{C}_\ell(\mathcal{D})$ of inverse Darboux transforms of \mathcal{D} , but it also gives spurious solutions $\mathcal{C} \in [\mathcal{S}]$ which are not CMV.

Take for instance $\nu = \vartheta$ as the Lebesgue measure and $\ell(z) = (z - \zeta)(z^{-1} - \bar{\zeta})$, $|\zeta| = 1$. Then $\mathcal{D} = \mathcal{S}$ and there is no solution μ of $\ell d\mu = d\nu$, i.e. $\mathcal{C}_\ell(\mathcal{S}) = \emptyset$. However, since $\ell d\vartheta$ is a positive measure on the unit circle, we know by Corollary 3.3 that $\mathcal{S} \in \mathcal{C}_\ell$ and, hence, there exist reversed Cholesky factorizations $\ell(\mathcal{S}) = A^+A$, $A \in \mathcal{F}$, leading to spurious solutions $\mathcal{C} = ADA^{-1}$. Actually, the functionals of these spurious solutions can be even non-Hermitian. This will be illustrated with an explicit example later on.

The appearance of spurious solutions can be made apparent by rewriting explicitly the reversed factorization $\ell(\mathcal{D}) = A^+A$ in terms of (27) and $\ell(z) = \alpha z + \beta + \bar{\alpha}z^{-1}$,

$$\begin{cases} r_n^2 + |s_n|^2 + |t_n|^2 = \beta - 2\operatorname{Re}(\alpha\bar{b}_n b_{n+1}), \\ s_n r_{n+1} + t_n s_{n+1} = \sigma_{n+1}(\bar{\alpha}b_n - \alpha b_{n+2}), \\ t_n r_{n+2} = \alpha\sigma_{n+1}\sigma_{n+2}, \end{cases} \quad n \geq 0. \quad (35)$$

In contrast to (31), these relations do not determine the coefficients r_n, s_n, t_n . Actually, expressing (35) as

$$r_{n+2} = |\alpha| \frac{\sigma_{n+1}\sigma_{n+2}}{\tau_n}, \quad s_{n+1} = \frac{\bar{\alpha}}{|\alpha|} \frac{\sigma_{n+1}(\bar{\alpha}b_n - \alpha b_{n+2}) - s_n r_{n+1}}{\tau_n}, \quad t_n = \frac{\alpha}{|\alpha|} \tau_n, \quad (36)$$

$$\tau_n = \sqrt{\beta - 2\operatorname{Re}(\alpha\bar{b}_n b_{n+1}) - r_n^2 - |s_n|^2},$$

shows that r_n, s_n, t_n are determined inductively for $n \geq 0$ by the three parameters r_0, s_0, r_1 , which enclose the freedom in the reversed factorization of $\ell(\mathcal{D})$. Nevertheless, this does not mean that any choice of r_0, s_0, r_1 should give such a reversed factorization, a fact that depends on the positivity of $\beta - 2\text{Re}(\alpha \bar{b}_n b_{n+1}) - r_n^2 - |s_n|^2$ at every induction step. The existence of a solution of (35) for at least one choice of the “free” parameters r_0, s_0, r_1 is equivalent to state that $\ell(\mathcal{D})$ is positive-definite (see Appendix 9).

Therefore, the general solution of a Darboux transformation with parameters depends on two real (r_0, r_1) and one complex (s_0) initial conditions, i.e. four real parameters. This is in striking contrast with the corresponding inverse Darboux transforms, which depend on at most two real parameters. Of course, this counting is indicative of the amount of spurious solutions only up to initial conditions r_0, s_0, r_1 giving no reversed factorization $\ell(\mathcal{D}) = A^+ A$.

Before presenting an explicit example of a spurious solution, we will give a characterization which allows us to distinguish the CMV solutions from the spurious ones with a few calculations. Such a characterization is based on Proposition 2.11 and some results concerning the hermiticity of the solutions u of $u\ell = v$.

Lemma 4.1. *Let ℓ be a Hermitian Laurent polynomial of degree one and v a Hermitian linear functional in Λ with $v[1] \neq 0$. Then, a linear functional solution u of $u\ell = v$ is Hermitian iff $u[l_{i*}] = \overline{u[l_i]}$ for a basis l_0, l_1 of \mathbb{L}_1 .*

Proof. The hermiticity of u on a basis of \mathbb{L}_1 , necessary for its hermiticity on Λ , is equivalent to its hermiticity on \mathbb{L}_1 . On the other hand, the relation $u\ell = v$, together with the hermiticity of v and ℓ , leads to

$$u[(\ell f)_*] = u[\ell f_*] = v[f_*] = \overline{v[f]} = \overline{u[\ell f]}, \quad f \in \Lambda,$$

which means that u is Hermitian in the subspace $\ell\Lambda$ of Λ . This, combined with the hermiticity of u in \mathbb{L}_1 , implies its hermiticity on $\Lambda = \mathbb{L}_1 \oplus \ell\Lambda$. \square

The previous lemma has the following consequences concerning the hermiticity of the functionals associated with the solutions of the Darboux transformations with parameters.

Proposition 4.2. *Let ℓ be a Hermitian Laurent polynomial of degree one, $\mathcal{D} \in \mathcal{C}_\ell$ with Schur parameters $b_n, \sigma_n = \sqrt{1 - |b_n|^2}$ and v a related functional. Consider an arbitrary solution $\mathcal{C} \in [\mathcal{S}]$ of the corresponding Darboux transformation with parameters, i.e. $\mathcal{C} = ADA^{-1}$ with $\ell(\mathcal{D}) = A^+ A$, $A \in \mathcal{T}$. If u is the functional of a zig-zag basis χ related to \mathcal{C} , then $u\ell = v$ up to a positive rescaling of ℓ . Besides, u is Hermitian iff any of the following equivalent conditions is satisfied:*

- (i) $u[z^{-1}] = \overline{u[z]}$.
- (ii) $u[\chi_{1*}] = 0$.
- (iii) $\chi_{1*} \in \text{span}\{\chi_1, \chi_2\}$.

(iv) The matrix A has the form (27) with

$$r_0^2 = \beta - 2\text{Re}(\alpha a), \quad a = b_1 + \frac{s_0}{r_1}\sigma_1. \quad (37)$$

Proof. Assume the hypothesis of the statement and let ω be the zig-zag basis which is orthonormal with respect to v . Using Proposition 2.3 we find that $\mathcal{C}A\omega = ADA^{-1}A\omega = AD\omega = zA\omega$, which means that $A\omega$ is a zig-zag basis of \mathcal{C} , so a positive rescaling of ℓ yields $\chi = A\omega$ and $\ell\omega = A^+\chi$. Since A has the structure (27), from the latter relation we conclude that $\ell\omega_n \in \text{span}\{\chi_n, \chi_{n+1}, \chi_{n+2}\}$ and the functional u of χ satisfies $u[\ell\omega_0] = r_0u[\chi_0] > 0$ because χ_0 is a positive constant and, by definition, $u[\chi_n\chi_{0*}] = \delta_{n,0}$. As a consequence, $u[\ell] > 0$ and $(u\ell - v)[\omega_n] = cu[\ell\omega_n] - v[\omega_n] = 0$ for $n \geq 1$ and any $c \in \mathbb{C}$. Since $(u\ell - v)[\omega_0] = 0$ iff $c = v[1]/u[\ell]$, we get $u\ell = v$ for such a positive value of c .

The characterizations of the hermiticity of u given in (i) and (ii) follow from the previous result, Lemma 4.1 and the fact that $u[\chi_1] = 0$ and $u[\chi_0] = u[\chi_{0*}] > 0$ for the functional u of χ .

Since $u[\chi_n]$ is non-zero only for $n = 0$, the relation $\chi_{1*} \in \text{span}\{1, z^{-1}\} \subset \mathbb{L}_2 = \text{span}\{\chi_0, \chi_1, \chi_2\}$ shows that $u[\chi_{1*}] = 0$ iff $\chi_{1*} \in \text{span}\{\chi_1, \chi_2\}$. This proves the equivalence (ii) \Leftrightarrow (iii).

As for the equivalence (iii) \Leftrightarrow (iv), note that $\chi_{1*} \in \text{span}\{\chi_1, \chi_2\}$ iff $\{\chi_1, \chi_2, \chi_{1*}\}$ is linearly dependent because $\{\chi_1, \chi_2\}$ is linearly independent. From (25) we know that $\sigma_2\omega_2 = \omega_{1*} + \bar{b}_2\omega_1$, which, combined with $\chi = A\omega$, provides the expansion

$$\chi_{1*} = \bar{s}_0\omega_0 + r_1\omega_{1*} = \bar{s}_0\omega_0 - r_1\bar{b}_2\omega_1 + r_1\sigma_2\omega_2.$$

Since $\chi_1 = s_0\omega_0 + r_1\omega_1$ and $\chi_2 = t_0\omega_0 + \bar{s}_1\omega_1 + r_2\omega_2$, the linear dependence of $\{\chi_1, \chi_2, \chi_{1*}\}$ reads as

$$\begin{vmatrix} s_0 & r_1 & 0 \\ t_0 & \bar{s}_1 & r_2 \\ \bar{s}_0 & -r_1\bar{b}_2 & r_1\sigma_2 \end{vmatrix} = 0.$$

Introducing the expressions of t_0 , s_1 , r_2 in terms of r_0 , s_0 , r_1 given in (36), this determinantal condition becomes (iv) after direct algebraic manipulations. \square

We are now ready to prove the characterization of the CMV solutions for the Darboux transformations with parameters.

Theorem 4.3. *Let ℓ be a Hermitian Laurent polynomial of degree one, $\mathcal{D} \in \mathcal{C}_\ell$ with Schur parameters b_n , $\sigma_n = \sqrt{1 - |b_n|^2}$ and v a related functional. Consider a reversed Cholesky factorization $\ell(\mathcal{D}) = A^+A$, $A \in \mathcal{T}$, the corresponding solution $\mathcal{C} = ADA^{-1}$ of the Darboux transformation with parameters and a related zig-zag basis χ . Then, \mathcal{C} is CMV iff any of the following equivalent conditions is satisfied:*

- (i) The first two rows of \mathcal{C} constitute an orthonormal system, i.e. the leading submatrix of $\mathcal{C}\mathcal{C}^+$ of order 2 is the identity.

(ii) There exists a linear functional u which solves $u\ell = v$, up to a positive rescaling of ℓ , and such that $u[\chi^{(1)}\chi^{(1)+}]$ is the identity, where $\chi^{(1)} = (\chi_0, \chi_1)^t$.

(iii) The functional u of χ satisfies $u[\chi_{1*}] = 0$ and $u[\chi_1\chi_{1*}] = 1$.

(iv) The basis χ satisfies $\chi_{1*} \in \text{span}\{\chi_1, \chi_2\}$ and

$$(\chi_1/\chi_0)(z) = \rho^{-1}(z + a), \quad |a| < 1, \quad \rho = \sqrt{1 - |a|^2}.$$

(v) If $\ell(z) = \alpha z + \beta + \bar{\alpha}z^{-1}$, the matrix A has the form (27) with

$$r_0^2 = \beta - 2\text{Re}(\alpha a), \quad a = b_1 + \frac{s_0}{r_1}\sigma_1, \quad r_1 = r_0 \frac{\sigma_1}{\rho}, \quad |a| < 1, \quad \rho^2 = \sqrt{1 - |a|^2}. \quad (38)$$

Moreover, the parameter a in (iv) and (v) is the first Schur parameter a_1 of \mathcal{C} .

Proof. From Proposition 2.7 and Corollary 2.9, the identity $\mathcal{C}\mathcal{C}^+ = I$ is equivalent to the unitarity of \mathcal{C} and thus to stating that it is CMV. Therefore, the equivalence with (i) follows directly from (34) and the second statement of Proposition 2.11, bearing in mind that $B = A^+$ and the number of zeros of ℓ is $N_z = 2$ in the present case.

Also, \mathcal{C} is CMV iff $u[\chi\chi^+] = I$ for a linear functional u in Λ , and in this case Theorem 3.2 states that $u\ell = v$ up to a positive rescaling of ℓ , thus u satisfies (ii). To prove the converse, consider the zig-zag basis ω which is orthonormal with respect to v . As in the proof of Proposition 4.2, $\chi = A\omega$ and $\ell\omega = A^+\chi$ up to a positive rescaling of ℓ . Suppose that u is an arbitrary linear functional solution of $u\ell = v$. Using Proposition 2.3 we get $A^+ = u\ell[\omega\omega^+]A^+ = u[\ell\omega(A\omega)^+] = A^+u[\chi\chi^+]$, so that

$$u[\chi\chi^+] = I + K, \quad A^+K = 0. \quad (39)$$

If $u[\chi^{(1)}\chi^{(1)+}]$ is the identity, then $u[\chi_0] = u[\chi_{0*}] > 0$ and $u[\chi_1] = 0 = u[\chi_{1*}]$. This guarantees the hermiticity of u due to Lemma 4.1. As a consequence, $K = K^+$, i.e. $K \in \mathcal{K}_{A^+}$. Then, (39) and the second statement of Proposition 2.11 ensure that $u[\chi\chi^+] = I$ whenever the leading submatrix of $u[\chi\chi^+]$ of order 2, which is $u[\chi^{(1)}\chi^{(1)+}]$, is the identity.

Concerning the equivalence with (iii), note that a linear functional u satisfying $u[\chi\chi^+] = I$ is necessarily the functional of χ , which therefore must satisfy $u[\chi_{1*}] = 0$ and $u[\chi_1\chi_{1*}] = 1$ when \mathcal{C} is unitary. Conversely, if the functional u of χ satisfies these conditions, then $u[\chi^{(1)}\chi^{(1)+}]$ is the identity because by definition $u[\chi_n\chi_{0*}] = \delta_{n,0}$. Proposition 4.2 ensures that the functional u of χ solves $u\ell = v$ up to a positive rescaling of ℓ , thus we conclude that u satisfies (ii) and hence \mathcal{C} is CMV.

As for the equivalence with (iv), let us write explicitly $(\chi_1/\chi_0)(z) = \kappa(z + a)$, $\kappa > 0$, $a \in \mathbb{C}$. If \mathcal{C} is CMV then, from $\mathcal{C}\chi = z\chi$ and (21), we get $a = a_1$ and $\kappa = \rho_1^{-1}$. Besides, (25) gives $\rho_2\chi_2 = \chi_{1*} + \bar{a}_2\chi_1$, so $\chi_{1*} \in \text{span}\{\chi_1, \chi_2\}$. For the converse, note that the functional u of χ satisfies $u[\chi_1] = u[\chi_2] = 0$, thus $u[\chi_{1*}] = 0$ under the condition $\chi_{1*} \in \text{span}\{\chi_1, \chi_2\}$. This condition also implies that $u[\chi_1\chi_{1*}] = \kappa^2(1 - |a|^2)$

because $\chi_1\chi_{1*} = \chi_0\chi_{0*}\kappa^2(1+|a|^2+\bar{a}z+az^{-1}) = \chi_0\chi_{0*}\kappa^2(1-|a|^2) + \kappa\bar{a}\chi_1\chi_{0*} + \kappa a\chi_0\chi_{1*}$ and $u[\chi_n\chi_{0*}] = \delta_{n,0}$. Bearing in mind the equivalence with (iii), this proves that \mathcal{C} is CMV under the hypothesis given in (iv).

Finally, let us prove the equivalence with (v). If ω is the orthonormal basis with respect to v we know that $\chi = A\omega$ is a zig-zag basis of \mathcal{C} and $\ell\omega = A^+\chi$. If \mathcal{C} is CMV, using Proposition 2.3 we get $u\ell[\omega\omega^+] = u[\omega(\ell\omega)^+] = A^{-1}u[\chi\chi^+]A = I$, so $v = u\ell$. From $\mathcal{C}\chi = z\chi$ and (21) we find that $\chi_1 = \chi_0\rho_1^{-1}(z + a_1)$, thus $\ell\chi_0\chi_{0*} = \alpha\rho_1\chi_1\chi_{0*} + \bar{\alpha}\chi_0\chi_{1*} + (\beta - 2\text{Re}(\alpha a_1))\chi_0\chi_{0*}$. Since $\chi_0 = r_0\omega_0$, we get $r_0^2 = v[\chi_0\chi_{0*}] = u[\ell\chi_0\chi_{0*}] = \beta - 2\text{Re}(\alpha a_1)$, which proves the first identity of (v) for $a = a_1$. The rest of the identities follow from (29). Let us prove now the converse. The Schur parametrization (21) of \mathcal{D} combined with $\mathcal{D}\omega = z\omega$ yields $-b_1\omega_0 + \sigma_1\omega_1 = z\omega_0$. Therefore, assuming (v), $\chi = A\omega$ yields

$$\chi_0 = r_0\omega_0, \quad \chi_1 = s_0\omega_0 + r_1\omega_1 = \frac{r_0}{\rho}(a\omega_0 - b_1\omega_0 + \sigma_1\omega_1) = \chi_0\frac{z+a}{\rho},$$

which gives the second condition in (iv). To prove that \mathcal{C} is CMV we only need to show additionally that $\chi_{1*} \in \text{span}\{\chi_1, \chi_2\}$. This is equivalent to (37), which coincides with part of the conditions given in (v). \square

Proposition 4.2.(iv) and Theorem 4.3.(v) provide characterizations of those reversed factorizations $\ell(\mathcal{D}) = A^+A$ leading respectively to Hermitian or positive definite functionals, via the Darboux transformation with parameters. Both characterizations are in the same spirit, i.e. they are given as restrictions on the initial parameters r_0, s_0, r_1 determining the matrix factor A . The (hermiticity) positive definiteness restrictions have the effect of reducing from four to (three) two the number of “free” real parameters. Of course, not any choice of the parameters r_0, s_0, r_1 satisfying such restrictions leads necessarily to a reversed factorization $\ell(\mathcal{D}) = A^+A$, so the only conclusion is that the solutions of the Darboux transformation with parameters leading to (Hermitian) positive definite functionals are parametrized by at most (three) two real parameters. This is in agreement with our previous discussion based on the interpretation of inverse Darboux transformations in terms of measures on the unit circle. We will refer to (37) and (38) as the hermiticity and the CMV conditions respectively. Note that the hermiticity condition (37) coincides with the first part of the CMV conditions (38). Therefore, the spurious solutions leading to Hermitian functionals are characterized by satisfying (37), but not the rest of the CMV conditions (38), i.e. either $|a| \geq 1$ or $r_1\rho \neq r_0\sigma_1$.

When the CMV conditions are taken into account, the Darboux transformations with parameters can be iterated because the set $\mathcal{C}_\ell(\mathcal{D})$ of CMV solutions is a subset of \mathcal{C}_ℓ , the class of CMV matrices \mathcal{C} which allow for a reversed Cholesky factorization of $\ell(\mathcal{C})$ (see Appendix 9).

We will finish this section illustrating in a detailed example the coexistence of CMV and spurious solutions of the Darboux transformation with parameters. To develop the example it will be useful to rewrite the relations (29) from the perspective

guaranteeing that $r_0^2 > 0$.

Due to the positivity requirement of x_n , the initial condition must satisfy $0 < x_0 < 2(1 + b^2) = \xi$. The recurrence relation (42) has two fixed points, $\xi_{\pm} = (1 \pm b)^2 < \xi$, which provide its constant positive solutions $x_n = \xi_{\pm}$. As for the non-constant solutions x_n , a quick analysis of their positivity follows from the interpretation of (42) as the implementation of Newton's method to the function h given by

$$h(x) = \frac{|x - \xi_-|^{\frac{\xi_+}{4b}}}{|x - \xi_+|^{\frac{\xi_-}{4b}}}.$$

The typical behaviour of h is represented in Figure 1. It is a non-negative concave function with a zero at ξ_- , a divergence at ξ_+ and a local minimum at $\xi = 2(1 + b^2)$. This guarantees that x_n converges monotonically to ξ_- whenever $x_0 < \xi_+$, being strictly increasing if $x_0 < \xi_-$ and strictly decreasing if $\xi_- < x_0 < \xi_+$. On the other hand, every $x_0 > \xi_+$ leads eventually to $x_n > \xi$ for some n , thus no such a choice yields a positive solution of (42). Therefore, the positive solutions x_n of (42) are those associated with an initial condition $x_0 \in (0, \xi_+]$, and, up to the fixed point $x_0 = \xi_+$, all these solutions converge monotonically to ξ_- .

From these results, and bearing in mind the restriction (43), we conclude that the solutions of the Darboux transformation with parameters corresponding to $s_1 = 0$ are given by (41) for any choice of r_1, r_2 with the constraints

$$0 < r_1, r_2 \leq 1 + b, \quad \frac{1 - b}{r_1^2} + \frac{1 + b}{r_2^2} < \frac{2}{(1 - b)^2}. \quad (44)$$

Let us discuss now the restrictions imposed on these solutions by the hermiticity and CMV conditions. According to (37), the solutions associated with an Hermitian linear functional are characterized by

$$r_0^2 = 2(1 + a), \quad a = b - \frac{\sigma^2(1 - b)}{r_1^2}. \quad (45)$$

Taking into account the expression of r_0 in terms of r_1, r_2 given in (41), the hermiticity condition becomes $r_1 = r_2$. In view of (44), the solutions of the Darboux transformations leading to Hermitian functionals are given by (41) with

$$r_1 = r_2 = r \in (1 - b, 1 + b],$$

and the related value of r_0 is

$$r_0^2 = 2(1 + b) \left(1 - \frac{(1 - b)^2}{r^2} \right).$$

The above solutions include the CMV ones which, according to (38), are characterized by the additional conditions $|a| < 1$ and $r\rho = r_0\sigma$, $\rho = \sqrt{1 - a^2}$. Using (45) we obtain

$$|a| < 1 \Leftrightarrow r > 1 - b, \quad r\rho = r_0\sigma \Leftrightarrow r = 1 + b.$$

giving the Cholesky factorization $\ell(\mathcal{C}) = AA^+$, has the form

$$A = \begin{pmatrix} r_0 & & & & \\ s_0 & r & & & \\ t & 0 & r & & \\ & t & 0 & r & \\ & & t & 0 & r \\ & & & \ddots & \ddots & \ddots \end{pmatrix}, \quad \begin{aligned} r_0 &= 2\sqrt{\frac{b}{1+b}}, & s_0 &= -\sigma\frac{1-b}{1+b}, \\ r &= 1+b, & t &= -(1-b). \end{aligned}$$

The CMV matrix \mathcal{D} is related to an absolutely continuous measure ν supported on the arc $\Gamma = \{e^{i\theta} : |\sin \frac{\theta}{2}| \geq b\}$ given by

$$d\nu(e^{i\theta}) = w(\theta) d\theta, \quad w(\theta) = \frac{\sqrt{\sin^2 \frac{\theta}{2} - b^2}}{|\sin \frac{\theta}{2}|}, \quad e^{i\theta} \in \Gamma.$$

Since $\ell(z) = (z-1)(z^{-1}-1)$ has a single zero at $z=1$, the above CMV solution \mathcal{C} of the Darboux transformation with parameters must be associated with a measure $d\mu = d\nu/\ell + m\delta_1$ for some $m \geq 0$.

To analyze the value of the mass m , let us have a look at the Schur function f of μ . The Schur function f has an analytic continuation through the essential gap $\{e^{i\theta} : |\sin \frac{\theta}{2}| < b\}$, and $z=1$ is a mass point of μ iff it is a solution of $zf(z) = 1$, i.e. $f(1) = 1$. On the other hand, Geronimus' theorem asserts that the application of the Schur algorithm to f generates the sequence $(-a, -b, -b, -b, \dots)$. Analogously, the Schur function g of ν is characterized by the constant sequence $(-b, -b, -b, \dots)$ arising from the Schur algorithm. Therefore, g is obtained from f after a single step of the Schur algorithm,

$$g(z) = \frac{1}{z} \frac{f(z) + a}{1 + af(z)}.$$

This relation implies that $f(1) = 1$ iff $g(1) = 1$, which is not possible because 1 is not a mass point of ν . We conclude that $m = 0$ and

$$d\mu(e^{i\theta}) = \frac{w(\theta)}{|e^{i\theta} - 1|^2} d\theta = \frac{w(\theta)}{2 \sin^2 \frac{\theta}{2}} d\theta, \quad e^{i\theta} \in \Gamma.$$

In other words, the arc Γ is the common spectrum of \mathcal{C} and \mathcal{D} , which are isospectral.

The spurious solutions can be also explicitly described. For instance, the choice $s_1 = 0$, $r_1 = 1-b$, $r_2 = 1+b$ satisfies (44) and yields $s_n = 0$, $r_{2n-1} = r_1$, $r_{2n} = r_2$ for $n \geq 1$, so that

$$A = \begin{pmatrix} r_0 & & & & \\ s_0 & r_1 & & & \\ -r_1 & 0 & r_2 & & \\ & -r_2 & 0 & r_1 & \\ & & -r_1 & 0 & r_2 \\ & & & \ddots & \ddots & \ddots \end{pmatrix}, \quad \begin{aligned} r_0 &= 2\sqrt{b}, & s_0 &= -\sigma, \\ r_1 &= 1-b, & r_2 &= 1+b. \end{aligned}$$

Since $r_1 \neq r_2$, the corresponding solution $\mathcal{C} = ADA^{-1}$ is related to a non-Hermitian

functional. Explicitly,

$$\mathcal{C} = \begin{pmatrix} 1 & \sigma\sqrt{b}\frac{2}{1-b} & 0 & 0 & 0 & 0 & \dots \\ \sigma\sqrt{b}\frac{2b}{1+b} & b & -\sigma b\frac{1-b}{1+b} & \sigma^2 & 0 & 0 & \dots \\ \sqrt{b}\frac{2(1+b^2)}{1-b} & \sigma\frac{b^2+4b-1}{(1-b)^2} & -b^2 & \sigma b\frac{1+b}{1-b} & 0 & 0 & \dots \\ \sigma\sqrt{b}\frac{4b}{(1+b)^2} & 0 & \sigma b\frac{b^2+4b-1}{(1+b)^2} & -b^2 & -\sigma b\frac{1-b}{1+b} & \sigma^2 & \dots \\ \sqrt{b}\frac{4b(1+b)}{(1-b)^2} & \sigma\frac{8b^2}{(1-b)^3} & \sigma^2 & \sigma b\frac{1+4b-b^2}{(1-b)^2} & -b^2 & \sigma b\frac{1+b}{1-b} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

which, solving $\mathcal{C}\chi = z\chi$, yields an associated zig-zag basis χ with

$$\chi_0 = 1, \quad \chi_1 = \frac{1-b}{2\sigma\sqrt{b}}(z-1), \quad \chi_2 = \frac{1}{2\sqrt{b}(1-b)}(bz+3b-1+z^{-1}), \quad \dots$$

Hence, the conditions $u[\chi_n\chi_{0*}] = \delta_{n,0}$ defining the associated functional u give

$$u[1] = 1 = u[z], \quad u[z^{-1}] = 1 - 4b \neq \overline{u[z]},$$

which show directly the lack of hermiticity of u .

5 Jacobi versus CMV

The standard procedure for the Darboux transformation of a Jacobi matrix

$$\mathcal{J} = \begin{pmatrix} * & + & & & \\ + & * & + & & \\ & + & * & + & \\ & & \dots & \dots & \dots \end{pmatrix} = \mathcal{J}^+$$

uses LU instead of Cholesky factorizations. Actually, the usual starting point is not the Jacobi matrix itself, but the tridiagonal one

$$\tilde{\mathcal{J}} = \begin{pmatrix} * & 1 & & & \\ + & * & 1 & & \\ & + & * & 1 & \\ & & \dots & \dots & \dots \end{pmatrix} \quad (46)$$

obtained by conjugating \mathcal{J} with a positive diagonal matrix $\Pi_{\mathcal{J}}$, i.e. $\tilde{\mathcal{J}} = \Pi_{\mathcal{J}}\mathcal{J}\Pi_{\mathcal{J}}^{-1}$. The Darboux transformation starts by choosing $\beta \in \mathbb{R}$ such that the LU factorization $\tilde{\mathcal{J}} + \beta I = \tilde{L}\tilde{U}$ is available, and then generates via the identity $\tilde{\mathcal{K}} + \beta I = \tilde{U}\tilde{L}$ a new tridiagonal matrix $\tilde{\mathcal{K}}$ with the shape (46), except that the entries of its lower diagonal can be signed. When such a diagonal is positive, $\tilde{\mathcal{K}}$ can be symmetrized by conjugation with a positive diagonal matrix $\Pi_{\mathcal{K}}$, leading to a Jacobi matrix $\mathcal{K} = \Pi_{\mathcal{K}}^{-1}\tilde{\mathcal{K}}\Pi_{\mathcal{K}}$.

Note that the LU factorization $\tilde{\mathcal{J}} + \beta I = \tilde{L}\tilde{U}$ exists simultaneously to that of $\mathcal{J} + \beta I$, whose LU factors are $L = \Pi_{\mathcal{J}}^{-1}\tilde{L}\Pi_{\mathcal{J}}$ and $U = \Pi_{\mathcal{J}}^{-1}\tilde{U}\Pi_{\mathcal{J}}$. Therefore, both LU factorizations are possible whenever $\mathcal{J} + \beta I$ is positive definite. We are going

to see that in this case the Darboux transformations of Jacobi matrices can be understood in terms of Cholesky factorizations, similarly to the developed Darboux transformations for CMV matrices. This will make it easier the comparison between Darboux for Jacobi and CMV.

The matrices involved in the above discussion are determined by two real sequences v_n and u_n defined by the lower and main diagonals of the factors \tilde{L} and \tilde{U} , respectively,

$$\tilde{L} = \begin{pmatrix} 1 & & & \\ v_1 & 1 & & \\ & v_2 & 1 & \\ & & \ddots & \ddots \end{pmatrix}, \quad \tilde{U} = \begin{pmatrix} u_1 & 1 & & \\ & u_2 & 1 & \\ & & u_3 & 1 \\ & & & \ddots & \ddots \end{pmatrix}.$$

More precisely, $\tilde{\mathcal{J}}$, $\tilde{\mathcal{K}}$ and their symmetrizing matrices $\Pi_{\mathcal{J}}$, $\Pi_{\mathcal{K}}$ are given by

$$\begin{aligned} \tilde{\mathcal{J}} + \beta I &= \begin{pmatrix} u_1 & 1 & & \\ v_1 u_1 & v_1 + u_2 & 1 & \\ & v_2 u_2 & v_2 + u_3 & 1 \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \quad \tilde{\mathcal{K}} + \beta I = \begin{pmatrix} v_1 + u_1 & 1 & & \\ v_1 u_2 & v_2 + u_2 & 1 & \\ & v_2 u_3 & v_3 + u_3 & 1 \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \\ \Pi_{\mathcal{J}}^2 &= \begin{pmatrix} 1 & & \\ \pi_1 & & \\ & \pi_2 & \\ & & \ddots \end{pmatrix}, \quad \Pi_{\mathcal{K}}^2 = \frac{1}{u_1} \Pi_{\mathcal{J}}^2 U_0, \quad U_0 = \begin{pmatrix} u_1 & & \\ & u_2 & \\ & & u_3 \\ & & & \ddots \end{pmatrix}, \quad \pi_n = \prod_{k=1}^n u_k v_k. \end{aligned} \quad (47)$$

By hypothesis, $\tilde{\mathcal{J}}$ is symmetrizable, i.e. $u_n v_n > 0$ for all n . Then, $\tilde{\mathcal{K}}$ is also symmetrizable iff $v_n u_{n+1} > 0$ for all n , which, bearing in mind the previous hypothesis, means that the sequences u_n and v_n have the same constant sign. This is the case when $\mathcal{J} + \beta I$ is positive definite, which is equivalent to state that $u_n, v_n > 0$ for all n and, thus, also to the positive definiteness of $\mathcal{K} + \beta I$.

It is known that LU and Cholesky factorizations of positive definite matrices are closely related. If $\mathcal{J} + \beta I$ is positive definite, its LU factors provide the Cholesky factorization

$$\mathcal{J} + \beta I = AA^+, \quad A = LU_0^{1/2}, \quad A^+ = U_0^{-1/2}U, \quad (48)$$

where $U_0^{1/2}$ is the positive square root of the matrix U_0 given in (47). Then, using (47) and (48) we get

$$\mathcal{K} + \beta I = \Pi_{\mathcal{K}}^{-1} \tilde{U} \tilde{\mathcal{K}} \Pi_{\mathcal{K}} = \Pi_{\mathcal{K}}^{-1} \Pi_{\mathcal{J}} U L \Pi_{\mathcal{J}}^{-1} \Pi_{\mathcal{K}} = U_0^{-1/2} U L U_0^{1/2} = A^+ A.$$

We conclude that the standard Darboux transformations relating Jacobi matrices \mathcal{J} and \mathcal{K} can be rewritten in the following way: take a real polynomial $\wp(x) = x + \beta$ such that $\wp(\mathcal{J})$ is positive definite. Then, a new Jacobi matrix \mathcal{K} is defined resorting to the Cholesky factorization $\wp(\mathcal{J}) = AA^+$, $A \in \mathcal{T}$, and reversing it as $\wp(\mathcal{K}) = A^+ A$. The above procedure can be directly generalized to an arbitrary real polynomial $\wp(x) = \alpha x + \beta$ of degree one such that $\wp(\mathcal{J})$ is positive definite, leading to a triangular matrix A with the 2-band structure

$$A = \begin{pmatrix} + & & & \\ \otimes & + & & \\ & \otimes & + & \\ & & & \ddots & \ddots \end{pmatrix},$$

whose lower diagonal has non-zero entries with the same sign as α . This defines the Darboux transformation $\mathcal{J} \xrightarrow{\varphi} \mathcal{K}$.

Jacobi matrices encode the three-term recurrence relation of orthogonal polynomials with respect to a positive Borel measure supported on an infinite subset of the real line ('measure on the real line' in short). Analogously to the CMV case, this links Jacobi matrices and positive definite linear functionals in the space of real polynomials. More generally, tridiagonal matrices with non-null entries in the lower and upper diagonals are related to polynomials which are orthogonal with respect to a quasi-definite linear functional not necessarily associated with a positive measure on the real line. An advantage of defining Darboux transformations via LU factorizations is that they apply directly to the general quasi-definite case. Although the Cholesky approach to Darboux transformations previously described has *a priori* the drawback of being applicable only to the positive definite case, Section 7.2 will show how to deal with Darboux transformations relating quasi-definite functionals by using generalized Cholesky factorizations.

Analogously to the CMV case, the Darboux transformations of Jacobi matrices based on Cholesky factorizations can be embedded into a general procedure for the factorization of a real polynomial φ evaluated on a Jacobi matrix \mathcal{J} . A key role is played by the basis $p = (p_0, p_1, p_2, \dots)^t$ of the real vector space $\mathbb{P} = \text{span}\{x^n\}_{n=0}^\infty = \mathbb{R}[x]$ of real polynomials defined up to a constant factor by

$$xp = \mathcal{J}p.$$

These polynomials are orthonormal with respect to a positive Borel measure on the real line, or equivalently, with respect to a positive definite linear functional on \mathbb{P} . Given any other basis $q = (q_0, q_1, q_2, \dots)^t$ of orthonormal polynomials related to a Jacobi matrix \mathcal{K} , the matrix $A \in \mathcal{T}$ and the lower Hessenberg type one B defined by

$$p = Aq, \quad \varphi q = Bp,$$

satisfy

$$\varphi(\mathcal{J}) = AB, \quad \varphi(\mathcal{K}) = BA, \quad \mathcal{J}A = AK, \quad B\mathcal{J} = \mathcal{K}B.$$

The result is valid for any arbitrary real polynomial φ of any degree, and the proof is a direct translation of that of Proposition 2.10. As in the CMV case, the factor A plays the role of a change of basis relating the Jacobi matrices by conjugation since $\mathcal{K} = A^{-1}\mathcal{J}A$ and $\mathcal{J} = A\mathcal{K}A^{-1}$. Nevertheless, in contrast to the CMV case, if φ has degree one, the relations $\varphi(\mathcal{J}) = AB$ and $\varphi(\mathcal{K}) = BA$ are enough to generate one of the matrices \mathcal{J}, \mathcal{K} starting from the other one.

When φ has degree one, $\varphi(\mathcal{J})$ is tridiagonal and B is strictly lower Hessenberg. The Darboux transformation amounts to choosing A as the left factor of the Cholesky factorization of $\varphi(\mathcal{J})$ when it is positive definite, so that A is 2-band with positive entries and $B = A^+$. Furthermore, Theorem 3.2 has an exact analogue for Jacobi matrices whose proof follows from similar arguments. In particular, the transformation $\mathcal{J} \xrightarrow{\varphi} \mathcal{K}$ is equivalent to multiplying by (a positive rescaling of) φ in terms of

orthogonality measures, which is known as a Christoffel transformation on the real line. As a consequence, the Darboux transformations of Jacobi matrices are almost isospectral since they preserve the spectrum up to the elimination of, at most, a single eigenvalue. This is similar to what happens to Darboux for CMV, although these transformations may delete two eigenvalues.

The only spectral change caused by the corresponding inverse transformations for Jacobi matrices is therefore the possible addition of an eigenvalue. The inverse Darboux transformations of Jacobi matrices are the matrix realization of the Geronimus transformations for measures on the real line, which divide a measure by a polynomial φ of degree one and add a Dirac delta at the zero of φ . On the other hand, in the language of Cholesky factorizations, the Darboux transformations with parameters for a Jacobi matrix \mathcal{K} start with a reversed factorization $\varphi(\mathcal{K}) = A^+A$, $A \in \mathcal{S}$, and then generate a new Jacobi matrix \mathcal{J} by either the relation $\varphi(\mathcal{J}) = AA^+$ or equivalently the conjugation $\mathcal{J} = A\mathcal{K}A^{-1}$.

At this point, a major difference appears concerning the comparison between Darboux for CMV and Jacobi matrices: if φ is of degree one, the freedom in the reversed Cholesky factorization $\varphi(\mathcal{K}) = A^+A$ is encoded by a single real parameter, the $(0,0)$ entry of A . This coincides with the number of free parameters in the Geronimus transformations on the real line, which are parametrized by the mass of the Dirac delta at the single zero of φ . Actually, contrary to the CMV case, every solution of the Jacobi version for the Darboux transformations with parameters corresponds to a solution of the corresponding inverse Darboux transformations. In other words, the Darboux transformations with parameters for Jacobi matrices present no spurious solutions, so that these transformations become an exact matrix realization of the Geronimus transformations for measures on the real line. This is in striking contrast with the CMV version of Darboux transformations with parameters.

Due to the absence of spurious solutions, the Jacobi version of Darboux transformations only involves band matrices and their inverses. Therefore, in view of Proposition 2.3, in the Jacobi case the associativity of matrix multiplications holds automatically for all the Darboux relations which exclude such inverses. This will be used in what follows to omit discussing associativity issues when dealing with relations involving only band matrices.

Despite the differences above mentioned, we have seen many similarities between the Jacobi and CMV case which justify the name ‘Darboux’ for the transformations of CMV matrices introduced in this paper. The analogy between the Jacobi and the CMV case can be highlighted by using the Szegő connection between orthogonal polynomials on the real line and on the unit circle. This will also reveal the close relation between the Cholesky factorizations behind the Darboux transformations of Jacobi and CMV matrices. This will be discussed in the following section, which uses the CMV approach to the Szegő connection given by Killip and Nenciu in [50] (see also [67]).

5.1 The Szegő connection

The Szegő connection is based on the map $z \mapsto x = z + z^{-1}$, which transforms the unit circle onto the interval $[-2, 2]$ mapping conjugated points onto the same image. This induces the Szegő projection $\mu \xrightarrow{\text{Sz}} \text{Sz}(\mu)$, a one-to-one correspondence between symmetric measures μ on the unit circle (i.e., measures which are invariant under conjugation $z \rightarrow \bar{z}$) and measures $\text{Sz}(\mu)$ on $[-2, 2]$, which is given by

$$\int P(x) d(\text{Sz}(\mu))(x) = \int P(z + z^{-1}) d\mu(z), \quad \forall P \in \mathbb{P}.$$

The symmetry of μ means that it has orthonormal polynomials φ_n with real coefficients, i.e. the corresponding Schur parameters a_n are real. Then, the relations

$$\begin{aligned} p_n^{(e)}(x) &= \frac{z^{-n}(\varphi_{2n}^*(z) + \varphi_{2n}(z))}{\sqrt{2(1 + a_{2n})}}, \\ p_{n-1}^{(o)}(x) &= \frac{z^{-n}(\varphi_{2n}^*(z) - \varphi_{2n}(z))}{\sqrt{2(1 - a_{2n})(z^{-1} - z)}}, \end{aligned} \quad x = z + z^{-1}, \quad (49)$$

provide two sequences of polynomials in the variable $x = z + z^{-1}$,

$$p^{(k)} = (p_0^{(k)}, p_1^{(k)}, p_2^{(k)}, \dots)^t, \quad k = e, o.$$

We use the convention $a_0 = 1$ so that $p_0^{(e)}(x) = \varphi_0(z)$. The sequences $p^{(e)}$ and $p^{(o)}$ are orthonormal with respect to the measures on $[-2, 2]$ given respectively by

$$\mu_e = \text{Sz}(\mu), \quad d\mu_o(x) = (4 - x^2) d\mu_e(x),$$

Due to the orthonormality with respect to a measure on the real line, there exist Jacobi matrices $\mathcal{J}_e, \mathcal{J}_o$, such that

$$xp^{(k)} = \mathcal{J}_k p^{(k)}, \quad k = e, o. \quad (50)$$

The relation between the measures μ_e and μ_o implies that \mathcal{J}_e and \mathcal{J}_o are connected by a Darboux transformation $\mathcal{J}_e \xrightarrow{4-x^2} \mathcal{J}_o$ generated by a polynomial of degree two (see Section 7.1 for the generalization of Darboux transformations to polynomials of higher degree).

For convenience, we will work with the matrix functions

$$\mathbf{p}_n = \begin{pmatrix} p_n & 0 \\ 0 & \tilde{p}_n \end{pmatrix}, \quad \begin{aligned} p_n(z) &= p_n^{(e)}(x), \\ \tilde{p}_n(z) &= (z^{-1} - z)p_n^{(o)}(x), \end{aligned} \quad x = z + z^{-1}, \quad (51)$$

which should be considered as functions of z instead of x due to the presence of the factor $z - z^{-1}$ in \tilde{p}_n . Bearing in mind (51), the recurrence relations (50) can be rewritten in terms of $p = (p_0, p_1, p_2, \dots)^t$ and $\tilde{p} = (\tilde{p}_0, \tilde{p}_1, \tilde{p}_2, \dots)^t$ as

$$xp(z) = \mathcal{J}_e p(z), \quad x\tilde{p}(z) = \mathcal{J}_o \tilde{p}(z), \quad x = z + z^{-1}.$$

Equivalently, using the notation $\mathbf{p} = (\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \dots)^t$,

$$x\mathbf{p}(z) = \mathcal{J}\mathbf{p}(z), \quad \mathcal{J} = \mathcal{J}_e \oplus \mathcal{J}_o, \quad x = z + z^{-1}, \quad (52)$$

with \mathcal{J}_e (\mathcal{J}_o) acting on even (odd) indices. More explicitly, \mathcal{J} is the 2×2 -block Jacobi matrix with entries

$$\mathcal{J}_{2i,2j} = (\mathcal{J}_e)_{i,j}, \quad \mathcal{J}_{2i+1,2j+1} = (\mathcal{J}_o)_{i,j}, \quad \mathcal{J}_{2i,2j+1} = \mathcal{J}_{2i+1,2j} = 0,$$

and can be understood as the block Jacobi matrix associated with the diagonal matrix of measures

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_e & 0 \\ 0 & \mu_o \end{pmatrix}.$$

On the other hand, (24) allows us to rewrite p_n and \tilde{p}_n in terms of the orthonormal Laurent polynomials χ_n associated with μ ,

$$p_n = \frac{\chi_{2n} + \chi_{2n*}}{\sqrt{2(1+a_{2n})}}, \quad \tilde{p}_{n-1} = \frac{\chi_{2n} - \chi_{2n*}}{\sqrt{2(1-a_{2n})}}. \quad (53)$$

If \mathcal{C} is the CMV matrix related to μ , we denote $\boldsymbol{\mu} = \mathbf{S}\mathbf{z}(\mu)$, $\mathbf{p} = \mathbf{S}\mathbf{z}(\chi)$, $\mathcal{J} = \mathbf{S}\mathbf{z}(\mathcal{C})$, referring to them as the matrix Szegő projections of μ , χ , \mathcal{C} , respectively, in contrast to the scalar Szegő projections $\mu_e = \text{Sz}(\mu)$, $p^{(e)} = \text{Sz}(\mu)$, $\mathcal{J}_e = \text{Sz}(\mathcal{C})$.

An explicit expression of the matrix Szegő projection $\mathcal{J} = \mathbf{S}\mathbf{z}(\mathcal{C})$ follows by inserting (25) into (53), which yields

$$p_n = \frac{\rho_{2n}\chi_{2n-1} + (1+a_{2n})\chi_{2n}}{\sqrt{2(1+a_{2n})}} = \frac{1}{\sqrt{2}}(\sqrt{1-a_{2n}}\chi_{2n-1} + \sqrt{1+a_{2n}}\chi_{2n}),$$

$$\tilde{p}_{n-1} = \frac{\rho_{2n}\chi_{2n-1} - (1-a_{2n})\chi_{2n}}{\sqrt{2(1-a_{2n})}} = \frac{1}{\sqrt{2}}(-\sqrt{1+a_{2n}}\chi_{2n-1} + \sqrt{1-a_{2n}}\chi_{2n}),$$

or, in matrix form,

$$\begin{pmatrix} \tilde{p}_{n-1} \\ p_n \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -\sqrt{1+a_{2n}} & \sqrt{1-a_{2n}} \\ \sqrt{1-a_{2n}} & \sqrt{1+a_{2n}} \end{pmatrix} \begin{pmatrix} \chi_{2n-1} \\ \chi_{2n} \end{pmatrix}.$$

The above relation can be compactly written as

$$\mathbf{p} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{S}\boldsymbol{\chi}, \quad \mathbf{S} = \begin{pmatrix} 1 & & & \\ & s_2 & & \\ & & s_4 & \\ & & & \ddots \end{pmatrix}, \quad S_n = \frac{1}{\sqrt{2}} \begin{pmatrix} -\sqrt{1+a_n} & \sqrt{1-a_n} \\ \sqrt{1-a_n} & \sqrt{1+a_n} \end{pmatrix}, \quad (54)$$

in terms of a block diagonal symmetric unitary matrix \mathbf{S} . Combining (54) with (52) and (6) we get

$$\mathcal{J}\mathbf{S}\boldsymbol{\chi} = \mathcal{J}\mathbf{p} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (z + z^{-1})\mathbf{p} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (z + z^{-1})\mathbf{S}\boldsymbol{\chi} = \mathbf{S}(\mathcal{C} + \mathcal{C}^+)\boldsymbol{\chi},$$

which, bearing in mind that $\mathbf{S}^2 = I$, leads to the following explicit expression for the matrix Szegő projection of a CMV matrix \mathcal{C} ,

$$\mathbf{Sz}(\mathcal{C}) = \mathcal{J} = \mathbf{S}(\mathcal{C} + \mathcal{C}^+)\mathbf{S}. \quad (55)$$

The Szegő connection is restricted to symmetric measures on the unit circle, i.e. to real CMV matrices because symmetric measures are characterized by having real Schur parameters. Therefore, in order to relate Darboux transformations on the unit circle and the real line through the Szegő connection we should consider a Darboux transformation $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ between two real CMV matrices \mathcal{C} , \mathcal{D} . This means that $d\nu = \ell d\mu$ for certain symmetric measures μ , ν on the unit circle related to \mathcal{C} , \mathcal{D} , respectively. If $\ell(z) = \alpha z + \beta + \bar{\alpha}z^{-1}$, the symmetry requirement for the measures implies that $\ell(z^{-1}) = \ell(z)$, thus $\alpha \in \mathbb{R}$ and

$$d\nu_e = \wp d\mu_e, \quad \wp(x) = \alpha x + \beta, \quad \mu_e = \mathbf{Sz}(d\mu), \quad \nu_e = \mathbf{Sz}(d\nu).$$

A similar relation is obviously valid for the measures $d\mu_o(x) = (4 - x^2) d\mu_e(x)$ and $d\nu_o(x) = (4 - x^2) d\nu_e(x)$. This implies that the Jacobi matrices $\mathcal{J}_e = \mathbf{Sz}(\mathcal{C})$ and $\mathcal{K}_e = \mathbf{Sz}(\mathcal{D})$ are related by a Darboux transformation $\mathcal{J}_e \xrightarrow{\wp} \mathcal{K}_e$, and analogously for the Jacobi matrices \mathcal{J}_o , \mathcal{K}_o associated with μ_o , ν_o , respectively. Since $\mathcal{J} = \mathbf{Sz}(\mathcal{C}) = \mathcal{J}_e \oplus \mathcal{J}_o$ and $\mathcal{K} = \mathbf{Sz}(\mathcal{D}) = \mathcal{K}_e \oplus \mathcal{K}_o$, we have the following Cholesky and reversed Cholesky factorizations,

$$\begin{aligned} \ell(\mathcal{C}) &= \mathbf{A}\mathbf{A}^+, & \ell(\mathcal{D}) &= \mathbf{A}^+\mathbf{A}, \\ \wp(\mathcal{J}_k) &= A_k A_k^+, & \wp(\mathcal{K}_k) &= A_k^+ A_k, & k &= e, o, \\ \wp(\mathcal{J}) &= \mathbf{A}\mathbf{A}^+, & \wp(\mathcal{K}) &= \mathbf{A}^+\mathbf{A}, & \mathbf{A} &= A_e \oplus A_o, \end{aligned}$$

where the direct sum $\mathbf{A} = A_e \oplus A_o$ is such that A_e (A_o) acts on even (odd) indices. Thus, \mathbf{A} has the structure

$$\mathbf{A} = \begin{pmatrix} + & & & & \\ 0 & + & & & \\ \otimes & 0 & + & & \\ & \otimes & 0 & + & \\ & & \ddots & \ddots & \ddots \end{pmatrix},$$

with the lower diagonal having constant sign equal to that of α . This shows that Darboux transformations for real CMV matrices, as defined in this paper, are mapped exactly into Darboux transformations for Jacobi matrices via the Szegő connection.

This is in agreement with [29], which shows that, in the case of symmetric measures, the linear spectral transformations associated with the Christoffel and Geronimus transformations on the unit circle are projected by the Szegő mapping onto similar transformations on the real line. However, we have seen that there exist Darboux transformations of CMV matrices which do not correspond to Christoffel or Geronimus on the unit circle, namely, those associated with a Laurent polynomial $\ell(z) = \alpha z + \beta + \alpha z^{-1}$, $\alpha, \beta \in \mathbb{R}$, whose zeros lie on the unit circle, i.e. $|\beta| \leq 2|\alpha|$. In

other words, there exist Christoffel and Geronimus transformations on the real line which are not the Szegő projection of a Christoffel or Geronimus transformation on the unit circle, but correspond via the Szegő mapping to general Laurent polynomial modifications of degree one on the unit circle.

The close relationship between Darboux for CMV and Jacobi is also highlighted by an explicit and very simple relation between the Darboux factor A and its matrix Szegő projection $\mathbf{A} = \mathbf{S}\mathbf{z}(A)$. Let us denote $\mathbf{p} = \mathbf{S}\mathbf{z}(\chi)$ and $\mathbf{q} = \mathbf{S}\mathbf{z}(\omega)$, where χ and ω are the orthonormal Laurent polynomials with respect to μ and ν , respectively. Together with (54) we have the analogous relation

$$\mathbf{q} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{T}\omega, \quad \mathbf{T} = \begin{pmatrix} 1 & & & \\ & T_2 & & \\ & & T_4 & \\ & & & \ddots \end{pmatrix}, \quad T_n = \frac{1}{\sqrt{2}} \begin{pmatrix} -\sqrt{1+b_n} & \sqrt{1-b_n} \\ \sqrt{1-b_n} & \sqrt{1+b_n} \end{pmatrix}, \quad (56)$$

where b_n are the Schur parameters of \mathcal{D} . On the other hand, from Theorem 3.2 and a similar result for Jacobi matrices we know that

$$\chi = A\omega, \quad p^{(e)} = A_e q^{(e)}, \quad p^{(o)} = A_o q^{(o)},$$

the last two identities being equivalent to

$$\mathbf{p} = \mathbf{A}\mathbf{q}.$$

Combining these relations with (54) and (56) we obtain

$$\mathbf{S}A\omega = \mathbf{S}\chi = \mathbf{p} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{A}\mathbf{q} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{A}\mathbf{T}\omega,$$

so that $\mathbf{S}A = \mathbf{A}\mathbf{T}$. Since $\mathbf{S}^2 = \mathbf{T}^2 = I$ we can rewrite this relation as

$$A = \mathbf{S}\mathbf{A}\mathbf{T}, \quad \mathbf{A} = \mathbf{S}A\mathbf{T}, \quad (57)$$

which is the referred connection between the Darboux factors for CMV and Jacobi matrices. The identification of \mathbf{A} given in (57) as the Darboux factor for $\mathcal{J} \xrightarrow{\wp} \mathcal{K}$ is consistent with the relation (55), which implies that

$$\wp(\mathcal{J}) = \mathbf{S}\ell(\mathcal{C})\mathbf{S} = \mathbf{S}A\mathbf{A}^+\mathbf{S} = \mathbf{A}\mathbf{A}^+, \quad \wp(\mathcal{K}) = \mathbf{T}\ell(\mathcal{D})\mathbf{T} = \mathbf{T}\mathbf{A}^+\mathbf{A}\mathbf{T} = \mathbf{A}^+\mathbf{A}.$$

Note that, analogously to the relation $\ell\omega = A^+\chi$, we have the identity $\wp\mathbf{q} = \mathbf{A}^+\mathbf{p}$ which provides the orthonormal polynomials $q^{(k)}$ related to the transformed Jacobi matrix \mathcal{K}_k in terms of those $p^{(k)}$ associated with \mathcal{J}_k .

We will illustrate the Szegő connection between Darboux for Jacobi and CMV with an explicit example.

Example 5.1. *The Szegő connection and the Darboux transformation for*

$$\ell(z) = (z+a)(z^{-1}+a) = a(z+z^{-1}) + 1 + a^2, \quad a \in (-1, 1),$$

and \mathcal{C} with Schur parameters $a_1 = a$ and $a_n = 0$ for $n \geq 2$,

$$\mathcal{C} = \begin{pmatrix} -a & \rho & & & & \\ 0 & 0 & 0 & 1 & & \\ \rho & a & 0 & 0 & & \\ & 0 & 0 & 0 & 1 & \\ & 1 & 0 & 0 & 0 & \\ & & 0 & 0 & 0 & 1 \\ & & 1 & 0 & 0 & 0 \\ & & & \dots & \dots & \dots \end{pmatrix}, \quad \rho = \sqrt{1 - a^2}.$$

It is known that \mathcal{C} is associated with the measure

$$d\mu(e^{i\theta}) = \frac{1}{|z + a|^2} \frac{d\theta}{2\pi} = \frac{1}{\ell(z)} \frac{d\theta}{2\pi},$$

whose orthonormal polynomials are the Bernstein-Szegő ones

$$\varphi_n(z) = \begin{cases} \rho, & n = 0, \\ z^{n-1}(z + a), & n \geq 1. \end{cases}$$

Therefore, the corresponding orthonormal zig-zag basis is given by

$$\chi_{2n-1}(z) = z^{n-1}(z + a), \quad \chi_{2n}(z) = \begin{cases} \rho, & n = 0, \\ z^{-n}(1 + az), & n \geq 1. \end{cases}$$

The Darboux transformation $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ is associated with the Christoffel transformation $\mu \xrightarrow{\ell} \nu$, where $d\nu(e^{i\theta}) = d\theta/2\pi$ is the Lebesgue measure. Hence, $\mathcal{D} = \mathcal{S}$ is the shift matrix given in (8) corresponding to the Schur parameters $b_n = 0$, and the related orthonormal zig-zag basis is

$$\omega_{2n-1}(z) = z^n, \quad \omega_{2n}(z) = z^{-n}.$$

The matrix A of the change of basis

$$\chi = A\omega, \quad A = \begin{pmatrix} \rho & & & & \\ a & 1 & & & \\ a & 0 & 1 & & \\ a & 0 & 1 & & \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

is the Darboux factor for the Cholesky factorizations $\ell(\mathcal{C}) = AA^+$ and $\ell(\mathcal{D}) = A^+A$.

Concerning the Szegő connection, in this case the matrices \mathbf{S} and \mathbf{T} coincide because $a_{2n} = b_{2n} = 0$, thus

$$S_n = T_n = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

The matrix Szegő projection $\mathbf{A} = \mathbf{S}\mathbf{A}\mathbf{T}$ of A

$$\mathbf{A} = \begin{pmatrix} \rho & & & & \\ 0 & 1 & & & \\ \sqrt{2}a & 0 & 1 & & \\ a & 0 & 1 & & \\ a & 0 & 1 & & \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} = A_e \oplus A_o, \quad A_e = \begin{pmatrix} \rho & & & & \\ \sqrt{2}a & 1 & & & \\ a & 1 & & & \\ a & 1 & & & \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}, \quad A_o = \begin{pmatrix} 1 & & & & \\ a & 1 & & & \\ a & 1 & & & \\ a & 1 & & & \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

An important ingredient in these interrelations is the symmetrization process for orthogonal polynomials on the unit circle [47, 56], which associates with any measure μ on the unit circle the only one $\hat{\mu}$ which is invariant under $z \rightarrow -z$ and satisfies

$$\int f(z) d\mu(z) = \int f(z^2) d\hat{\mu}(z), \quad \forall f \in \Lambda.$$

We will express this relation as $\mu \xrightarrow{z \mapsto z^2} \hat{\mu}$. The corresponding orthonormal polynomials, φ_n and $\hat{\varphi}_n$, are related by

$$\hat{\varphi}_{2n}(z) = \varphi_n(z^2), \quad \hat{\varphi}_{2n+1}(z) = z\varphi_n(z^2), \quad (58)$$

so that the induced transformation of CMV matrices $\mathcal{C} \xrightarrow{z \mapsto z^2} \hat{\mathcal{C}}$ amounts to the transformation of Schur parameters $(a_n)_{n \geq 1} \xrightarrow{z \mapsto z^2} (\hat{a}_n)_{n \geq 1} = (0, a_1, 0, a_2, 0, a_3, \dots)$.

Since $z \mapsto z^2$ reads as $x \mapsto x^2 - 2$ if $x = z + z^{-1}$, the combination of the Szegő projection and the above symmetrization process maps the measure $\mu_e = \text{Sz}(\mu)$ onto the only one $\hat{\mu}_e = \text{Sz}(\hat{\mu})$ on $[-2, 2]$ which is symmetric under $x \rightarrow -x$ and satisfies

$$\int P(x) d\mu_e(x) = \int P(x^2 - 2) d\hat{\mu}_e(x), \quad \forall P \in \mathbb{P}.$$

We will write $\mu_e \xrightarrow{x \mapsto x^2 - 2} \hat{\mu}_e$. The orthonormal polynomials $\hat{p}_n^{(e)}$ of $\hat{\mu}_e$ are given by

$$\hat{p}_{2n}^{(e)}(x) = p_n^{(e)}(x^2 - 2), \quad \hat{p}_{2n+1}^{(e)}(x) = x q_n^{(e)}(x^2 - 2), \quad (59)$$

where $p_n^{(e)}$ and $q_n^{(e)}$ are orthonormal with respect to μ_e and its Christoffel transform $(x+2)d\mu_e(x)$, respectively [10]. The Jacobi matrices \mathcal{J}_e and \mathcal{K}_e of $p_n^{(e)}$ and $q_n^{(e)}$ are related by the Darboux transformation $\mathcal{J}_e \xrightarrow{x+2} \mathcal{K}_e$. Since $\mathcal{J}_e = \text{Sz}(\mathcal{C})$, we conclude that $\mathcal{K}_e = \text{Sz}(\mathcal{D})$ with \mathcal{D} the CMV matrix given by the Darboux transformation $\mathcal{C} \xrightarrow{z+z^{-1}+2} \mathcal{D}$. The relations (59) imply that the Jacobi matrix $\hat{\mathcal{J}}_e$ of $\hat{p}_n^{(e)}$ satisfies

$$\hat{\mathcal{J}}_e^2 - 2I = \mathcal{J}_e \oplus \mathcal{K}_e,$$

where \mathcal{J}_e and \mathcal{K}_e act on even and odd indices respectively. We will express the above connections by writing $\mathcal{K}_e \xleftarrow{x+2} \mathcal{J}_e \xrightarrow{x \mapsto x^2 - 2} \hat{\mathcal{J}}_e$, the double line in the right arrow indicating that both, \mathcal{J}_e and \mathcal{K}_e , are involved in $\hat{\mathcal{J}}_e$.

The DVZ connection starts from a known factorization of CMV matrices into block-diagonal unitary factors [11],

$$\mathcal{C} = \mathcal{M}\mathcal{L}, \quad \mathcal{L} = \begin{pmatrix} \Theta_1 & & & \\ & \Theta_3 & & \\ & & \Theta_5 & \\ & & & \ddots \end{pmatrix}, \quad \mathcal{M} = \begin{pmatrix} a_0 & & & \\ & \Theta_2 & & \\ & & \Theta_4 & \\ & & & \ddots \end{pmatrix}, \quad \Theta_n = \begin{pmatrix} -a_n & \rho_n \\ \rho_n & \bar{a}_n \end{pmatrix}.$$

Here a_n are the Schur parameters of \mathcal{C} and, as previously, we use the convention $a_0 = 1$. When these parameters are real, a Jacobi matrix \mathcal{K} is defined by

$$\mathcal{K} = \mathcal{L} + \mathcal{M} = \begin{pmatrix} a_0 - a_1 & \rho_1 & & & \\ \rho_1 & a_1 - a_2 & \rho_2 & & \\ & \rho_2 & a_2 - a_3 & \rho_3 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix}. \quad (60)$$

This establishes the DVZ connection $\mathcal{C} \xrightarrow{\text{DVZ}} \mathcal{K}$. Derevyagin, Vinet and Zhedanov identify in [22] the measure and orthonormal polynomials of \mathcal{K} in terms of those of \mathcal{C} . This identification appears surprisingly when combining a Christoffel transformation on the real line with a connection between measures on the unit circle and on the real line due to Delsarte and Genin [18]. A more direct approach to this problem can be found in [10], which also uncovers the role of the symmetrization process and the Szegő connection in the DVZ connection. If μ and φ_n are the measure and orthonormal polynomials related to \mathcal{C} , then

$$(x+2)d\hat{\mu}_e(x), \quad q_n(x) = \frac{z^{-n}(\varphi_n^*(z^2) + z\varphi_n(z^2))}{\sqrt{2}(1+z)}, \quad x = z + z^{-1},$$

are the measure and orthonormal polynomials associated with \mathcal{K} . Here, $\hat{\mu}_e = \text{Sz}(\hat{\mu})$ is the Szegő projection of the symmetrization $\hat{\mu}$ of μ . It turns out that $\mathcal{K} = \text{Sz}(\hat{\mathcal{D}})$ is the Szegő projection of the CMV matrix $\hat{\mathcal{D}}$ arising from the Darboux transformation $\hat{\mathcal{C}} \xrightarrow{z+z^{-1}+2} \hat{\mathcal{D}}$, where $\hat{\mathcal{C}}$ comes from the symmetrization $\mathcal{C} \xrightarrow{z \mapsto z^2} \hat{\mathcal{C}}$.

These results can be summarized in the following commutative diagram, where Darboux transformations play a prominent role. It shows explicitly that DVZ follows from a mix of symmetrization, Szegő and Darboux.

$$\begin{array}{ccccccc}
 \mathcal{D} & \xleftarrow{z+z^{-1}+2} & \mathcal{C} & \xrightarrow{z \mapsto z^2} & \hat{\mathcal{C}} & \xrightarrow{z+z^{-1}+2} & \hat{\mathcal{D}} \\
 \downarrow \text{Sz} & & \downarrow \text{Sz} & & \downarrow \text{Sz} & \searrow \text{DVZ} & \downarrow \text{Sz} \\
 \mathcal{K}_e & \xleftarrow{x+2} & \mathcal{J}_e & \xrightarrow{x \mapsto x^2-2} & \hat{\mathcal{J}}_e & \xrightarrow{x+2} & \mathcal{K} = \mathcal{L} + \mathcal{M}
 \end{array}$$

6 Darboux versus QR

Prior to Cholesky factorizations, other matrix factorizations have been related to Laurent polynomial modifications of measures on the unit circle. In the same work [74] that introduces systematically for the first time the CMV matrices, D. Watkins links Christoffel transformations on the unit circle to QR factorizations of shifted CMV matrices. This can be seen as the unitary analogue of a similar connection between Jacobi matrices and polynomial modifications of measures on the real line first established by Kautsky and Golub [49] (see also [6, 7, 32, 74]). A version of this result for Hessenberg matrices related to orthogonal polynomials on the unit circle appears in [17, 28, 43]. We will describe the QR interpretation of Christoffel transformations on the unit circle based on CMV matrices to compare with the Darboux transformations developed in the present paper.

Given a measure μ on the unit circle, consider a general Christoffel transformation $\mu \xrightarrow{\ell} \nu$, i.e.

$$\ell = q_*q, \quad q(z) = z - \zeta, \quad |\zeta| < 1. \quad (61)$$

Let \mathcal{C} , \mathcal{D} be the CMV matrices related to μ , ν , respectively, and χ , ω the corresponding orthonormal zig-zag bases. We know that $\chi = A\omega$ with A the 3-band

lower triangular factor appearing in the Darboux transformation $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$.

Since $d\nu(z) = |z - \zeta|^2 d\mu(z)$, we have that $(1 - |\zeta|)^2 \mu \leq \nu \leq (1 + |\zeta|)^2 \mu$. Thus, the sets L_μ^2 and L_ν^2 of square integrable functions are identical as topological spaces, although they have different Hilbert structures given by the inner products

$$\langle f, g \rangle_\mu = \int f \bar{g} d\mu, \quad \langle f, g \rangle_\nu = \int f \bar{g} d\nu.$$

In consequence, the following mappings between Hilbert spaces define bounded linear operators with a bounded inverse,

$$\begin{array}{llll} \mathbf{A}: L_\mu^2 \rightarrow L_\nu^2 & \mathbf{U}: L_\nu^2 \rightarrow L_\mu^2 & \mathbf{UA}: L_\mu^2 \rightarrow L_\mu^2 & \mathbf{AU}: L_\nu^2 \rightarrow L_\nu^2 \\ f \mapsto f & f \mapsto q_* f & f \mapsto q_* f & f \mapsto q_* f \end{array}$$

Indeed, \mathbf{U} is unitary because $\langle \mathbf{U}f, \mathbf{U}g \rangle_\mu = \langle f, g \rangle_\nu$ for every $f, g \in L_\nu^2$ and $\text{ran}(\mathbf{U}) = L_\mu^2$.

To generate matrix representations of these operators, we will choose the zig-zag bases χ and ω , respectively, as orthonormal bases for the Hilbert spaces L_μ^2 and L_ν^2 connected by such operators. Since $(\mathbf{A}\chi_0, \mathbf{A}\chi_1, \dots)^t = \chi = A\omega$, the corresponding matrix for the operator \mathbf{A} is simply the Darboux factor A , so that \mathbf{A} reads in coordinates as $X \mapsto Y = XA$. Here X and Y stand for row vectors of coordinates with respect to χ and ω respectively, i.e. $f = X\chi$ and $\mathbf{A}f = f = Y\omega$. Analogously, $q_*(\mathcal{C}) = (\mathcal{C} - \zeta I)^+$ and $q_*(\mathcal{D}) = (\mathcal{D} - \zeta I)^+$ are the matrices of the operators \mathbf{UA} and \mathbf{AU} respectively.

Contrary to the previous discussions, in this context we do not need to worry about the associativity of matrix products because it is guaranteed by the associativity of compositions of operators in Hilbert spaces. Likewise, A has a unique bounded inverse, that one representing the bounded operator \mathbf{A}^{-1} in the bases ω and χ . Hence, the linear operator $X \rightarrow XA$ has a trivial kernel in the space \mathfrak{H} of square-summable sequences. This result is not in contradiction with Proposition 2.11, which refers to $\ker_R(B) = \ker_R(A^+) = \ker_L(A)^+$, since this formal kernel can have no non-trivial square-summable vectors.

Let U be the matrix of the unitary operator \mathbf{U} , i.e. the unitary matrix defined by the relation

$$q_*\omega = U\chi.$$

Since \mathbf{UA} and \mathbf{AU} are obtained by composition of \mathbf{A} and \mathbf{U} , their matrices are products of A and U ,

$$\begin{aligned} (\mathcal{C} - \zeta I)^+ \chi = q_* \chi = q_* A \omega = AU \omega &\Rightarrow (\mathcal{C} - \zeta I)^+ = AU, \\ (\mathcal{D} - \zeta I)^+ \omega = q_* \omega = U \chi = UA \omega &\Rightarrow (\mathcal{D} - \zeta I)^+ = UA. \end{aligned}$$

Therefore,

$$q(\mathcal{C}) = \mathcal{C} - \zeta I = QR, \quad q(\mathcal{D}) = \mathcal{D} - \zeta I = RQ, \quad Q = U^+, \quad R = A^+,$$

with Q unitary and R upper triangular. In other words, $q(\mathcal{C})$ and $q(\mathcal{D})$ are connected by a QR factorization and a reversed one, much in the same way as $\ell(\mathcal{C}) = AA^+$ and

$\ell(\mathcal{D}) = A^+A$ are connected by Cholesky factorizations. Moreover, these two kinds of factorizations are closely related because R is the adjoint of the Darboux factor A . Actually, the Cholesky factorizations are a direct consequence of the QR ones because, due to the unitarity of Q ,

$$\begin{aligned} q(\mathcal{C}) = QR &\Rightarrow \ell(\mathcal{C}) = q_*(\mathcal{C})q(\mathcal{C}) = R^+R, \\ q(\mathcal{D}) = RQ &\Rightarrow \ell(\mathcal{C}) = q(\mathcal{D})q_*(\mathcal{D}) = RR^+. \end{aligned}$$

Like the upper triangular factor R , which is 3-band, the unitary factor Q has a remarkable structure. Bearing in mind that $R = A^+$ has a bounded inverse, we find that $Q = (\mathcal{C} - \zeta I)R^{-1}$. This implies that Q has an upper CMV structure, i.e. it is upper Hessenberg type with the shape

$$Q = \begin{pmatrix} * & * & * & * & * & * & \dots \\ * & * & * & * & * & * & \dots \\ * & * & * & * & * & * & \dots \\ & & * & * & * & * & \dots \\ & & & * & * & * & \dots \\ & & & & * & * & \dots \\ & & & & & * & \dots \\ & & & & & & \dots \end{pmatrix}. \quad (62)$$

So far, we have described a QR interpretation of Christoffel transformations on the unit circle based on the use of CMV matrices. This result was first obtained by D.Watkins in [74] using a different approach. Indeed, he proves that the transformations $d\mu(x) \rightarrow |x - \beta|^2 d\mu(x)$, $\beta \in \mathbb{R}$, on the real line have a similar QR interpretation in terms of Jacobi matrices (see also [6, 7, 49]). Therefore, QR suggests that the unit circle version of such transformations on the real line should be the Christoffel transformations on the unit circle. On the contrary, the Darboux transformations developed in the present paper highlight the role of the Christoffel transformations on the unit circle as the counterpart of the Christoffel transformations $d\mu(x) \rightarrow (x - \beta) d\mu(x)$ on the real line.

As in the QR case, the Darboux transformation associated with a Christoffel transformation $\mu \xrightarrow{\ell} \nu$ can be interpreted in terms of factorizations of operators in Hilbert spaces. To show this, let us introduce the following bounded linear operators with a bounded inverse

$$\begin{array}{lll} \mathbf{B}: L_\nu^2 \rightarrow L_\mu^2 & \mathbf{BA}: L_\mu^2 \rightarrow L_\mu^2 & \mathbf{AB}: L_\nu^2 \rightarrow L_\nu^2 \\ f \mapsto \ell f & f \mapsto \ell f & f \mapsto \ell f \end{array}$$

The equality $\langle \mathbf{B}f, g \rangle_\mu = \langle f, \mathbf{A}g \rangle_\nu$ for every $f \in L_\nu^2$ and $g \in L_\mu^2$ identifies $\mathbf{B} = \mathbf{A}^+$ as the adjoint operator of \mathbf{A} . Therefore, A^+ is the matrix of \mathbf{B} with respect to the bases χ of L_μ^2 and ω of L_ν^2 , so that we recover the identity $\ell\omega = (\mathbf{B}\omega_0, \mathbf{B}\omega_1, \dots)^t = A^+\chi$. The corresponding matrices of \mathbf{BA} and \mathbf{AB} are $\ell(\mathcal{C})$ and $\ell(\mathcal{D})$, respectively. The fact that these two last operators are compositions of \mathbf{A} and \mathbf{B} translates into the Cholesky factorizations related to the Darboux transformation $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$. The relation between the QR and Cholesky factorizations is simply the matrix realization of the operator identities $\mathbf{BA} = \mathbf{A}^+\mathbf{U}^+\mathbf{U}\mathbf{A}$ and $\mathbf{AB} = \mathbf{A}\mathbf{U}\mathbf{U}^+\mathbf{A}^+$, which use the fact that $\mathbf{U}^+ = \mathbf{U}^{-1}$ due to the unitarity of \mathbf{U} .

From the operator point of view, the main difference between QR and Darboux is that QR looks for the unitarity of one of the factors, U , requiring no symmetry for the factorized operators, UA and AU , while Darboux needs the self-adjointness of the factorized operators, BA and AB .

We have seen that the QR interpretation of Christoffel transformations on the unit circle yields also the Darboux transformations of the corresponding CMV matrices. Nevertheless, QR does not provide all the richness given by Darboux, which allows more general transformations of measures $\mu \xrightarrow{\ell} \nu$, not necessarily with the form (61) corresponding to ℓ with zeros outside of the unit circle. In particular, QR can never deal with the case of ℓ having two different zeros on the unit circle, a situation covered by Darboux as long as ℓ is non-negative on the support of μ . As for the case of a double zero on the unit circle, i.e. (61) with $|\zeta| = 1$, it can always be tackled by Darboux but not by QR. When the double zero does not lie in the support of μ the situation is as good as for Christoffel transformations concerning QR. Otherwise the operator U is certainly isometric but not necessarily unitary because $\text{ran}(U) = L_\mu^2$ iff $1/q \in L_\nu^2$.

Let us clarify from the operator point of view the situation which distinguishes Darboux from QR, i.e. the case of ℓ having its zeros on the unit circle. Suppose ℓ non-negative on the support of μ , so that $\ell(\mathcal{C})$ is positive definite and Darboux can be implemented. Then, $d\nu = \ell d\mu$ is such that any function in L_μ^2 lies also in L_ν^2 , but the converse is not necessarily true. On the other hand, the supports of μ and ν may differ only in the zeros of ℓ , where μ may have mass points but ν not. Hence, ℓL_ν^2 makes sense as a subset of L_μ^2 . All this ensures that A and B are well defined bounded operators which are adjoint of each other, but now the first of these factors may have a non-trivial kernel. It is not difficult to see that $\ker(B) = \{0\}$, while $\ker(A)$ is spanned by the characteristic functions of the zeros of the ℓ , which is non-trivial iff μ has a mass point at some of these zeros. From general results (see for instance [75, Theorem 5.39]) we know that $BA = A^+A$ and $AB = AA^+$ are unitarily equivalent when restricted to the orthogonal complement of $\ker(A^+A) = \ker(A)$ and $\ker(AA^+) = \ker(A^+) = \{0\}$, respectively. We conclude that BA and AB are isospectral unless μ has a mass point at a zero of ℓ , in which case the spectrum of BA is obtained by adding the null eigenvalue to the spectrum of AB . Bearing in mind that, using the notation (20), $BA = \ell(U_\mu)$ and $AB = \ell(U_\nu)$, this agrees with the fact that the spectrum of U_μ differs from that of U_ν only in the mass points of μ located at the zeros of ℓ . Note that $\chi(z)^+$ are the coordinates of the characteristic function of $\{z\}$ with respect to the orthonormal basis χ of L_μ^2 . Therefore, the previous comments show that the square-summable vectors of the formal kernel $\ker_R(B) = \ker_R(A^+) = \ker_L(A)^+$ are spanned by the evaluations of χ at the zeros of ℓ .

7 Further extensions

The Darboux transformations of CMV matrices introduced in this paper admit a couple of generalizations of interest. One of them has to do with the degree of the Hermitian Laurent polynomial involved in the transformation. The results of the previous sections carry over directly to polynomial modifications of higher degree. Indeed, announcing this generalization, some of the results previously obtained were proved for Laurent polynomials of arbitrary degree (see Section 2.2).

On the other hand, as we pointed out previously, Darboux transformations of Jacobi matrices can deal also with tridiagonal matrices whose lower and upper diagonals are not necessarily positive, but simply non-null. This connects such transformations with general quasi-definite functionals on the real line or, in other words, with orthogonal polynomials with respect to signed measures supported on the real line. Although this generalization seems to be missing in the Cholesky approach to Darboux transformations, the quasi-definite case can be tackled resorting to generalized Cholesky factorizations AEA^+ , $A \in \mathcal{T}$, involving diagonal matrices E whose diagonal entries are ± 1 ('sign matrices' in short). The matrices playing the role of CMV for this generalization constitute a wider class of matrices \mathcal{C} which can be non-unitary, but are always quasi-unitary [76] with respect to a certain sign matrix E , i.e.

$$\mathcal{C}E\mathcal{C}^+ = E = \mathcal{C}^+E\mathcal{C}.$$

We will call quasi-CMV matrices to such a generalization of CMV, given by (67). Although not named in this way, quasi-CMV matrices were introduced for the first time by some of the authors in the paper [11] which rediscovered CMV a decade after Watkins' one [74]. Actually, [11] introduces directly quasi-CMV matrices, from which CMV arises as a subclass.

In this section we will briefly discuss these two generalizations of Darboux transformations for CMV matrices.

7.1 Higher degree transformations

Darboux transformations of CMV matrices can be associated with any Hermitian Laurent polynomial of arbitrary degree. They are defined exactly as in the case of degree one. Since Section 2.2, central for the rest of the paper, is developed for an arbitrary degree, the generalization to higher degree of the results previously obtained for degree one is straightforward. Therefore, the aim of this section is not a detailed discussion of higher degree Darboux transformations, but to give an account of those features which depend on the degree.

Let ℓ be a Hermitian Laurent polynomial of degree d , i.e. with $N_z = 2d$ zeros counting multiplicity. Given a CMV matrix \mathcal{C} , the Cholesky factorization $\ell(\mathcal{C}) = AA^+$, $A \in \mathcal{T}$, exists iff $\ell(\mathcal{C})$ is positive definite, and then it is unique (see Appendix 9). Due to the CMV shape (9) of \mathcal{C} , the Hermitian matrix $\ell(\mathcal{C})$ is $2N_z + 1$ -diagonal with non-null entries in the upper and lower diagonals, so A is $N_z + 1$ -band with non-null

entries in the lower diagonal, i.e.

$$A = \begin{pmatrix} \begin{matrix} + \\ * + \\ * * + \\ * * * + \\ \vdots \vdots \vdots \vdots \\ * * \dots * * + \\ \otimes * * \dots * * + \\ \otimes * * \dots * * + \\ \otimes * * \dots * * + \\ \dots \dots \dots \end{matrix} \\ \dots \end{pmatrix}. \quad (63)$$

This defines the higher degree Darboux transformations $\mathcal{C} \xrightarrow{\ell} \mathcal{D} = A^{-1}\mathcal{C}A$ of CMV matrices. As in the case of degree one, they constitute the matrix translation of the mapping $u \xrightarrow{\ell} v = u\ell$ between positive definite functionals. Actually, the positive definiteness of $\ell(\mathcal{C})$, which guarantees its Cholesky factorization, is also equivalent to the positive definiteness of $u\ell$, as well as to the unitarity of \mathcal{D} , which is therefore CMV. Besides, the same factor A appearing in the Cholesky factorization of $\ell(\mathcal{C})$ gives the reversed Cholesky factorization $\ell(\mathcal{D}) = A^+A$.

Since every non-constant Hermitian Laurent polynomial factorizes into degree one ones, the functional interpretation of higher degree Darboux transformations implies their equivalence with compositions of degree one transformations. However, this is not the case for higher degree Darboux transformations with parameters, which exhibit a higher richness of spurious solutions than mere compositions of degree one transformations. To show this, consider a Hermitian Laurent polynomial ℓ with N_z zeros, a CMV matrix \mathcal{D} and a $N_z + 1$ -band triangular matrix $A \in \mathcal{T}$. Let us rewrite $\ell(\mathcal{D})$ and A as $N_z \times N_z$ -block matrices,

$$\ell(\mathcal{C}) = \begin{pmatrix} L_0 & M_0^+ \\ M_0 & L_1 & M_1^+ \\ & M_1 & L_2 & M_2^+ \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \quad L_n = L_n^+, \quad M_n = \begin{pmatrix} \otimes & * & * & \dots & * & * \\ & \ddots & & & \vdots & \\ & & \otimes & * & * & \vdots \\ & & & \ddots & & \vdots \\ & & & & \otimes & * \\ & & & & & \ddots \\ & & & & & & \otimes & * \\ & & & & & & & \ddots \\ & & & & & & & & \otimes & * \\ & & & & & & & & & \ddots \end{pmatrix},$$

$$A = \begin{pmatrix} R_0 \\ S_0 & R_1 \\ S_1 & R_2 \\ & \ddots & \ddots \end{pmatrix}, \quad R_n = \begin{pmatrix} + \\ * + \\ * * + \\ * * * + \\ \vdots \vdots \vdots \vdots \\ * * \dots * * + \\ * * \dots * * + \\ * * \dots * * + \\ \dots \dots \dots \end{pmatrix}, \quad S_n = \begin{pmatrix} \otimes & * & * & \dots & * & * \\ & \ddots & & & \vdots & \\ & & \otimes & * & * & \vdots \\ & & & \ddots & & \vdots \\ & & & & \otimes & * \\ & & & & & \ddots \\ & & & & & & \otimes & * \\ & & & & & & & \ddots \\ & & & & & & & & \otimes & * \\ & & & & & & & & & \ddots \end{pmatrix}.$$

Note that R_0 is the $N_z \times N_z$ leading submatrix of A highlighted in (63), and M_n , R_n , S_n are non-singular for all $n \geq 0$. Then, the reversed Cholesky factorization $\ell(\mathcal{D}) = A^+A$ reads as

$$R_n^+ R_n + S_n^+ S_n = L_n, \quad R_{n+1}^+ S_n = M_n, \quad n \geq 0. \quad (64)$$

Since S_n is non-singular, it has a unique polar decomposition $S_n = V_n |S_n|$ with V_n unitary and $|S_n|$ the positive definite square root of $S_n^+ S_n$. This allows us to rewrite (64) as

$$|S_n|^2 = L_n - R_n^+ R_n, \quad |S_n| M_n^{-1} = V_n^+ (R_n^+)^{-1}.$$

Therefore, R_n determines $|S_n|$ via the first equation, while $|S_n|$ fixes V_n and R_{n+1} , provided by the QR factors of $|S_n|M_n^{-1}$ as the second equation shows. This implies that all the blocks of A are inductively generated starting from a choice of R_0 , which is only restricted by the condition that $|S_n|$ must be positive definite for all $n \geq 0$. Therefore, apart from this constraint, the entries of R_0 can be considered as ‘free’ parameters describing all the reversed Cholesky factorizations of $\ell(\mathcal{D})$. In the case $N_z = 2$, using the notation (27) we get $R_0 = \begin{pmatrix} r_0 & 0 \\ s_0 & r_1 \end{pmatrix}$, which encodes the ‘free’ parameters already known for degree one.

The number of non-trivial entries of R_0 is $N_z(N_z + 1)/2$, among them N_z positive (the diagonal ones) and the rest are, in general, complex numbers. Therefore, the number of ‘free’ real parameters of a Darboux transformation with parameters is N_z^2 , where N_z is the number of zeros of the corresponding Hermitian Laurent polynomial. This in contrast with the composition of degree one Darboux transformations with parameters associated with $d = N_z/2$ factors of the Hermitian Laurent polynomial, which only leads to $4d = 2N_z$ ‘free’ real parameters.

Although the composition of degree one Darboux transformations with parameters gives in principle less spurious solutions than higher degree ones, this iterative procedure to deal with higher degree inverse Darboux has a drawback because such a composition is not feasible in most of the cases. The reason is that every single degree one step gives many spurious solutions which are not CMV neither unitary, preventing us from composing with a new CMV Darboux transformation. This forces us to take into account the CMV conditions (38) at any step. Then, the resulting constrained iterative procedure should end up with a set of solutions parametrized by d complex numbers in the open unit disk, playing the role of the first d Schur parameters of the CMV solutions.

Another reason could increase even more the interest of higher degree Darboux transformations with parameters for CMV: The spurious solutions might have their own interest if associated with non-standard orthogonal polynomials on the unit circle. There is a strong reason to expect an interesting connection with non-standard orthogonality. Higher degree Darboux transformations with parameters for Jacobi matrices have been already considered in the literature [24, 31]. Contrary to the case of degree one, they present spurious solutions which are related to Sobolev orthogonal polynomials on the real line [20, 21].

The parametric analysis of spurious solutions for Jacobi follows the same arguments given previously, since they work for any matrix which, like $\ell(\mathcal{C})$, can be considered as a block-Jacobi one. Then, if a real polynomial \wp has N_z zeros (coinciding now with the degree), the corresponding Darboux transformations with parameters of a Jacobi matrix \mathcal{J} are described by $N_z(N_z + 1)/2$ ‘free’ parameters, those encoded in the $N_z \times N_z$ leading submatrix of $\wp(\mathcal{J})$. The only quirks of the Jacobi case is that all these parameters are now real and that, in contrast to the CMV case, any (not necessarily even) number N_z of zeros is possible. Actually, the fact that the minimal step for Jacobi and CMV is $N_z = 1$ and $N_z = 2$, respectively, is the cause of the possibility of avoiding spurious solutions in the Jacobi but not in the CMV case.

Due to the equivalence with polynomial transformations of measures, higher degree Darboux transformations of Jacobi and CMV matrices are both almost isospectral. More precisely, the Darboux and inverse Darboux transformations may delete or add, respectively, at most N_z eigenvalues.

7.2 Quasi-CMV matrices

Quasi-CMV matrices are the matrix representations of the multiplication operator (4) with respect to zig-zag bases which are ‘orthonormal’ with respect to quasi-definite functionals on the unit circle. Such functionals are those Hermitian linear functionals u on Λ whose Gram matrix in $\mathbb{P}_n = \text{span}\{1, z, z^2, \dots, z^n\}$ is non-singular for all $n = 0, 1, 2, \dots$ (equivalently, the Gram matrix of u in \mathbb{P} with respect to a basis $p_n \in \mathbb{P}_n \setminus \mathbb{P}_{n-1}$ is quasi-definite, i.e. all its leading submatrices are non-singular). This property, independent of the basis chosen for the Gram matrix, is equivalent to the existence of a sequence of polynomials $\varphi_n \in \mathbb{P}_n \setminus \mathbb{P}_{n-1}$ which are ‘orthonormal’ with respect to u , i.e. $u[\varphi_n \varphi_m^*] = e_n \delta_{n,m}$ with $e_n \in \{1, -1\}$. The polynomials φ_n satisfy forward and backward recurrence relations which generalize (22),

$$\begin{cases} \rho_n \varphi_n(z) = z \varphi_{n-1}(z) + a_n \varphi_{n-1}^*(z), & \rho_n = \sqrt{|1 - |a_n|^2|}, \\ \hat{\rho}_n z \varphi_{n-1}(z) = \varphi_n(z) - a_n \varphi_n^*(z), & \hat{\rho}_n = \frac{e_n}{e_{n-1}} \rho_n = \text{sg}(1 - |a_n|^2) \rho_n, \end{cases} \quad (65)$$

where $\text{sg}(x) = x/|x|$, $x \in \mathbb{R} \setminus \{0\}$, is the sign function and the Schur parameters a_n are no longer restricted to the open unit disk, but they can lie anywhere in the complex plane outside of the unit circle.

The previous extension of orthogonal polynomials on the unit circle goes back at least to the works of Geronimus [33]. However, the connection of the quasi-definite case with ‘orthonormal’ Laurent polynomials and quasi-CMV matrices, which we summarize below, is of much more recent vintage [11].

Since $f(z) \mapsto z^{-[n/2]} f(z)$ is a one-to-one correspondence between \mathbb{P}_n and \mathbb{L}_n which preserves the sesquilinear form $(f, g) \mapsto u[fg^*]$, the quasi-definiteness of u also means that its Gram matrix in \mathbb{L}_n is non-singular for all $n = 0, 1, 2, \dots$ (equivalently, the Gram matrix of u in Λ with respect to a basis $l_n \in \mathbb{L}_n \setminus \mathbb{L}_{n-1}$ is quasi-definite). Therefore, analogously to the previous discussion, the quasi-definiteness of u is equivalent to the existence of a zig-zag basis χ which is ‘orthonormal’ with respect to u , i.e. $u[\chi_n \chi_m^*] = e'_n \delta_{n,m}$, $e'_n \in \{1, -1\}$. Actually, as in the positive definite case, the relation between the orthonormal polynomials φ_n and the orthonormal Laurent polynomials χ_n is given by (24). This implies that $e'_n = e_n$, so that the orthonormality condition for χ can be rewritten as $u[\chi \chi^+] = E$, where E denotes the sign matrix

$$E = \begin{pmatrix} e_0 & & & \\ & e_1 & & \\ & & e_2 & \\ & & & \ddots \end{pmatrix}, \quad \begin{aligned} e_0 &= \text{sg}(u[1]), \\ e_n &= e_0 \prod_{k=1}^n \text{sg}(1 - |a_k|^2), \quad n \geq 1. \end{aligned} \quad (66)$$

If u is a quasi-definite functional associated with the quasi-CMV matrix \mathcal{C} and E is the corresponding sign matrix, then we get

$$u\ell[\chi\chi^+] = u[\ell\chi\chi^+] = \ell(\mathcal{C})u[\chi\chi^+] = \ell(\mathcal{C})E.$$

As a consequence, we obtain the quasi-definite version of Corollary 3.3: $u\ell$ is quasi-definite iff $\ell(\mathcal{C})E$ is quasi-definite. The last condition is also equivalent to the existence of a generalized Cholesky factorization (see Appendix 9)

$$\ell(\mathcal{C})E = AFA^+, \quad A \in \mathcal{T}, \quad F \in \mathcal{S} = \text{set of infinite sign matrices}, \quad (68)$$

which is also unique. Since $\ell(\mathcal{C})$ has the five-diagonal structure (14), as in the CMV case, A has the 3-band shape (15). The factorization (68) is the starting point for the Darboux transformations of quasi-CMV matrices.

To introduce such transformations we just follow the same steps as in the case of CMV matrices. From Proposition 2.10, the zig-zag basis ω given by $\chi = A\omega$ and the matrix B defined by $\ell\omega = B\chi$ yield the factorization $\ell(\mathcal{C}) = AB$. Comparing this with (68), taking care of the associativity with Proposition 2.3 and bearing in mind that $\ker_R(A) = \{0\}$, we conclude that B is the 3-band upper triangular matrix

$$B = FA^+E. \quad (69)$$

Using again Proposition 2.10 we also find that $\mathcal{D} = A^{-1}\mathcal{C}A$ is the matrix of the multiplication operator with respect to ω and satisfies

$$F\ell(\mathcal{D}) = A^+EA, \quad A \in \mathcal{T}, \quad E \in \mathcal{S}, \quad (70)$$

which is a generalized reversed Cholesky factorization. Besides, the associativity properties of Proposition 2.3 and (69) yield

$$u\ell[\omega\omega^+] = u[\ell\omega\omega^+] = Bu[\chi\chi^+](A^+)^{-1} = F.$$

This identity has several consequences: The zig-zag basis ω is orthonormal with respect to the quasi-definite functional $v = u\ell$, thus \mathcal{D} is a quasi-CMV matrix and F is the corresponding sign matrix.

Sumarizing, given a quasi-CMV matrix \mathcal{C} with sign matrix E and a Hermitian Laurent polynomial ℓ such that $\ell(\mathcal{C})E$ is quasi-definite, we have found a transformation $\mathcal{C} \xrightarrow{\ell} \mathcal{D} = A^{-1}\mathcal{C}A$ which generates a new quasi-CMV matrix \mathcal{D} . The ingredients of this transformation, i.e. the lower triangular matrix A and the sign matrix F of \mathcal{D} , are given by the generalized Cholesky factorization (68) of $\ell(\mathcal{C})E$. The matrix $F\ell(\mathcal{D})$ satisfies a generalized reversed Cholesky factorization obtained from that of $\ell(\mathcal{C})E$ by changing $A \rightarrow A^+$ and $F \rightarrow E$. This is what we call the Darboux transformation for quasi-CMV matrices. In perfect agreement with the CMV case, this transformation is the matrix translation of the transformation $u \xrightarrow{\ell} v = u\ell$ connecting quasi-definite functionals.

Concerning inverse Darboux transformations, starting from a quasi-CMV matrix \mathcal{D} and its sign matrix F , the matrix transformation $\mathcal{C} = A\mathcal{D}A^{-1}$, with A and E given by (70), defines the Darboux transformations with parameters for quasi-CMV matrices. This matrix procedure solves the inverse problem, which consists in finding all the quasi-CMV matrices \mathcal{C} such that $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$, but also yields spurious solutions \mathcal{C} which are neither quasi-CMV nor quasi-unitary. Analogously to the CMV case, the solutions of the inverse Darboux transformations coincide with the quasi-CMV solutions of the Darboux transformations with parameters. To prove this note that, if ω is an orthonormal zig-zag basis for \mathcal{D} and $\chi = A\omega$ then, from (70) and Proposition 2.3, $\ell\omega = \ell(\mathcal{D})\omega = FA^+EA\omega = B\chi$ with B given by (69). In view of Proposition 2.10, we find that $\mathcal{D} = A^{-1}\mathcal{C}A$ with $\ell(\mathcal{C}) = AB = AFA^+E$. Since $\ell(\mathcal{C})E = AFA^+$ is Hermitian, Lemma 7.1 proves that, whenever \mathcal{C} is quasi-CMV, E is its sign matrix up to a sign. Therefore, $\mathcal{C} \xrightarrow{\ell} \mathcal{D}$ for such a quasi-CMV matrix \mathcal{C} .

The theory of Darboux transformations for quasi-CMV matrices can be developed just by mimicking the steps and arguments given for the CMV case. All the results obtained in the previous sections have a natural generalization to the quasi-CMV case, the details are left to the reader. We will simply mention some specific results of interest for the example discussed below.

Assuming the notation (27) for the matrix coefficients of the factor A and denoting by f_n the diagonal entries of the sign matrix F , the factorization (68) reads as

$$\begin{cases} f_{n-2}|t_{n-2}|^2 + f_{n-1}|s_{n-1}|^2 + f_n r_n^2 = e_n (\beta - 2\operatorname{Re}(\alpha \bar{a}_n a_{n+1})), \\ f_{n-1} s_{n-1} \bar{t}_{n-1} + f_n r_n s_n = e_{n+1} \rho_{n+1} (\bar{\alpha} a_n - \alpha a_{n+2}), \\ f_n r_n t_n = \alpha e_{n+2} \rho_{n+1} \rho_{n+2}, \end{cases} \quad n \geq 0. \quad (71)$$

These equations, together with the initial conditions $t_{-2} = t_{-1} = s_{-1} = 0$, make explicit the Darboux transformations for quasi-CMV matrices. Besides, generalizing the identities in (29), (30) and (40) we get the following direct relations between the matrix coefficients of A and the Schur parameters of the quasi-CMV matrices involved in a Darboux transformation,

$$\begin{aligned} \frac{r_{n-1}}{r_n} &= \frac{\rho_n}{\sigma_n}, & \frac{t_{n-1}}{t_n} &= \frac{\hat{\sigma}_n}{\hat{\rho}_{n+2}}, \\ a_n &= b_n + s_{n-1} \frac{\sigma_n}{r_n}, & \alpha a_{n+2} &= \bar{\alpha} b_n - e_{n+1} f_n s_n \frac{r_{n+1}}{\sigma_{n+1}}. \end{aligned} \quad (72)$$

As for the Darboux transformations with parameters, inserting (27) into the factorization (70) shows that they are explicitly given by

$$\begin{cases} e_n r_n^2 + e_{n+1} |s_n|^2 + e_{n+2} |t_n|^2 = f_n (\beta - 2\operatorname{Re}(\alpha \bar{b}_n b_{n+1})), \\ e_{n+1} s_n r_{n+1} + e_{n+2} t_n s_{n+1} = f_n \sigma_{n+1} (\bar{\alpha} b_n - \alpha b_{n+2}), \\ e_{n+2} t_n r_{n+2} = \alpha f_n \sigma_{n+1} \sigma_{n+2}, \end{cases} \quad n \geq 0, \quad (73)$$

where b_n are the Schur parameters of \mathcal{D} and $\sigma_n = \sqrt{|1 - |b_n|^2|}$. These relations determine e_n, r_n, s_n, t_n once e_0, r_0, s_0, e_1, r_1 are chosen, as follows from rewriting

(73) as

$$\begin{aligned}
e_{n+2} &= \text{sg}(\varepsilon_n), & r_{n+2} &= |\alpha| \frac{\sigma_{n+1}\sigma_{n+2}}{\tau_n}, & t_n &= \frac{\alpha}{|\alpha|} e_{n+2} f_n \tau_n, \\
s_{n+1} &= \frac{\bar{\alpha}}{|\alpha|} \frac{\sigma_{n+1}(\bar{\alpha}b_n - \alpha b_{n+2}) - e_{n+1} f_n s_n r_{n+1}}{\tau_n}, \\
\tau_n &= \sqrt{|\varepsilon_n|}, & \varepsilon_n &= f_n (\beta - 2\text{Re}(\alpha \bar{b}_n b_{n+1})) - e_n r_n^2 - e_{n+1} |s_n|^2.
\end{aligned} \tag{74}$$

Therefore, the solutions of the Darboux transformations with parameters can be parametrized by $e_0 r_0^2$, s_0 , $e_1 r_1^2$, which give again four real ‘free’ parameters. The existence of such a solution for a given value of these free parameters is equivalent to state that $\varepsilon_n \neq 0$ for all $n \geq 0$. Among these solutions, the quasi-CMV ones are characterized by a generalization of the CMV conditions (38), namely,

$$\begin{aligned}
e_0 f_0 r_0^2 &= \beta - 2\text{Re}(\alpha a), & r_1 &= r_0 \frac{\sigma_1}{\rho}, & |a| &\neq 1, \\
a &= b_1 + \frac{s_0}{r_1} \sigma_1, & e_1 &= e_0 \text{sg}(1 - |a|^2), & \rho &= \sqrt{|1 - |a|^2|}.
\end{aligned} \tag{75}$$

The conditions in the second column of (75) can be combined into $e_1 f_1 r_1^2 (1 - |a|^2) = e_0 f_0 r_0^2 (1 - |b_1|^2)$. Like in the CMV case, a is the first Schur parameter a_1 of the corresponding quasi-CMV solution, while the pair of equalities in the first column of the quasi-CMV conditions (75) select the solutions associated with (not necessarily quasi-definite) Hermitian functionals.

Example 7.2. *Darboux transformation with parameters for a general Hermitian Laurent polynomial of degree one $\ell(z) = \alpha z + \beta + \bar{\alpha} z^{-1}$ and $\mathcal{D} = \mathcal{S}$ the shift matrix given in (8).*

We will illustrate with this example the coexistence of CMV, quasi-CMV and spurious solutions in the Darboux transformations with parameters. The shift matrix $\mathcal{D} = \mathcal{S}$ is a CMV with Schur parameters $b_n = 0$ (so, $\sigma_n = 1$) for $n \geq 1$, thus $F = I$ is a sign matrix for \mathcal{S} . We will search for solutions \mathcal{C} of the corresponding Darboux transformation with parameters, considered in the generalized sense described in this section. This allows the appearance of quasi-CMV solutions \mathcal{C} with a non-trivial sign matrix E . Without loss of generality we can take $|\alpha| = 1$, which simply means that we avoid a trivial positive rescaling of the factor A in the reversed factorization $\ell(\mathcal{S}) = A^+ E A$, where

$$\ell(\mathcal{S}) = \begin{pmatrix} \beta & \alpha & \bar{\alpha} & & & & \\ \bar{\alpha} & \beta & 0 & \alpha & & & \\ \alpha & 0 & \beta & 0 & \bar{\alpha} & & \\ & \bar{\alpha} & 0 & \beta & 0 & \alpha & \\ & & \alpha & 0 & \beta & 0 & \bar{\alpha} \\ & & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{76}$$

The diagonal entries e_n of the sign matrix E , as well as the coefficients r_n , s_n , t_n giving and the three diagonals of A , are determined by (74). In what follows we

will restrict our attention to the subset of solutions with $s_1 = 0$, a condition which becomes a constraint between the parameters $e_0r_0^2$, s_0 , $e_1r_1^2$ describing the whole set of solutions. Therefore, the referred subset of solutions can be parametrized by $e_0r_0^2$ and $e_1r_1^2$, although instead we will use $e_1r_1^2$ and $e_2r_2^2$ for convenience. The corresponding parametrization of A and E , as follows from (74), is given by

$$\begin{aligned} s_0 &= \bar{\alpha} \frac{e_1}{r_1}, & e_0r_0^2 &= \beta - e_1|s_0|^2 - \frac{1}{e_2r_2^2} = \beta - \frac{1}{e_1r_1^2} - \frac{1}{e_2r_2^2}, \\ s_n &= 0, & e_{n+2}r_{n+2}^2 &= \frac{1}{\beta - e_nr_n^2}, & t_{n-1} &= \alpha \frac{e_{n+1}}{r_{n+1}}, & n &\geq 1. \end{aligned} \quad (77)$$

These equations are governed by a single recurrence relation

$$x_{n+1} = \frac{1}{\beta - x_n}, \quad x_n = \begin{cases} e_{2n+1}r_{2n+1}^2, \\ \text{or} \\ e_{2n+2}r_{2n+2}^2, \end{cases} \quad n \geq 0, \quad (78)$$

which provides separately the odd and even coefficients r_n , t_n depending whether we use the initial condition $x_0 = e_1r_1^2$ or $x_0 = e_2r_2^2$. The initial conditions generating true solutions of the Darboux transformation with parameters are those giving $x_n \neq \beta$ for $n \geq 0$ and guaranteeing that $e_0r_0^2 \neq 0$, i.e.

$$\frac{1}{e_1r_1^2} + \frac{1}{e_2r_2^2} \neq \beta. \quad (79)$$

Recurrence (78) yields the continued fraction expansion

$$x_n = \cfrac{1}{\cfrac{1}{\beta} - \cfrac{1}{\cfrac{1}{\beta} - \cfrac{1}{\cfrac{1}{\beta} - \cfrac{1}{\beta} - \dots - \cfrac{1}{\beta} - \cfrac{1}{\beta - x_0}}}}.$$

Using basic results from the theory of continued fractions [72], we obtain

$$x_n = \frac{U_{n-1}(\frac{\beta}{2}) - x_0 U_{n-2}(\frac{\beta}{2})}{U_n(\frac{\beta}{2}) - x_0 U_{n-1}(\frac{\beta}{2})}, \quad n \geq 1,$$

where U_n are the second kind Chebyshev polynomials, given by

$$U_{n+1}(t) = 2tU_n(t) - U_{n-1}(t), \quad U_{-1} = 0, \quad U_0 = 1. \quad (80)$$

Therefore, if $s_1 = 0$, the reversed Cholesky factorization $\ell(\mathcal{C}) = A^+EA$ is possible exactly for those choices of $e_1r_1^2$ and $e_2r_2^2$ satisfying (79) and

$$e_1r_1^2, e_2r_2^2 \neq \frac{U_n(\frac{\beta}{2})}{U_{n-1}(\frac{\beta}{2})}, \quad n \geq 1. \quad (81)$$

The solutions associated with a Hermitian functional are characterized by the two equalities in the first column of (75), which in this case become

$$e_0 r_0^2 = \beta - \frac{2}{e_1 r_1^2}. \quad (82)$$

Using the expression for r_0 given in (77), the above hermiticity condition turns out to be $e_1 = e_2$ and $r_1 = r_2$. Therefore, bearing in mind (79) and (81), the solutions leading to Hermitian functionals are those generated by (78) with initial conditions satisfying

$$\frac{2}{\beta} \neq x_0 \neq \frac{U_n(\frac{\beta}{2})}{U_{n-1}(\frac{\beta}{2})}, \quad e_1 r_1^2 = x_0 = e_2 r_2^2, \quad n \geq 1. \quad (83)$$

The quasi-CMV solutions are those satisfying the previous hermiticity conditions together with those summarized in the second and third column of (75). In our case, these additional conditions read as

$$|a| \neq 1, \quad x_0(1 - |a|^2) = e_0 r_0^2, \quad a = \frac{\bar{\alpha}}{x_0},$$

which, using (82), become

$$|x_0| \neq 1, \quad x_0 = \frac{1}{\beta - x_0}.$$

This means that the quasi-CMV solutions are associated with those fixed points $x_n = x_0$ of (78) lying outside of the unit circle. Since $s_n = 0$ for $n \geq 1$, we find from (72) that the Schur parameters of these quasi-CMV solutions are $a_1 = a$ and $a_n = 0$ for $n \geq 2$. In other words, all the quasi-CMV solutions with $s_1 = 0$ are Bernstein-Szegő type.

The fixed points $x_n = x_0$ of (78) are $x_0 = \frac{\beta}{2} \pm \sqrt{(\frac{\beta}{2})^2 - 1}$, which are inverse of each other, while the requirement $|x_0| \neq 1$ translates into the inequality $|\beta| > 2$. It is convenient to rewrite these two fixed points as

$$x_0^\pm = \text{sg}(\beta) \left(\left| \frac{\beta}{2} \right| \pm \sqrt{\left(\frac{\beta}{2} \right)^2 - 1} \right), \quad (84)$$

so that $|x_0^\pm| = \left| \frac{\beta}{2} \right| \pm \sqrt{\left(\frac{\beta}{2} \right)^2 - 1}$ for $|\beta| > 2$. Then, $|x_0^+| > 1$ and $|x_0^-| < 1$. Both fixed points satisfy the hermiticity conditions (83), thus they give rise to two different quasi-CMV solutions characterized by the Schur parameters $(a_\pm, 0, 0, 0, \dots)$, with $a_\pm = \bar{\alpha}/x_0^\pm = \bar{\alpha}x_0^\mp$. Since $|a_+| < 1$ and $|a_-| > 1$, the Bernstein-Szegő solution related to the initial condition $x_0 = x_0^+$ is positive definite, while that one corresponding to $x_0 = x_0^-$ is quasi-definite but not positive definite. The related CMV matrix, \mathcal{C}_+ ,

subset. The simplest ones should be those obtained by fixing $e_1 r_1^2 = e_2 r_2^2 = x_0$ as a fixed point of (78) for $|\beta| \leq 2$. These fixed points lie on the unit circle, so they are candidates to generate spurious solutions only when $x_0 = \pm 1$, which corresponds to $\beta = \pm 2$. However, choosing $e_1 r_1^2 = e_2 r_2^2 = \pm 1$ for $\beta = \pm 2$ gives no solution of the Darboux transformation with parameters because, using (77), we find that $e_0 r_0^2 = 0$. Thus, no spurious solution with $s_1 = 0$ has an eventually constant sequence $e_n r_n^2$.

We can reinterpret the previous results by stating that, in the case $s_1 = 0$, the quasi-CMV solutions are those whose factors A and E have constant diagonals up to finitely many entries. Therefore, the solutions given by factors A and E with eventually periodic diagonals are spurious whenever the period is greater than one. These are the simplest spurious solutions, which are generated by the periodic solutions of the recurrence (78).

The periodic solutions of (78) with period j are characterized by the relation

$$x_0 = \frac{U_{j-1}(\frac{\beta}{2}) - x_0 U_{j-2}(\frac{\beta}{2})}{U_j(\frac{\beta}{2}) - x_0 U_{j-1}(\frac{\beta}{2})}.$$

Using (80), this equation can be rewritten as

$$(x_0^2 + \beta x_0 + 1) U_{j-1}(\frac{\beta}{2}) = 0,$$

which means that either $x_0^2 + \beta x_0 + 1 = 0$ or $U_{j-1}(\frac{\beta}{2}) = 0$.

The zeros of $x_0^2 + \beta x_0 + 1$ are the fixed points $x_0 = x_0^\pm$ given in (84) and leading to the quasi-CMV solutions previously found. Therefore, if $U_j(\frac{\beta}{2}) \neq 0$ for every $j \geq 1$, the spurious solutions we are interested in may arise only from two alternatives: $e_1 r_1^2 = x_0^+$, $e_2 r_2^2 = x_0^-$ or $e_1 r_1^2 = x_0^-$, $e_2 r_2^2 = x_0^+$. However, (77) shows that both options lead to $e_0 r_0^2 = \beta - x_0^+ - x_0^- = 0$, thus none of them yield a solution of the Darboux transformation with parameters.

Therefore, the kind of spurious solutions we are looking for only appear when $\beta = \beta_{j,k}$ is a zero of $U_{j-1}(\frac{\beta}{2})$,

$$\beta_{j,k} = 2 \cos \left(\frac{k}{j} \pi \right), \quad k = 1, 2, \dots, j-1.$$

Actually, for these values of β all the solutions are spurious because quasi-CMV solutions require $|\beta| > 2$.

Consider for instance the simplest case, $j = 2$, so that $\beta = 0$. Then, (78) generates a 2-periodic sequence $x_0, -x_0^{-1}, x_0, -x_0^{-1}, \dots$ for every initial value $x_0 \neq 0$. Therefore, in the case $s_1 = 0$, the solutions of the Darboux transformation with parameters whose factors A and E have eventually 4-periodic diagonals are obtained for $\beta = 0$. The parameters $e_1 r_1^2, e_2 r_2^2 \in \mathbb{R} \setminus \{0\}$ describing these solutions are only restricted by (79), which in this case reads as $e_1 r_1^2 + e_2 r_2^2 \neq 0$. Although these solutions are necessarily spurious, the associated functionals are Hermitian when $e_1 r_1^2 = e_2 r_2^2$, as required by hermiticity condition (83). In the Hermitian case, the factors A and E ,

case, the minimal step of Darboux for CMV involves polynomials with two zeros. These facts cause a number of differences between Darboux for Jacobi and CMV which we summarize below, together with the open problems that they suggest:

- Darboux transformations of Jacobi matrices are based on a factorization of a matrix which retains the hermiticity of the Jacobi one. However, the Darboux transformations of CMV matrices introduced in this paper need the factorization of matrices which do not share the unitarity of CMV, but are also Hermitian. The reason for this is that the method we have used to obtain a unitary version of Darboux is to relate a unitary matrix with a Hermitian one to which we apply the standard Darboux procedure for self-adjoint operators. Is it possible to develop a new theory of Darboux transformations for unitary operators without abandoning unitarity (i.e. based on splittings of unitary operators built out of the previous ones)?
- Darboux transformations of Jacobi matrices are usually implemented via LU factorizations of non-symmetric tridiagonal matrices associated with monic instead of orthonormal polynomials. This allows for the unified treatment of the general quasi-definite case. Nevertheless, in the positive definite case they can be formulated in terms of Cholesky factorizations. This opens a way for the CMV generalization of Darboux transformations as defined in this paper, while Section 7.2 shows that generalized Cholesky factorizations permit their extension to the quasi-definite case. However, a question remains: Is there any CMV analogue of the LU version for the Darboux transformations (i.e. based on matrix representations with respect “monic” zig-zag bases)?
- Darboux transformations for Jacobi matrices can be implemented via Cholesky factorizations and reversed ones or, equivalently, using the Cholesky factor as a change of basis relating Jacobi matrices by conjugation. While the latter method is directly generalizable to CMV matrices, the first one seems hard to tackle in the CMV case due to the drawbacks in recovering an infinite matrix from its image by an Hermitian Laurent polynomial. This leads to the following question: Given a Hermitian Laurent polynomial of degree one, is there any simple procedure to recover a CMV matrix from the result of evaluating the Laurent polynomial on such a matrix?
- In the Jacobi case, the Darboux transformations with parameters can be identified with the inverse Darboux transformations, which correspond exactly to the Geronimus transformations of measures on the real line. The Geronimus transformations of measures on the unit circle are a particular case of dividing a measure by an Hermitian Laurent polynomial and adding Dirac deltas at their zeros. These more general transformations of measures on the unit circle correspond exactly to the inverse Darboux transformations of CMV matrices, whose matrix realization is given by the Darboux transformations with parameters. However, these latter transformations present spurious solutions

which are not CMV neither unitary matrices, and can be eventually associated with non-Hermitian functionals. Is there any interesting interpretation of these spurious solutions? Can they be understood in terms of orthonormal Laurent polynomials with respect to non-standard or matrix inner products?

A positive answer to the last item is suggested by the recently found connection between higher degree Darboux transformations of Jacobi matrices and Sobolev orthogonal polynomials on the real line [20, 21]. A similar connection for the CMV case should be the starting point for a new approach to non-standard orthogonality on the unit circle. Other promising lines of future research are indicated by the success of Darboux techniques for Jacobi matrices in problems of numerical linear algebra, the study of integrable systems or the search for bispectral situations. The recent connection found between CMV matrices and the so called quantum walks [9] also suggests a possible use of Darboux for CMV in the study of discrete quantum dynamical systems.

9 Appendix: Cholesky factorizations

Given an infinite Hermitian complex matrix M , the existence of a generalized Cholesky factorization (GCF)

$$M = AEA^+, \quad A \in \mathcal{T}, \quad E \in \mathcal{S}, \quad (86)$$

is equivalent to state that M is quasi-definite, i.e. all its leading submatrices are non-singular. Obviously the condition is necessary because, denoting by the subindex n the leading submatrix of order n , (86) implies that $M_n = A_n E_n A_n^+$. The sufficiency of the condition follows by an inductive reasoning on the order of the leading submatrices of M . More precisely, if $M_n = A_n E_n A_n^+$ is the GCF of the non-singular leading submatrix M_n and

$$M_{n+1} = \begin{pmatrix} M_n & Y_n^+ \\ Y_n & y_n \end{pmatrix}, \quad Y_n \in \mathbb{C}^n, \quad y_n \in \mathbb{R},$$

$$A_{n+1} = \begin{pmatrix} A_n & 0 \\ X_n & x_n \end{pmatrix}, \quad X_n \in \mathbb{C}^n, \quad x_n > 0; \quad E_{n+1} = \begin{pmatrix} E_n & 0 \\ 0 & e_n \end{pmatrix}, \quad e_n \in \{1, -1\},$$

then

$$M_{n+1} = A_{n+1} E_{n+1} A_{n+1}^+ \Leftrightarrow \begin{cases} A_n E_n X_n^+ = Y_n^+, \\ X_n E_n X_n^+ + e_n x_n^2 = y_n, \end{cases} \Leftrightarrow \begin{cases} X_n^+ = E_n A_n^{-1} Y_n^+, \\ e_n x_n^2 = y_n - Y_n M_n^{-1} Y_n^+. \end{cases}$$

These equations have solutions $X_n \in \mathbb{C}^n$, $e_n \in \{1, -1\}$, $x_n > 0$, whenever M_{n+1} is non-singular because, then, $y_n \neq Y_n M_n^{-1} Y_n^+$ since otherwise $(-Y_n M_n^{-1}, 1) M_{n+1} = 0$. This proves that the quasi-definiteness of M guarantees the existence of the GCF of its leading submatrices, obtained by enlarging those of the smaller leading submatrices.

because $XM X^+ = XA^+AX^+ = \|AX^+\|^2 > 0$ for every finitely non-null vector X since $\ker_R(A) = \{0\}$.

In the case of finite matrices, positive definiteness is also characterized by the positivity of all the principal minors corresponding to a south-east corner submatrix. As a consequence, the existence of CF and reversed CF takes place simultaneously and is equivalent to positive definiteness. However, establishing a similar result for infinite matrices requires more technical hypothesis and arguments to deal with the non-trivial relation between the reversed factorizations of an infinite matrix and its leading submatrices. This is accounted for in the following proposition, which also summarizes the previous results for convenience.

Proposition 9.1. *If M is an infinite Hermitian complex matrix, then*

$$\begin{aligned} M \text{ has a GCF} &\Leftrightarrow M \text{ is quasi-definite} \\ M \text{ has a CF} &\Leftrightarrow M \text{ is positive definite} \Leftarrow M \text{ has a reversed CF} \end{aligned}$$

If, besides, M is band matrix with bounded entries and its lower and upper diagonals have only non-null entries, then

$$M \text{ has a reversed CF} \Leftrightarrow M \text{ is positive definite}$$

Proof. It only remains to prove the left-handed implication of the last statement. The positive definiteness of M is equivalent to the existence of a reversed CF, $M_n = A_n^+ A_n$, for every leading submatrix M_n of M . The subindex n in A_n only refers to the dependence on the order of the leading submatrix M_n . In other words, we cannot assume that all the factors A_n are leading submatrices of the same infinite matrix A because A_k is not necessarily a leading submatrix of A_n for $k < n$. Our strategy consists in proving that, despite the non-trivial relation between the different matrix factors A_n , a subsequence of them “converges” in some sense to an infinite matrix A which provides a reversed CF, $M = A^+ A$.

Bearing in mind that M is a band matrix, the boundedness of its entries is equivalent to state that it represents a bounded operator $X \mapsto MX$ in the Hilbert space \mathfrak{H} of square-summable sequences. If P_n is the orthogonal projection of \mathfrak{H} onto the subspace \mathfrak{H}_n spanned by the first n canonical vectors of \mathfrak{H} , then

$$\hat{M}_n := P_n M P_n = M_n \oplus I = \hat{A}_n^+ \hat{A}_n, \quad \hat{A}_n := A_n \oplus I,$$

where M_n and A_n act on \mathfrak{H}_n , while the identity acts on its orthogonal subspace. Since $\|\hat{A}_n\|^2 = \|\hat{M}_n\| \leq \|M\|$, there exists a subsequence of \hat{A}_n weakly converging to some bounded operator A on \mathfrak{H} [52, Theorem 5.70]. The proposition is proved if we show that $M = A^+ A$ and $A \in \mathcal{T}$, where, as in the case of M , we identify A with its matrix representation in the canonical basis of \mathfrak{H} .

Suppose that M is $2N + 1$ -diagonal. Then, each M_n is also $2N + 1$ -diagonal, hence A_n is $N + 1$ -band lower triangular with positive diagonal entries. Since weak convergence implies convergence of the matrix coefficients, we conclude that A is

$N + 1$ -band lower triangular with non-negative diagonal entries. To prove that $A \in \mathcal{I}$ we only need to see that its diagonal entries are non-null. This follows directly from the identity $M = A^+A$, which we will prove below, and the hypothesis that the lower and upper diagonals of M have no null entries.

As for the identity $M = A^+A$, note first that $\hat{M}_{n+N}X = MX$ for every $X \in \mathfrak{H}_n$ because M is $2N + 1$ -diagonal. In consequence, \hat{M}_n is strongly convergent to M because

$$\|(\hat{M}_{n+N} - M)X\| = \|(\hat{M}_{n+N} - M)(X - P_nX)\| \leq 2\|M\|\|X - P_nX\|$$

and $\|X - P_nX\| \rightarrow 0$ for every $X \in \mathfrak{H}$.

Let us denote also by \hat{A}_n the subsequence weakly converging to A , no misunderstanding will appear from this abuse of language. In contrast with strong convergence, weak limits do not commute with compositions of operators, so we cannot conclude directly that $M = A^+A$ just taking weak limits in $\hat{M}_n = \hat{A}_n^+\hat{A}_n$. Luckily enough, the weak convergence of \hat{A}_n implies its strong convergence due to the $N + 1$ -band structure of \hat{A}_n and the uniform bound $\|\hat{A}_n\| \leq \|M\|$ (see Lemma 9.2 below). Bearing in mind that strong convergence preserves compositions [48], taking strong limits in $\hat{M}_n = \hat{A}_n^+\hat{A}_n$ we finally obtain $M = A^+A$. \square

The proof of the previous proposition makes use of the following lemma.

Lemma 9.2. *Let B_n be a sequence of infinite band matrices with uniformly bounded entries and a uniformly bounded size of the band. Then, if B_n is weakly convergent, it is also strongly convergent.*

Proof. First, $\|B_n\|$ is bounded since B_n is uniformly banded with uniformly bounded entries. Therefore, the weak limit B of B_n is also bounded. Since the entries of B are the limits of the entries of B_n , the weak limit B inherits the band structure of B_n . Due to this band structure, if $X \in \mathfrak{H}$ is finitely non-null, $(B_n - B)X$ only involves a finite n -independent number of matrix entries of both, B_n and B . Therefore, $\|(B_n - B)X\| \rightarrow 0$ for every finitely non-null vector X . Given an arbitrary vector $X \in \mathfrak{H}$, consider the inequalities

$$\begin{aligned} \|(B_n - B)X\| &\leq \|(B_n - B)(X - P_kX)\| + \|(B_n - B)P_kX\| \\ &\leq (\sup_n \|B_n\| + \|B\|)\|X - P_kX\| + \|(B_n - B)P_kX\|. \end{aligned}$$

We know that $\|X - P_kX\| \xrightarrow{k} 0$, while $\|(B_n - B)P_kX\| \xrightarrow{n} 0$ because P_kX is finitely non-null for any fixed k . Thus, we can make both summands of the last line arbitrarily small by choosing k and n big enough. This implies that $\|(B_n - B)X\| \rightarrow 0$, proving that B is a strong limit of B_n . \square

The inductive proof of the existence of GCF given at the beginning of this section also shows that it is unique. On the contrary, the proof of Proposition 9.1 suggests that a positive definite Hermitian matrix can have multiple reversed CF, at least so

many as limit points has the sequence \hat{A}_n arising from the CF, $M_n = \hat{A}_n^+ A_n$, of the leading submatrices M_n . The examples presented in the previous sections prove that this is indeed the case, and reversed GCF are in general not unique.

A more abstract proof of the uniqueness of a GCF sheds light on the lack of uniqueness for reversed factorizations and its relation with the algebra of infinite matrices. Suppose that

$$A_1 E_1 A_1^+ = A_2 E_2 A_2^+, \quad A_i \in \mathcal{T}, \quad E_i \in \mathcal{S}.$$

Using the associativity laws of Proposition 2.3, we get $E_1 A_1^+ (A_2^+)^{-1} = A_1^{-1} A_2 E_2$. The left-hand side of this equality is upper triangular, while the right-hand side is lower triangular, so both must be diagonal. Thus, $A_1^+ (A_2^+)^{-1} = D = A_1^{-1} A_2$ is a positive diagonal matrix and $E_1 = E_2$. Taking care of the associativity again with Proposition 2.3, we find that $D^+ D = (A_1^+ (A_2^+)^{-1})^+ (A_1^{-1} A_2) = I$, which implies that $D = I$ and $A_2 = A_1$.

The above proof of the uniqueness of GCF works equally well for finite or infinite matrices thanks to the associativity of the involved operations, ensured by Proposition 2.3 in the case of infinite matrices. A similar proof of uniqueness also works for reversed GCF of finite matrices, but fails in the case of infinite matrices because Proposition 2.3 does not guarantee the associativity of the required operations. To show this, assume that

$$A_1^+ E_1 A_1 = A_2^+ E_2 A_2, \quad A_i \in \mathcal{T}, \quad E_i \in \mathcal{S},$$

where A_i are band matrices so that the products are admissible. Proposition 2.3 allows us to deduce that

$$A_1^+ E_1 = (A_1^+ E_1 A_1) A_1^{-1} = (A_2^+ E_2 A_2) A_1^{-1} = A_2^+ (E_2 A_2 A_1^{-1}).$$

However, when trying to move to the left the factor A_2^+ we only can conclude that

$$(A_2^+)^{-1} A_1^+ E_1 = (A_2^+)^{-1} (A_1^+ E_1) = (A_2^+)^{-1} A_2^+ (E_2 A_2 A_1^{-1}),$$

but the right-hand side cannot be further simplified using Proposition 2.3 because $(A_2^+)^{-1}$ is upper triangular and A_1^{-1} is lower triangular. Thus, the lack of associativity for products of infinite matrices is the algebraic reason for the lack of uniqueness of reversed GCF. This problem does not appear for finite matrices, whose reversed GCF are unique.

If \mathcal{C} is a CMV matrix and ℓ is a Hermitian Laurent polynomial with N_z zeros, then $\ell(\mathcal{C})$ is a Hermitian $2N_z + 1$ -diagonal matrix with bounded entries due to the unitarity of \mathcal{C} . Besides, the lower and upper bands of $\ell(\mathcal{C})$ have no null entries, as follows from the zig-zag shape (9) of \mathcal{C} . Therefore, all the previous results apply to the matrices $\ell(\mathcal{C})$ whose CF and reversed CF are involved in Darboux for CMV.

Acknowledgements

The work of the first, third and fourth authors has been supported in part by the research project MTM2011-28952-C02-01 from the Ministry of Science and Innovation of Spain and the European Regional Development Fund (ERDF), and by Project E-64 of Diputación General de Aragón (Spain).

The work of the second author has been partially supported by the research project MTM2012-36732-C03-01 from Dirección General de Investigación Científica y Técnica, Ministerio de Economía y Competitividad of Spain.

References

- [1] Aharonov, Y., Davidovich, L., Zagury, N., Quantum random walks, *Phys. Rev. A* **48** (1993), no. 2, 1687–1690.
- [2] Álvarez-Fernández, C., Mañas, M., Orthogonal Laurent polynomials on the unit circle, extended CMV ordering and 2D Toda type integrable hierarchies, *Adv. Math.* **240** (2013), 132–193.
- [3] Ammar, G. S., Gragg, W. B., Schur flows for orthogonal Hessenberg matrices, Hamiltonian and gradient flows, algorithms and control, pp. 27–34, *Fields Inst. Commun.* **3**, Amer. Math. Soc., Providence, R. I., 1994.
- [4] Ammar, G. S., Gragg, W. B., Reichel L., On the eigenproblem for orthogonal matrices, *Proc. 25th Conference on Decision and Control (Athens, 1986)*, pp. 1963–1966, New York, 1986.
- [5] Bueno, M. I., Marcellán, F., Darboux transformations and perturbations of linear functionals, *Linear Algebra Appl.* **384** (2004), 215–242.
- [6] Bueno, M. I., Marcellán, F., Polynomial perturbations of bilinear functionals and Hessenberg matrices, *Linear Algebra Appl.* **414** (2006), 64–83.
- [7] Buhmann, M. D., Iserles, A., On orthogonal polynomials transformed by the QR algorithm, *J. Comput. Appl. Math.* **43** (1992), no. 1-2, 117–134.
- [8] Bunse-Gerstner, A., Elsner, L., Schur parameters pencils for the solution of the unitary eigenproblem, *Linear Algebra Appl.* **154/156** (1991), 741–778.
- [9] Cantero, M. J., Grünbaum, F. A., Moral, L., Velázquez, L., Matrix-valued Szegő polynomials and quantum random walks, *Comm. Pure Appl. Math.* **63** (2010), no. 4, 464–507.
- [10] Cantero M. J., Marcellán F., Moral L., Velázquez L., in preparation.
- [11] Cantero, M. J., Moral, L., Velázquez, L., Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, *Linear Algebra Appl.* **362** (2003), 29–56.

- [12] Cantero, M. J., Moral, L., Velázquez, L., Minimal representations of unitary operators and orthogonal polynomials on the unit circle, *Linear Algebra Appl.* **408** (2005), 40–65.
- [13] Castillo, K., Garza, L., Marcellán, F., Laurent polynomials perturbations of linear functionals. An inverse problem, *Electron. Trans. Numer. Anal.* **36** (2009/10), 83–98.
- [14] Cooke, R. G., *Infinite matrices and sequence spaces*, reprinted edition, Dover Publ., New York, 2014.
- [15] Darboux G., Sur la representation sphériques de surfaces, *Comp. Hebd. Rend. Acad. Sci.* **94** (1882), 1343–1345.
- [16] Darboux, G., Sur une proposition relative aux équations linéaires, *Comp. Hebd. Rend. Acad. Sci.* **94** (1882), 1456–1459.
- [17] Daruis, L., Hernández, J., Marcellán, F., Spectral transformations for Hermitian Toeplitz matrices, *J. Comput. Appl. Math.* **202** (2007), no. 2, 155–176.
- [18] Delsarte, P., Genin, Y., The split Levinson algorithm, *IEEE Trans. Acoust. Speech Signal Process.* **34** (1986), no. 3, 470–478.
- [19] Denisov, S. A., Continuous analogs of polynomials orthogonal on the unit circle and Kreĭn systems, *IMRS Int. Math. Res. Surv.* 2006, Art. ID 54517, 148 pp.
- [20] Derevyagin, M., García-Ardila, J. C., Marcellán, F., Multiple Geronimus transformations, *Linear Algebra Appl.* **454** (2014), 158–183.
- [21] Derevyagin, M., Marcellán, F., A note on the Geronimus transformation and Sobolev orthogonal polynomials, *Numer. Algorithms* **67** (2014), no. 2, 271–287.
- [22] Derevyagin, M., Vinet, L., Zhedanov, A., CMV matrices and Little and Big -1 Jacobi Polynomials, *Constr. Approx.* **36** (2012), no. 3, 513–535.
- [23] Dirac, P. A. M., *The Principles of Quantum Mechanics*, second edition, Oxford, Clarendon Press, 1935.
- [24] Elhay, S., Kautsky, J., Jacobi matrices for measures modified by a rational factor, *Numer. Algorithms* **6** (1994), no. 3-4, 205–227.
- [25] Faybusovich, L., Gekhtman, M., On Schur flows, *J. Phys. A* **32** (1999), no. 25, 4671–4680.
- [26] Fernández D. J., Mielnik B., Rosas-Ortiz O., Samsonov B., Nonlocal supersymmetric deformations of periodic potentials, *J. Phys. A: Math. Gen.* **35** (2002), no. 19, 4279–4291.

- [27] Galant, D., An implementation of Christoffel's theorem in the theory of orthogonal polynomials, *Math. Comp.* **25** (1971), 111–113.
- [28] Garza, L., Hernández, J., Marcellán, F., Orthogonal polynomials and measures on the unit circle. The Geronimus transformations, *J. Comput. Appl. Math.* **233** (2010), no. 5, 1220–1231.
- [29] Garza, L., Hernández, J., Marcellán, F., Spectral transformations of measures supported on the unit circle and the Szegő transformation, *Numer. Algorithms* **49** (2008), no. 1-4, 169–185.
- [30] Garza, L., Marcellán, F., Linear spectral transformations and Laurent polynomials, *Mediterr. J. Math.* **6** (2009), no. 3, 273–289.
- [31] Gautschi, W., An algorithmic implementation of the generalized Christoffel theorem, in *Numerical Integration* (G. Hämmerlin, ed.), *Internat. Ser. Numer. Math.* **57**, Birkhäuser, Basel, 1982, pp. 89–106.
- [32] Gautschi, W., The interplay between classical analysis and (numerical) linear algebra—a tribute to Gene H. Golub, *Electron. Trans. Numer. Anal.* **13** (2002), 119–147.
- [33] Geronimus, Ya. L., Polynomials orthogonal on a circle and their applications, *Amer. Math. Soc. Translation* **1954** (1954), no. 104, 79 pp.
- [34] Golinskiĭ, L. B., Schur flows and orthogonal polynomials on the unit circle (Russian) *Mat. Sb.* **197** (2006), no. 8, 41–62; translation in *Sb. Math.* **197** (2006), no. 7-8, 1145–1165.
- [35] Golub, G. H., Kautsky, J., Calculation of Gauss quadratures with multiple free and fixed knots, *Numer. Math.* **41** (1983), no. 2, 147–163.
- [36] Gragg, W. B., The QR algorithm for unitary Hessenberg matrices, *J. Comput. Appl. Math.* **16** (1986), no. 1, 1–8.
- [37] Grenander, U., Szegő, G., *Toeplitz forms and their applications*, second edition, Chelsea Publishing Co., New York, 1984.
- [38] Grünbaum, F. A., The Darboux process and a noncommutative bispectral problem: some explorations and challenges, *Geometric aspects of analysis and mechanics*, pp. 161–177, *Progr. Math.*, 292, Birkhäuser/Springer, New York, 2011.
- [39] Grünbaum, F. A., Haine, L., Orthogonal polynomials satisfying differential equations: the role of the Darboux transformation, *Symmetries and integrability of difference equations* (Estrel, PQ, 1994), pp. 143–154, *CRM Proc. Lecture Notes* **9**, Amer. Math. Soc., Providence, R. I., 1996.

- [40] Grünbaum, F. A., Haine, L., Bispectral Darboux transformations: an extension of the Krall polynomials, *Internat. Math. Res. Notices* 1997, no. 8, 359–392.
- [41] Grünbaum, F. A., Haine, L., Horozov, E., Some functions that generalize the Krall-Laguerre polynomials, *J. Comput. Appl. Math.* **106** (1999), no. 2, 271–297.
- [42] Harnard, J., Kasman, A. (editors), *The bispectral problem* (Montreal, PQ, 1997), CRM Proc. Lecture Notes, 14, Amer. Math. Soc., Providence, R. I., 1998.
- [43] Humet, M., Van Barel, M., Algorithms for the Geronimus transformation for orthogonal polynomials on the unit circle, *J. Comput. Appl. Math.* **267** (2014), 195–217.
- [44] Infeld, L., On a new treatment of some eigenvalue problems, *Phys. Rev.* **59** (1941), no. 9, 737–747.
- [45] Infeld, L., Hull, T. E., The factorization method, hydrogen intensities and related problems, *Phys. Rev.* **74** (1948), no. 8, 905–909.
- [46] Infeld, L., Hull, T. E., The factorization method, *Rev. Modern Physics* **23** (1951), no. 1, 21–68.
- [47] Ismail M. E. H., Li, X., On sieved orthogonal polynomials IX: Orthogonality on the unit circle, *Pacific J. Math.* **153** (1992), no. 2, 289–297.
- [48] Kato, T., *Perturbation Theory for Linear Operators*, second edition, Grundlehren der Mathematischen Wissenschaften, Band 132, Springer-Verlag, Berlin-New York, 1976.
- [49] Kautsky, J., Golub, G. H., On the calculation of Jacobi matrices, *Linear Algebra Appl.* **52/53** (1983), 439–455.
- [50] Killip, R., Nenciu, I., Matrix models for circular ensembles, *Int. Math. Res. Not.* (2004), no. 50, 2665–2701.
- [51] Killip, R., Nenciu, I., CMV: The unitary analogue of Jacobi matrices, *Comm. Pure Appl. Math.* **60** (2007), no. 8, 1148–1188.
- [52] Kubrusly, C. S., *The Elements of Operator Theory*, second edition, Birkhäuser/Springer, New York, 2011.
- [53] Li, L.-C., Nenciu, I., The periodic defocusing Ablowitz-Ladik equation and the geometry of Floquet CMV matrices, *Adv. Math.* **231** (2012), no. 6, 3330–3388.
- [54] Maddox, I. J., *Infinite matrices of operators*, Lecture Notes in Mathematics **786**, Springer, Berlin, 1980.
- [55] Marcellán F., Hernández, J., Christoffel transforms and Hermitian linear functionals, *Mediterr. J. Math.* **2** (2005), no. 4, 451–458.

- [56] Marcellán, F., Sansigre, G., Orthogonal polynomials on the unit circle: Symmetrization and quadratic decomposition, *J. Approx. Theory* **65** (1991), no. 1, 109–119.
- [57] Matveev, V. B., Salle, M. A., Darboux transformations and solitons, Springer Series in Nonlinear Dynamics, Springer-Verlag, Berlin, 1991.
- [58] Meyer, D. A., From quantum cellular automata to quantum lattice gases, *J. Stat. Phys.* **85** (1996), no. 5-6, 551–574.
- [59] Mielnik, B., Rosas-Ortiz, O., Factorization: little or great algorithm?, *J. Phys. A* **37** (2004), no. 43, 10007–10035.
- [60] Moutard, T., Sur la construction des équations de la forme $\frac{1}{z} \frac{d^2 z}{dx dy} = \lambda(x, y)$, qui admettent une intégrale générale explicite, *J. de l'Ecole Poly.* **45** (1878), 1–11.
- [61] Mukaihira, A., Nakamura, Y., Schur flow for orthogonal polynomials on the unit circle and its integrable discretization, *J. Comput. Appl. Math.* **139** (2002), no. 1, 75–94.
- [62] Nenciu, I., Lax pairs for the Ablowitz-Ladik system via orthogonal polynomials on the unit circle, *Int. Math. Res. Not.* 2005, no. 11, 647–686.
- [63] Nenciu, I., CMV matrices in random matrix theory and integrable systems: a survey, *J. Phys. A* **39** (2006), no. 28, 8811–8822.
- [64] Schrödinger, E., A method of determining quantum-mechanical eigenvalues and eigenfunctions, *Proc. Roy. Irish Acad. Sect. A* **46** (1940), 9–16.
- [65] Schrödinger, E., Further studies on solving eigenvalue problems by factorization, *Proc. Roy. Irish Acad. Sect. A* **46** (1941), 183–206.
- [66] Shivakumar, P. N., Sivakumar, K. C., A review of infinite matrices and their applications, *Linear Algebra Appl.* **430** (2009), no. 4, 976–998.
- [67] Simon, B., Orthogonal polynomials on the unit circle. Part 1 and 2, *Amer. Math. Soc. Colloq. Publ. Series* **54**, Vol. 1 and 2, Amer. Math. Soc., Providence, R. I., 2005.
- [68] Simon, B., OPUC on one foot, *Bull. Amer. Math Soc.* **42** (2005), no. 4, 431–460.
- [69] Simon, B., CMV matrices: Five years after, *J. Comput. Appl. Math.* **208** (2007), no. 1, 120–154.
- [70] M.H. Stone, *Linear Transformations in Hilbert Spaces and Their Applications to Analysis*, Amer. Math. Soc. Colloq. Publ. **15**, Amer. Math. Soc., New York, 1932.

- [71] Teschl, G., Jacobi operators and completely integrable nonlinear lattices, *Mathematical Surveys and Monographs* **72**, Amer. Math. Soc., Providence, R. I., 2000.
- [72] Wall, S. H., *Analytic Theory of Continued Fractions*, D. Van Nostrand Company, Inc., New York, N. Y., 1948.
- [73] Wang, T.-L., Gragg, W. B., Convergence of the shifted QR algorithm for unitary Hessenberg matrices, *Math. Comp.* **71** (2002), no. 240, 1473–1496.
- [74] Watkins, D. S., Some perspectives on the eigenvalue problem, *SIAM Rev.* **35** (1993), no. 3, 430–471.
- [75] Weidmann, J., *Linear Operators in Hilbert Spaces*, Graduate Texts in Mathematics **68**, Springer-Verlag, New York-Berlin, 1980.
- [76] Williamson, J., Quasi-unitary matrices, *Duke Math. J.* **3** (1937), no. 4, 715–725.
- [77] Yoon, G. J., Darboux transforms and orthogonal polynomials, *Bull. Korean Math. Soc.* **39** (2002), no. 3, 359–376.
- [78] Zhedanov, A., Rational spectral transformations and orthogonal polynomials, *J. Comput. Appl. Math.* **85** (1997), no. 1, 67–86.