Peakons and Cauchy Biorthogonal Polynomials

M. Bertola †‡ [1], M. Gekhtman [a] [2], J. Szmigielski [b] [3]

† *Centre de recherches mathématiques, Université de Montréal*
C. P. 6128, succ. centre ville, Montréal, Québec, Canada H3C 3J7
E-mail: bertola@crm.umontreal.ca

‡ *Department of Mathematics and Statistics, Concordia University*
7141 Sherbrooke W., Montréal, Québec, Canada H4B 1R6

[a] *Department of Mathematics 255 Hurley Hall, Notre Dame, IN 46556-4618, USA*
E-mail: Michael.Gekhtman.1@nd.edu

[b] *Department of Mathematics and Statistics, University of Saskatchewan*
106 Wiggins Road, Saskatoon, Saskatchewan, S7N 5E6, Canada

E-mail: szmigiel@math.usask.ca

Abstract

Peakons are non-smooth soliton solutions appearing in certain nonlinear partial differential equations, most notably the Camassa-Holm equation and the Degasperis-Procesi equation. In the latter case the construction of peakons leads to a new class of biorthogonal polynomials. The present paper is the first in the series of papers aimed to establish a general framework in which to study such polynomials. It is shown that they belong to a class of biorthogonal polynomials with respect to a pairing between two Hilbert spaces with measures $\mathrm{d}\alpha, \mathrm{d}\beta$ on the positive semi-axis $\mathbb{R}_+$ coupled through the the Cauchy kernel $K(x,y) = \frac{1}{x+y}$. Fundamental properties of these polynomials are proved: their zeros are interlaced, they satisfy four-term recurrence relations and generalized Christoffel-Darboux identities, they admit a characterization in terms of a 3 by 3 matrix Riemann-Hilbert problem. The relevance of these polynomials to a third order boundary value problem (the cubic string) is explained. Moreover a connection to certain two-matrix random matrix models, elaborated on in subsequent papers, is pointed out.

[1]Work supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), Grant. No. 261229-03 and by the Fonds FCAR du Québec No. 88353.

[2]Work supported in part by NSF Grant DMD-0400484.

[3]Work supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), Grant. No. 138591-04

Contents

1 Introduction

This paper deals with a class of biorthogonal polynomials $\{p_n(x)\}_{\mathbb{N}}, \{q_n(y)\}_{\mathbb{N}}$ of exact degree n satisfying the biorthogonality relations

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} p_n(x)q_m(y)\frac{\mathrm{d}\alpha(x)\mathrm{d}\beta(y)}{x+y} = \delta_{mn}, \tag{1-1}$$

where $\mathrm{d}\alpha, \mathrm{d}\beta$ are positive measures supported on $\mathbb{R}_+$ such that all the bimoments are finite.

We present two main reasons why such polynomials are of interest: one source of interest is the weakly dispersive equation

$$u_t - u_{xxt} + 3uu_x = 2u_x u_{xx} + uu_{xxx}, \quad (x,t) \in \mathbb{R}^2, \tag{1-2}$$

which was proposed in the early 1990s by Camassa and Holm [1] as a model shallow water wave equation. This equation admits weak solutions of the form:

$$u(x,t) = \sum_{i=1}^{n} m_i(t)\, e^{-|x-x_i(t)|}, \tag{1-3}$$

where the positions $x_i(t)$ and the heights $m_i(t)$ are determined by the system of nonlinear ODEs:

$$\begin{aligned} \dot{x}_k &= \sum_{i=1}^{n} m_i e^{-|x_k-x_i|}, \\ \dot{m}_k &= \sum_{i=1}^{n} m_k m_i \,\mathrm{sgn}(x_k - x_i)\, e^{-|x_k-x_i|}, \end{aligned} \tag{1-4}$$

for $k = 1, \dots, n$. On account of the non-smooth character, and the presence of sharp peaks at $\{x_k\}$, these solutions were named *peakons*. The peakon solutions to the CH equation were subsequently constructed using inverse scattering techniques by Beals, Sattinger and Szmigielski [2, 3]. In particular, it was shown in these works that rapidly decaying at large x solutions to the CH equation can be constructed by solving the inverse problem for an inhomogeneous string equation:

$$-\phi''(\xi,z) = zm^*(\xi)\phi(\xi,z), \qquad -1 < \xi < 1, z \in \mathbb{C}. \tag{1-5}$$

In this equation, ξ is related to the spacial variable x by $\xi = \tanh(x)$ and $m^*(\xi)$ is simply related to $m(x) = u - u_{xx}$. In the peakon case, $m = \sum_{i=1}^{n} m_i \delta_{x_i}$, where δ_{x_i} is the Dirac measure concentrated at the position of the peak x_i, and $m^* = \sum_{i=1}^{n} m_i^* \delta_{x_i^*}$ respectively [2]. The connection to orthogonal polynomials became transparent once it was realized that the Weyl function for the problem (1-5),

$$W(z) = \frac{\phi'(1,z)}{z\phi(1,z)},$$

admits a continued fraction expansion of Stieltjes' type:

$$W(z) = \cfrac{1}{l_n^* z + \cfrac{1}{m_n^* + \cfrac{1}{l_{n-1}^* z + \cdots}}}, \tag{1-6}$$

where $l_j^* = x_{j+1}^* - x_j^*$ is the distance between peaks in the ξ coordinate. In this way the peakon solutions became intimately linked to the origins of the modern theory of orthogonal polynomials [4]. One of the most immediate applications of orthogonal polynomials to peakons was the proof of the absence of the so-called *triple collisions* [3], roughly stating that the peaks x_j can only coalesce in pairs. The essential part of the argument could be traced back to the fact that orthogonal polynomials satisfy three-term recurrence relations. Yet, the CH equation can be viewed as belonging to a family of PDEs

$$u_t - u_{xxt} + (b+1)uu_x = bu_x u_{xx} + uu_{xxx}, \quad (x,t) \in \mathbb{R}^2, \tag{1-7}$$

for which the peakon ansatz (1-3) leads to a similar system of ODEs as (1-4):

$$\begin{aligned} \dot{x}_k &= \sum_{i=1}^n m_i e^{-|x_k - x_i|}, \\ \dot{m}_k &= (b-1) \sum_{i=1}^n m_k m_i \operatorname{sgn}(x_k - x_i)\, e^{-|x_k - x_i|}. \end{aligned} \tag{1-8}$$

Two cases of (1-7) are known to be integrable. In addition to the case $b = 2$, namely the original CH equation, the case $b = 3$ is also known to be integrable. The latter equation was discovered by Degasperis and Procesi [5] and was later shown to be integrable by Degasperis, Holm and Hone [6, 7]. We will refer to the case $b = 3$ of (1-7) as the DP equation. The construction of the DP peakons was given by Lundmark and Szmigielski, first in a short note [8], then a complete construction in the longer paper [9]. The main steps of this construction, presented with an emphasis on its connection to biorthogonal polynomials, is the subject of Section 3.

The other motivation comes from random matrix theory: while this topic will be dealt with in depth in the subsequent papers, we indicate the main reasons behind this aspect of our interest in biorthogonal polynomials.

It is well known [10] that the Hermitean matrix model is intimately related to (in fact, solved by) orthogonal polynomials (OPs). Not so much is known about the role of biorthogonal polynomials (BOPs). However, certain biorthogonal polynomials somewhat similar to the ones in the the present paper appear prominently in the analysis of "the" two–matrix model after reduction to the spectrum of eigenvalues [11, 12, 13, 14]; in that case the pairing is of the form

$$\int\int p_n(x) q_m(y) \mathrm{e}^{-xy} \mathrm{d}\alpha(x) \mathrm{d}\beta(y) = \delta_{mn}. \tag{1-9}$$

We refer to these BOPs as the "Itzykson–Zuber BOPs" (IZ-BOPs) due to the relevance of the Itzykson–Zuber–Harish-Chandra formula for the matrix model they derive from. Several algebraic structural properties of these polynomials and their recurrence relation (both multiplicative and differential) have been thoroughly analyzed in the previously cited papers for densities of the form $\mathrm{d}\alpha(x) = \mathrm{e}^{-V_1(x)}\mathrm{d}x,\ \mathrm{d}\beta(y) = \mathrm{e}^{-V_2(y)}\mathrm{d}y$ for *polynomials potentials* $V_1(x),\ V_2(y)$ and for potentials with rational derivative (and hard–edges) in [15].

We recall that while ordinary OPs satisfy a multiplicative three–term recurrence relation, the BOPs defined by (1-9) solve a longer recurrence relation of length related to the degree of the differential $\mathrm{d}V_j(x)$ over the Riemann sphere [15]; a direct (although not immediate) consequence of the finiteness of the recurrence relation is the fact that these BOPs (and certain integral transforms of them) are characterized by a Riemann–Hilbert problem for a matrix of size equal to the length of the recurrence relation (minus one). The BOPs we deal with in this paper share all these features, although in some respects they are closer to the ordinary orthogonal polynomials than to the IZBOPs.

We now list the main properties of biorthogonal polynomials studied in this paper:

- they are linked to the spectral and inverse spectral problem for the **cubic string**, similar to the ordinary polynomials being linked to the theory of an inhomogeneous string of M.G. Krein [16]
- they solve a **four–term** recurrence relation for *any* pair of measures $\mathrm{d}\alpha, \mathrm{d}\beta$ as specified after (1-1);
- they have **positive and simple** zeroes;
- the **zeroes** of $p_n(x)$ $(q_n(y))$ **are interlaced** with the zeroes of the neighboring polynomials;
- they are characterized by a 3×3 Riemann–Hilbert problem;
- they satisfy interesting **Christoffel–Darboux identities** which pair them naturally with other sequences of polynomials which solve a **dual Riemann–Hilbert problem**.

In the first part of the paper which comprises sections 3.1 through 3.3 we carry out a detailed study of a discrete cubic string with a variety of boundary conditions and establish the main source of biorthogonality in the form of a generalization of the Parseval Identity (Theorem 3.4). This part of the paper is of interest *per se* and can be read independently of the remainder of the paper, even though it is conceptually important to understand the deeper reasons for the relevance of biorthogonality from the ODE point of view. This part of the paper addresses the first item on the list above, but it also motivates many concepts introduced later.

The second part of this paper, starting with section 4.1 onward, is the detailed analysis of the remaining points from the list.

In the follow-up paper we will explain the relation of the BOPs introduced in this paper with

- a new two–matrix model with its relevant diagrammatic expansion for large size [17];
- a rigorous asymptotic analysis for continuous (varying) measures $d\alpha, d\beta$ using the nonlinear steepest descent method [17];
- generalizations covering a multi-matrix model [18].

1.1 Relationship with random matrix models

As a preview of the forthcoming papers we would like to point out the relevant two–matrix model our polynomials are related to.

Consider the set of pairs $\mathcal{H}_+^{(2)} := \{(M_1, M_2)\}$ of Hermitean *positive-definite* matrices; it is a cone in the direct sum of the vector spaces of Hermitean matrices, endowed with the ($U(N)$–invariant) Lebesgue measure, which we short-handedly denote by $dM_1 dM_2$. Consider the following positive measure on this space

$$d\mu(M_1, M_2) = \frac{1}{\mathcal{Z}_N^{(2)}} \frac{\alpha'(M_1)\beta'(M_2) dM_1 dM_2}{\det(M_1 + M_2)^N} \tag{1-10}$$

where $\mathcal{Z}_N^{(2)}$ (the *partition function* is a normalization constant crafted so as to have a unit total mass). As a result the measure space $(\mathcal{H}_+^{(2)}, d\mu)$ becomes a probability space, and the matrices M_1, M_2 are *random matrices*. The notation $\alpha'(M_1), \beta'(M_2)$ stands for the product of the densities α', β' (the Radon–Nikodym derivatives of the measures $d\alpha, d\beta$ with respect to the Lebesgue measure) over the (positive) eigenvalues of M_j.

This probability space is similar to the so–called two–matrix model, where the coupling between matrices instead of $\det(M_1 + M_2)^{-N}$ is $e^{N \mathrm{Tr} M_1 M_2}$ [19]. The connection with our BOPs (1-1) is precisely on the same footing as the connection between ordinary orthogonal polynomials and the Hermitean Random matrix model [10], namely the probability space over $\mathcal{H}_N$ given by the measure

$$d\mu_1(M) := \frac{1}{\mathcal{Z}_N^{(1)}} \alpha'(M) dM \ . \tag{1-11}$$

In particular we will show in the forthcoming paper how the statistics of the eigenvalues of the two matrices M_j can be described in terms of the biorthogonal polynomials we are introducing in the present work. A prominent role in the description of said statistics will be played by the Christoffel–Darboux identities that we develop in Section 6. Finally, it remains an open question as to what the nature of the precise connection between what we are proposing and the work of Mark Adler and Pierre van Moerbeke (e.g. [20]) is, and that item certainly merits further investigation.

2 List of symbols and notations commonly used

- Given variables $x_1, \ldots, x_n$ we will denote them collectively as X, regardless of what the labeling set is.
- S_n denotes the group of permutations of n elements.
- Given a permutation $\sigma \in S_n$ we will denote by X_σ the permutation of the indices of the variables.
- Given a permutation σ we denote by $\epsilon(\sigma)$ its sign.
- $C(X) := \prod_{j=1}^{n} x_j$
- $\Delta(X) := \prod_{i>j}(x_i - x_j) = \det[x_i^{j-1}]_{i,j\le n}$
- Given a square matrix A we denote by $\widetilde{A}$ the matrix of its cofactors (the adjoint matrix).
- α_j, β_j = moments of the measures $\mathrm{d}\alpha, \mathrm{d}\beta$.
- $K(x,y)$= totally positive kernel.
- $\langle x^k | y^\ell \rangle = I_{k,\ell} = \iint x^i y^\ell \mathrm{d}\alpha(x)\mathrm{d}\beta(y) K(x,y)$ bimoments associated to the kernel $K(x,y)$ and the measures $\mathrm{d}\alpha, \mathrm{d}\beta$.
- $I = [I_{k,\ell}]$=matrix of bimoments.
- $D_n := \det[I_{k,\ell}]_{k,\ell=0,\ldots,n-1}$, principal minors of the bimoment matrix.
- Λ = semiinfinite upper shift matrix.
- $p_n(x), q_n(y)$, biorthogonal polynomials satisfying $\langle p_n | q_m \rangle = \delta_{mn}$.
- Given any sequence $\{x_j\}$ we will denote by $\mathbf{x}$ the column vector $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix}$, while $\mathbf{x}^T$ will denote the corresponding row vector.
- $\boldsymbol{\pi} = \int \mathbf{p}\mathrm{d}\alpha$, $\boldsymbol{\eta} = \int \mathbf{q}\mathrm{d}\beta$.

3 Degasperis-Procesi equation and a cubic string

3.1 DP equation

We begin by summarizing basic facts about the DP peakons. The Degasperis-Procesi (DP) equation:

$$u_t - u_{xxt} + 4uu_x = 3u_x u_{xx} + uu_{xxx}, \qquad (x,t) \in \mathbb{R}^2, \tag{3-1}$$

admits weak n-peakon solutions. They are obtained by substituting the peakon ansatz (1-3) into (3-1). Then careful analysis shows that the solution exists in a weak sense if and only if:

$$\begin{aligned} \dot{x}_k &= \sum_{i=1}^{n} m_i e^{-|x_k - x_i|}, \\ \dot{m}_k &= 2\sum_{i=1}^{n} m_k m_i \operatorname{sgn}(x_k - x_i)\, e^{-|x_k - x_i|}. \end{aligned} \tag{3-2}$$

This system of ODEs can be successfully analyzed with the help of another crucial ingredient: the DP equation admits a Lax formulation, first proposed in [6]. Thus the DP equation follows from the compatibility condition for the the system

$$(\partial_x - \partial_x^3)\psi = z\, m\psi, \tag{3-3a}$$

$$\partial_t \psi = \left[z^{-1}(1-\partial_x^2) + u_x - u\partial_x\right]\psi. \tag{3-3b}$$

where $z \in \mathbb{C}$. In the case of peakons, $m = \sum_{i=1}^n m_i \delta_{x_i}$, and, as a result, the equation (3-3) is assumed to hold in the sense of distributions. Furthermore, (3-3) can be solved quite explicitly [9]. To this end it is useful to perform a Liouville transformation on (3-3). This is fully explained in [9] and here we only need the main thread of the reasoning that results in the appearance of the cubic string boundary value problem.

Lemma 3.1. *Under the change of variables*

$$\xi = \tanh\frac{x}{2}, \qquad \psi(x) = \frac{2\,\phi(\xi)}{1-\xi^2}, \tag{3-4}$$

the DP spectral problem (3-3a) *is equivalent to the cubic string problem*

$$\begin{aligned} -\phi_{\xi\xi\xi}(\xi) &= z\, g(\xi)\,\phi(\xi) \quad \textit{for } \xi \in (-1,1), \\ \phi(-1) &= \phi_\xi(-1) = 0, \\ \phi(1) &= 0, \end{aligned} \tag{3-5}$$

where

$$\left(\frac{1-\xi^2}{2}\right)^3 g(\xi) = m(x). \tag{3-6}$$

In the discrete case, when $m(x) = 2\sum_1^n m_i\,\delta_{x_i}$, *equation* (3-6) *should be interpreted as*

$$g(\xi) = \sum_{i=1}^{n} g_i\,\delta_{\xi_i}, \quad \textit{where} \quad \xi_i = \tanh\frac{x_i}{2}, \quad g_i = \frac{8m_i}{\left(1-\xi_i^2\right)^2}. \tag{3-7}$$

Remark 3.1. *The specific boundary conditions mentioned in the lemma have been chosen to deal with the peakon problem. We will choose later a different set of boundary conditions to reflect the focus of this paper (see 3.2)*

3.2 Discrete cubic string

We slighly generalize the cubic string discussed in the previous section in connection with the peakon problem. In an ordinary string problem different boundary conditions correspond to different ways of tying down the ends of the string. For us, different boundary conditions will eventually lead to different spectral measures with respect to which we will define the biorthogonal polynomials.

Definition 3.1. *We define the cubic string boundary value problems (BVP) of three types*

$$\begin{aligned}
&-f_{\xi\xi\xi}(\xi) = z\,g(\xi)\,f(\xi) \quad \textit{for } \xi \in (0,1),\\
&f(0) = f_\xi(0) = 0,\\
&\textit{Type 0 (Peakon case)}: f(1) = 0, \quad \textit{Type 1}: f_\xi(1) = 0 \quad \textit{Type 2}: f_{\xi\xi}(1) = 0
\end{aligned} \tag{3-8}$$

Remark 3.2. *For simplicity we have adjusted the length of the string; it is now* 1 *rather than* 2.

We are only interested in the case where the mass distribution consists of a finite collection of point-masses:

$$g(\xi) = \sum_{i=1}^{n} g_i\,\delta_{\xi_i}, \quad \text{where} \quad 0 < g_i, \quad 0 < \xi_1 < \xi_2 < \cdots < \xi_n < 1. \tag{3-9}$$

We will consider all three boundary value problems mentioned above with this mass distribution, as well as one degenerate case in which the last mass is placed at 1 (i.e. $\xi_n = 1$: in the latter case we take the right hand limit to compute the derivatives of f at 1. Moreover, for that case, we consider only the BVP of type 2.)

We will collectively refer to all these cases as *the discrete cubic string.*

We will also use an accompanying initial value problem, which is the same for all three types.

Definition 3.2. *The cubic string initial value problem (IVP) is defined by the following equations*

$$\begin{aligned}
&-\phi_{\xi\xi\xi}(\xi) = z\,g(\xi)\,\phi(\xi) \quad \textit{for } \xi \in (0,1),\\
&\phi(0) = \phi_\xi(0) = 0, \phi_{\xi\xi}(0) = 1
\end{aligned} \tag{3-10}$$

The boundary value problems in Definition 3.1 are not self-adjoint and the adjoint boundary value problems play an important role.

Definition 3.3. *The adjoint cubic string boundary value problems are described by the following relations*

$$f^*_{\xi\xi\xi}(\xi) = z\, g(\xi)\, f^*(\xi) \quad \textit{for } \xi \in (0,1) \tag{3-11}$$

1. *Type 0*

$$f^*(0) = 0,$$
$$f^*(1) = f^*_\xi(1) = 0,$$

2. *Type 1*

$$f^*(0) = 0,$$
$$f^*(1) = f^*_{\xi\xi}(1) = 0,$$

3. *Type 2*

$$f^*(0) = 0,$$
$$f^*_\xi(1) = f^*_{\xi\xi}(1) = 0.$$

The corresponding initial value problem is:

Definition 3.4. *The adjoint cubic string initial value problems are gives as*

$$\begin{aligned} \phi^*_{\xi\xi\xi}(\xi) = z\, g(\xi)\, \phi^*(\xi) \quad \textit{for } \xi \in (0,1), \\ \textit{with nonzero initial values as follows} \\ \textit{Type 0: } \phi^*_{\xi\xi}(1) = 1, \textit{ Type 1: } \phi^*_\xi(1) = 1, \textit{ Type 2: } \phi^*(1) = 1 \end{aligned} \tag{3-12}$$

To avoid cluttering the notation we will use the same symbol ϕ^* in all three cases whenever the context clearly identifies one of the three adjoint boundary/initial value problems. When necessary, we will attach an index $a = 0, 1, 2$ referring to the type, for example, ϕ^*_0 will refer to the BVP/IVP of type 0 etc.

In the process of integration by parts one identifies the relevant bilinear symmetric form:

Definition 3.5. *Given any twice differentiable f, h the bilinear concomitant is defined as the bilinear form:*

$$B(f,h)(\xi) = f_{\xi\xi}h - f_\xi h_\xi + f h_{\xi\xi}. \tag{3-13}$$

This bilinear symmetric form induces a bilinear symmetric form (denoted also by B) on triples $F^T := (f, f_\xi, f_{\xi\xi})$, namely

$$B(F,H) = F^T J H, \quad J := \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{3-14}$$

We also define a natural L^2 space associated with g, denoted $L^2[0,1]_g$, equipped with the inner product: $(f,h)_g = \int_0^1 f(\xi)h(\xi)g(\xi)\,d\xi$. Since all initial value problems 3.2, 3.4 can be solved for arbitrary $z \in \mathbb{C}$, ϕ and ϕ^* are functions of the spectral parameter z. The following theorem establishes a relation between these two functions and the relevant boundary value problems.

Lemma 3.2. *Suppose $\phi(\xi; z)$ and $\phi^*(\xi; \lambda)$ are solutions to the IVPs 3.2, 3.4 with spectral parameters z and λ.*

1. *Type 0: the spectrum is determined by the zeros of $\phi(1,z) = 0$. Moreover,*

$$-B(\phi(\xi;z), \phi^*(\xi,\lambda)|_0^1 = \phi^*(0;\lambda) - \phi(1;z) = (z-\lambda)(\phi(\bullet;z), \phi^*(\bullet;\lambda))_g. \tag{3-15}$$

2. *Type 1: the spectrum is determined by the zeros of $\phi_\xi(1,z) = 0$. Moreover,*

$$-B(\phi(\xi;z), \phi^*(\xi,\lambda)|_0^1 = \phi^*(0;\lambda) + \phi_\xi(1;z) = (z-\lambda)(\phi(\bullet;z), \phi^*(\bullet;\lambda))_g. \tag{3-16}$$

3. *Type 2: the spectrum is determined by the zeros of $\phi_{\xi\xi}(1,z) = 0$. Moreover,*

$$-B(\phi(\xi;z), \phi^*(\xi,\lambda)|_0^1 = \phi^*(0;\lambda) - \phi_{\xi\xi}(1;z) = (z-\lambda)(\phi(\bullet;z), \phi^*(\bullet;\lambda))_g. \tag{3-17}$$

In addition, in all three cases,

$$-B(\phi(\xi;z), \phi(\xi;\lambda)|_0^1 = (z+\lambda)(\phi(\bullet;z), \phi(\bullet;\lambda))_g. \tag{3-18}$$

Proof. Indeed (3-10) and two integrations by parts imply that

$$-\int_0^1 \phi_{\xi\xi\xi}(\xi;z)\phi^*(\xi;\lambda)\,d\xi = -B(\phi,\phi^*)|_0^1 + \int_0^1 \phi(\xi;z)\phi^*_{\xi\xi\xi}(\xi;\lambda)\,d\xi =$$

$$z\int_0^1 \phi(\xi;z)\phi^*(\xi;\lambda)g(\xi)\,d\xi.$$

Consequently, using equation (3-12) we obtain:

$$-B(\phi,\phi^*)|_0^1 = (z-\lambda)\int_0^1 \phi(\xi;z)\phi^*(\xi;\lambda)g(\xi)\,d\xi,$$

which in view of the initial conditions implies the claim. A similar computation works for the second identity. □

It is now easy to see that

Corollary 3.1. *ϕ and ϕ^* satisfy the following relations:*

1. *Case 0: $\phi^*(0;z) = \phi(1;z)$. Case 1: $\phi^*(0;z) = -\phi_\xi(1;z)$, Case 2: $\phi^*(0;z) = \phi_{\xi\xi}(1;z)$.*
2. *Case 0: $-\phi_z(1;z) = (\phi(\bullet;z), \phi^*(\bullet;z))_g$. Case 1: $\phi_{\xi z}(1;z) = (\phi(\bullet;z), \phi^*(\bullet;z))_g$. Case 2: $-\phi_{\xi\xi z}(1;z) = (\phi(\bullet;z), \phi^*(\bullet;z))_g$.*

We give below a complete characterization of the spectra and of the eigenfunctions for all three BVPs.

Theorem 3.1. *Consider a cubic string with a finite measure g as in (3-9).*

1. *Let $z_{a,j}$ denote the eigenvalues of the BVP of type $a = 0, 1, 2$. In each of the three cases, the spectrum is positive and simple.*
2. *For any pair of BVPs of type 0,1,2 the spectra are interlaced in the following order:*

$$z_{2,j} < z_{1,j} < z_{0,j}, \quad j = 1, \dots n$$

3. *The eigenfunctions $\phi(\xi; z_{a,j}) \equiv \phi_{a,j}(\xi)$ can be chosen to be real valued and they are linearly independent. Moreover, for the following combinations of BVPs, $(\phi_{a,i}, \phi_{b,j})_g$ factorizes as:*
   (a) Type 00:
   $$(\phi_{0,i}, \phi_{0,j})_g = \frac{\phi_{0,i,\xi}(1)\phi_{0,j,\xi}(1)}{z_{0,i} + z_{0,j}} \tag{3-19}$$
   (b) Type 01
   $$(\phi_{0,i}, \phi_{1,j})_g = -\frac{\phi_{0,i,\xi\xi}(1)\phi_{1,j}(1)}{z_{0,i} + z_{1,j}} \tag{3-20}$$
   (c) Type 12
   $$(\phi_{1,i}, \phi_{2,j})_g = -\frac{\phi_{1,i,\xi\xi}(1)\phi_{2,j}(1)}{z_{1,i} + z_{2,j}} \tag{3-21}$$
   (d) Type 22
   $$(\phi_{2,i}, \phi_{2,j})_g = \frac{\phi_{2,i,\xi}(1)\phi_{2,j,\xi}(1)}{z_{2,i} + z_{2,j}} \tag{3-22}$$

Proof. It is easy to check (see Section 4.1 in [9]) that

$$\begin{pmatrix} \phi(1;z) \\ \phi_\xi(1;z) \\ \phi_{\xi\xi}(1;z) \end{pmatrix} = L_n G_n(z)\, L_{n-1}\, G_{n-1}(z) \cdots L_1\, G_1(z)\, L_0 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3-23}$$

where

$$G_k(z) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -z\, g_k & 0 & 1 \end{pmatrix},$$

$$L_k = \begin{pmatrix} 1 & l_k & l_k^2/2 \\ 0 & 1 & l_k \\ 0 & 0 & 1 \end{pmatrix},$$

and

$$l_k = \xi_{k+1} - \xi_k, \qquad \xi_0 = 0, \qquad \xi_{n+1} = 1. \tag{3-24}$$

We prove the statement about the spectra by using the results obtained in [9]. By Theorem 3.5 in [9] $\phi(1; z)$ has n distinct positive zeros and so do $\phi_\xi(1; z)$ and $\phi_{\xi\xi}(1; z)$ (denoted there ϕ_y, ϕ_{yy}). Indeed this follows from observing that $\frac{\phi_\xi(1;z)}{\phi_z(1;z)}$ and $\frac{\phi_{\xi\xi}(1;z)}{\phi_z(1;z)}$ are strictly positive on the spectrum of Type 0, which implies that both $\phi_\xi(1; z)$ and $\phi_{\xi\xi}(1; z)$ change signs n times, hence all three spectra are simple. Furthermore, $\phi_\xi(1; -z) > 0$, $\phi_{\xi\xi}(1; -z) > 0$ for $z > 0$, so the zeros of $\phi_\xi(1; z)$ and $\phi_{\xi\xi}(1; z)$ are strictly positive and consequently they interlace with the zeros of $\phi(1; z)$. Thus the spectra of type 0 and 1, as well as 0 and 2 interlace. To see that the spectrum of type 1 interlaces with the spectrum of type 2 we proceed as follows. By (3-18), after evaluating at $z = z_{2,i}, \lambda = -z_{2,i}$, we obtain

$$\phi_\xi(1; z_{2,i})\phi_\xi(1; -z_{2,i}) = \phi_{\xi\xi}(1; -z_{2,_i})\phi(1; z_{2,i}),$$

which gives

$$\phi_\xi(1; z_{2,i}) = \frac{\phi_{\xi\xi}(1; -z_{2,i})}{\phi_\xi(1; -z_{2,i})}\phi(1; z_{2,i}) \tag{3-25}$$

For $z > 0$, $\mathrm{sgn}(\phi_\xi(1; -z)) = \mathrm{sgn}(\phi_{\xi\xi}(1; -z)) = +1$ because both are strictly positive there. Since the zeros of $\phi(1; z)$ interlace with the zeros of $\phi_{\xi\xi}(1; z)$, $\mathrm{sgn}(\phi(1; z_{2,i}))$ alternates, which in turn implies that $\mathrm{sgn}(\phi_\xi(1; z_{2,i}))$ alternates as well. Thus the proof that the zeros of $\phi_\xi(1; z)$ are simple and they interlace with the zeros of $\phi_{\xi\xi}(1; z)$ is complete.

The relative position of the spectra of the three types is best inferred from the fact that $\phi_\xi(1 : z_{0,1})$ and $\phi_{\xi\xi}(1 : z_{0,1})$ are both negative since $\frac{\phi_\xi(1,z)}{\phi_z(1,z)}$ and $\frac{\phi_{\xi\xi}(1,z)}{\phi_z(1,z)}$ are strictly positive on the spectrum of Type 0. Thus the spectra of type 1 and 2 are shifted to the left relative to the spectrum of type 0. In particular, $\phi(1, z_{2,1}) > 0$ and so is $\phi_\xi(1, z_{2,1})$ by (3-25). Thus, at least the first zero of $\phi_{\xi\xi}$ occurs to the left of the zeros of ϕ_ξ and ϕ. So $z_{2,1} < z_{1,1} < z_{0,1}$. Suppose this holds for the $(j-1)$st eigenvalues. Then we know that both $z_{0,j-1} < z_{1,j} < z_{0,j}$ and $z_{0,j-1} < z_{2,j} < z_{0,j}$. If $z_{1,j} < z_{2,j}$ then $z_{2,j-1} < z_{1,j-1} < z_{1,j} < z_{2,j}$, thus contradicting that the spectra of type 1 and 2 interlace.

The statements about the eigenfunctions follow immediately from equation (3-18) after setting $z = z_{a,i}, \lambda = z_{b,j}$. We turn now to the linear independence. We observe that the cubic string boundary value problem 3.1 can be equivalently written as an integral equation:

$$\phi(\xi; z) = z \int_0^1 G(\xi, \tau)\phi(\tau; z)g(\tau)\, d\tau, \tag{3-26}$$

where $G(\xi, \tau)$ is the Green's function satisfying the boundary conditions of 3.1. Then the linear independence of eigenfunctions corresponding to distinct eigenvalues follows from the general result about the eigenfunctions of a linear operator.

□

We now briefly analyze the degenerate case with the mass m_n at the end point $x_n = 1$. The remark right below the next theorem partially explains its relevance . First, we state the main theorem for the degenerate case.

Theorem 3.2. *Let us consider a cubic string with a finite measure g as in (3-9) with $x_n = 1$, and the BVP of type 2. Then*

1. *the spectrum is positive and simple*
2. *the zeros $\{z_{0,j}\}_j^{n-1}$ of $\phi(1; z)$ interlace with the zeros $\{z_{2,j}\}$ of $\phi_{\xi\xi}(1; z)$ and*

$$0 < z_{2,1} < z_{0,1} < \cdots z_{0,n-1} < z_{2,n}$$

   holds,

3. *the eigenfunctions $\phi(\xi; z_{2,j}) := \phi_j(\xi)$ can be chosen real, they satisfy*

$$(\phi_i, \phi_j)_g = \frac{\phi_{i,\xi}(1)\phi_{j,\xi}(1)}{z_{2,i} + z_{2,j}}, \tag{3-27}$$

   and they are linearly independent.

Remark 3.3. *We notice that* $\deg \phi(1; z) = \deg \phi_\xi(1; z) = n-1$ *while* $\deg \phi_{\xi\xi}(1; z) = n$*. This is in contrast to the previous cases with all positions $x_1, \ldots, x_n$ inside the interval* $[0, 1]$ *for which all polynomials have the same degree n.*

Proof. The spectrum is clearly given by the zeros of $\phi_{\xi\xi}(1; z)$. Let us first consider the case when m_n is placed slightly to the left of the point 1. Thus, initially, $l_n > 0$ (see (3-23)). By Theorem 3.1 $\phi(1; z)$ has n distinct positive zeros and so does $\phi_{\xi\xi}$ and they interlace. We subsequently take the limit $l_n \to 0$ in the above formulas. We will use the same letters for the limits to ease the notation. By simple perturbation argument, $z_{n,0} \to \infty$. Since $z = 0$ is not in the spectrum, $z_{2,1}$ has to stay away from 0. This shows that the spectrum is positive. Furthermore, in the limit $z_{0,1}, \ldots, z_{0,n-1}$ approach simple zeros of the BVP of type 0 for $n-1$ masses. Indeed, using (3-23) with $l_n = 0$ there, we obtain:

$$\phi(1+0; z) = \phi(1-0; z), \quad \phi_\xi(1+0; z) = \phi_\xi(1-0; z), \quad \phi_{\xi\xi}(1+0; z) = -zm_n\phi(1-0; z) + \phi_{\xi\xi}(1-0; z), \tag{3-28}$$

where 1 ± 0 refers to the right hand, or the left hand limit at 1. To see that the spectrum is simple we observe that if in the limit two successive eigenvalues coalesce, namely $z_{2,i} = z_{2,i+1}$, then necessarily $z_{2,i} = z_{0,i}$ because of the interlacing property. However now equation (3-28) implies that $\phi(1-0; z_{0,i}) = \phi_{\xi\xi}(1-0; z_{0,i}) = 0$ which contradicts Theorem 3.1 for the BVP of type 0 for $n-1$ masses. Thus, the zeros of $\phi(1+0, z)$ and $\phi_{\xi\xi}(1+0, z)$ interlace and we have

$$0 < z_{2,1} < z_{0,1} < \cdots < z_{2,n-1} < z_{0,n-1} < z_{2,n}.$$

To prove the statement about the eigenfunctions we use (3-18) and after setting $z = z_i, \lambda = z_j$ in that formula we obtain the required identity. The linear independence is proven by the same type of argument as in the proof of Theorem 3.1. □

We immediately have several results about the adjoint cubic string 3.3.

Corollary 3.2. *Given a discrete finite measure g*

1. *the adjoint cubic string (Definition 3.3) and the cubic string (Definition 3.1) have identical spectra.*
2. *Let $\phi^*_{a,i}$ be the eigenfunction corresponding to the eigenvalue $z_{a,i}$. Then the families of functions $\{\phi_{a,j}\}$ and $\{\phi^*_{a,j}\}$ are biorthogonal, that is:*

$$(\phi_{a,i}, \phi^*_{a,j})_g = 0 \qquad \textit{whenever} \qquad i \neq j. \tag{3-29}$$

3. *For $i = j$,*

$$(\phi_{a,i}, \phi^*_{a,i})_g = \begin{cases} -\phi_z(1; z_{0,i}) \neq 0, a = 0 \\ \phi_{\xi z}(1; z_{1,i}) \neq 0, a = 1 \\ -\phi_{\xi\xi z}(1; z_{2,i}) \neq 0, a = 2 \end{cases} \tag{3-30}$$

holds.

Proof. The first equality in Corollary 3.1 implies that the spectra of the cubic string and its adjoint are identical. The biorthogonality follows immediately from equations (3-15), (3-16) and (3-17). For $i = j$ we use Corollary 3.1. Finally, since the spectrum is simple the required derivatives with respect to z are nonzero. □

We conclude this section with the definition and some fundamental properties of two important functions which play a significant role in the remainder of the paper.

Definition 3.6. *We call*

$$W(z) := \frac{\phi_\xi(1; z)}{\phi(1; z)}, \qquad Z(z) = \frac{\phi_{\xi\xi}(1; z)}{\phi(1; z)} \tag{3-31}$$

the Weyl functions associated with the cubic string boundary value problem 3.1 of type 0. We call

$$W(z) := -\frac{\phi(1; z)}{\phi_\xi(1; z)}, \qquad Z(z) = \frac{\phi_{\xi\xi}(1; z)}{\phi_\xi(1; z)} \tag{3-32}$$

the Weyl functions associated with the cubic string boundary value problem 3.1 of type 1. We call

$$W(z) := -\frac{\phi_\xi(1; z)}{\phi_{\xi\xi}(1; z)}, \qquad Z(z) = -\frac{\phi(1; z)}{\phi_{\xi\xi}(1; z)} \tag{3-33}$$

the Weyl functions associated with the cubic string boundary value problem 3.1 of type 2.

Remark 3.4. *The definition of the Weyl functions for the BVP of type 2 in the degenerate case is identical to the one given above for the BVP of type 2.*

The Weyl functions W and Z are not independent, they are related by an identity which was originally formulated for the DP peakons in [9]. As an example we formulate such an identity for the BVP of type 2 (both the degenerate as well as the nondegenerate case).

Lemma 3.3. *Consider the BVP of type 2. Then the corresponding Weyl functions satisfy:*

$$W(z)W(-z) + Z(z) + Z(-z) = 0 \tag{3-34}$$

Proof. By formula (3-18)

$$B(\phi(\xi; z), \phi(\xi; -z))|_0^1 = 0, \tag{3-35}$$

which, when written out explicitly, gives the identity:

$$\phi_{\xi\xi}(1; z)\phi(1; -z) - \phi_\xi(1; z)\phi_\xi(1; -z) + \phi_{\xi\xi}(1; -z)\phi(1; z) = 0. \tag{3-36}$$

Upon dividing the last equation by $\phi_{\xi\xi}(1; z)\phi_{\xi\xi}(1; -z)$ we obtain the claim. □

We state now the fundamental theorem with regards to $W(z)$ and $Z(z)$. We state only the relevant results for the BVP of type 2 in the degenerate case , the remaining cases being merely variations of this, most transparent case.

Theorem 3.3. *Consider the BVP of type 2 (degenerate case). Then the Weyl functions W and Z have the following (Stieltjes) integral representations:*

$$W(z) = \int \frac{1}{z-y} d\beta(y), \quad Z(z) = \int \frac{1}{(z-y)(x+y)} d\beta(y) d\beta(x), \tag{3-37}$$

where $d\beta = \sum_{i=1}^n b_i \delta_{z_{2,i}}$, $b_i = \frac{-\phi(1;z_{2,i})}{\phi_{\xi\xi z}(1;z_{2,i})} > 0$.

Proof. Since $\phi(1; z), \phi_{\xi\xi}(1; z)$ have simple, interlacing zeros, and $\deg \phi(1; z) = n - 1$ while $\deg \phi_{\xi\xi}(1; z) = n$, $W(z)$ admits a partial fraction decomposition with simple factors:

$$W(z) = \sum_{i=1}^n \frac{b_i}{z - z_{2,i}},$$

where, by the residue calculus, $b_i = \frac{-\phi(1;z_{2,i})}{\phi_{\xi\xi z}(1;z_{2,i})}$. Moreover the b_is are all of the same sign because the zeros of $\phi(1; z)$ and $\phi_{\xi\xi}(1; z)$ interlace and, consequently, it suffices to check the sign of $\frac{-\phi(1;z_{2,i})}{\phi_{\xi\xi z}(1;z_{2,i})}$ at the first zero $z_{2,1}$. By Theorem 3.1 $\operatorname{sgn}(\phi(1; z_{2,1}) = 1$, and thus $b_1 > 0$ because on the first zero $\phi_{\xi\xi z}$ must be negative. Consequently, all $b_i > 0$.

Likewise, $Z(z)$ admits a partial fraction decomposition:

$$Z(z) = \sum_{i=1}^{n} \frac{c_i}{z - z_{2,i}},$$

and again, it follows from the second item in Theorem 3.1 that $c_i > 0$. Finally, by residue calculus, it follows from Lemma 3.3 that

$$c_i = \sum_{j=1}^{n} \frac{b_i b_j}{z_{2,i} + z_{2,j}},$$

which proves the integral representation for $Z(z)$.

□

3.3 Generalized Fourier transform and biorthogonality

Since $\phi_{a,i}$ are linearly independent we can decompose any $f \in L^2_g[0,1]$ in the basis of $\{\phi_{a,i}\}$ and use the dual family $\{\phi^*_{a,i}\}$ to compute the coefficients in the expansion:

$$f = \sum_i C_{a,i}\phi_{a,i}, \quad C_{a,i} = \frac{(\phi^*_{a,i}, f)_g}{(\phi^*_{a,i}, \phi_{a,i})_g}.$$

For each pair a, b for which $(\phi_{a,i}, \phi_{b,j})_g$ factorizes (item 3 in Theorem 3.1) we define two Hilbert spaces $H_\alpha := L^2(\mathbb{R}, d\alpha)$ and $H_\beta := L^2(\mathbb{R}, d\beta)$ where the measures $d\alpha$ and $d\beta$ are chosen by splitting the numerator of $(\phi_{a,i}, \phi_{b,j})_g$ followed by a unique choice of measures corresponding to the BVP of type a and b associated to the respective Weyl functions with matching numerators.

Example 3.1. *For types 00, item 3 in Theorem 3.1 and Definition 3.6 imply*

$$\mathrm{d}\alpha = d\beta = \sum_i \frac{\phi_{0,i.\xi}}{\phi_z(1; z_{0,i})}\delta_{z_{0,i}}$$

For types 01, item 3 in Theorem 3.1 and Definition 3.6 imply

$$\mathrm{d}\alpha = \sum_i \frac{\phi_{0,i,\xi\xi}}{\phi_z(1; z_{0,i})}\delta_{z_{0,i}}, \mathrm{d}\beta = -\sum_i \frac{\phi_{1,i}}{\phi_{\xi z}(1; z_{1,i})}\delta_{z_{1,i}}.$$

Furthermore, for every pair a, b specified above, we define a natural pairing between H_α and H_β, namely,

Definition 3.7.

$$\langle p|q\rangle = \int \frac{p(x)q(y)}{x+y} d\alpha(x)\beta(y).$$

We now introduce a family of generalized Fourier transforms adapted to each of the three types of BVPs

Definition 3.8. *Given $f \in L^2_g[0,1]$ and $a = 0,1,2$*

$$\hat{f}_a(z) := (-1)^a \int_0^1 \phi_a^*(\xi; z) f(\xi) g(\xi) d\xi. \tag{3-38}$$

Remark 3.5. *Observe that*

$$\hat{f}_a(z) = ((-1)^a \phi^*(\bullet, z), f(\bullet))_g,$$

and, in particular,

$$\hat{f}_a(z_{a,i}) = ((-1)^a \phi_{a,i}^*, f)_g$$

whenever z equals to one of the points of the spectrum of (3-8).

Remark 3.6. *A map of this type was introduced in the context of an inhomogeneous string problem by I.S. Kac and M.G. Krein in [16] as a generalization of the Fourier transform.*

The main property of this map is captured in the following theorem.

Theorem 3.4. *For every pair a, b for which $(\phi_{a,i}, \phi_{b,j})_g$ factorizes (item 3 in Theorem 3.1) the generalized Fourier transforms satisfy Parseval's identity, that is, for every $f, h \in L^2_g[0,1]$*

$$(f,h)_g = \langle \hat{f}_a | \hat{h}_b \rangle = \langle \hat{h}_a | \hat{f}_b \rangle. \tag{3-39}$$

Proof. Let us fix a, b for which $(\phi_{a,i}, \phi_{b,j})_g$ factorizes. Consider two functions $f, h \in L^2_g[0,1]$. Writing their expansions in the bases $\{\phi_{a,i}\}$, $\{\phi_{b,i}\}$ respectively, we obtain

$$f = \sum_i \frac{(\phi_{a,i}^*, f)_g}{(\phi_{a,i}^*, \phi_{a,i})_g} \phi_{a,i}, \quad h = \sum_j \frac{(\phi_{b,j}^*, h)_g}{(\phi_{b,j}^*, \phi_{b,j})_g} \phi_{b,j}.$$

Hence their inner product reads:

$$(f,h)_g = \sum_{i,j} \frac{(\phi_{a,i}^*, f)_g}{(\phi_{a,i}^*, \phi_{a,i})_g} \frac{(\phi_{b,j}^*, h)_g}{(\phi_{b,j}^*, \phi_{b,j})_g} (\phi_{a,i}, \phi_{b,j})_g$$

Applying now item 3 from Theorem 3.1 as well as item 2 from Lemma 3.1 we obtain

$$\begin{aligned}
Type\,00 :(f,h)_g &= \sum_{i,j} \frac{(\phi_{0,i}^*, f)_g}{\phi_z(1; z_{0,i})} \frac{(\phi_{0,j}^*, h)_g}{\phi_z(1; z_{0,j})} \frac{\phi_\xi(1; z_{0,i}) \phi_\xi(1; z_{0,j})}{z_{0,i} + z_{0,j}}, \\
Type\,01 :(f,h)_g &= \sum_{i,j} \frac{(\phi_{0,i}^*, f)_g}{\phi_z(1; z_{0,i})} \frac{(\phi_{1,j}^*, h)_g}{\phi_{\xi z}(1; z_{1,j})} \frac{\phi_{\xi\xi}(1; z_{0,i}) \phi(1; z_{1,j})}{z_{0,i} + z_{1,j}}, \\
Type\,12 :(f,h)_g &= \sum_{i,j} \frac{(\phi_{1,i}^*, f)_g}{\phi_{\xi z}(1; z_{1,i})} \frac{(\phi_{2,j}^*, h)_g}{\phi_{\xi\xi z}(1; z_{2,j})} \frac{\phi_{\xi\xi}(1; z_{1,i}) \phi(1; z_{2,j})}{z_{1,i} + z_{2,j}}, \\
Type\,22 :(f,h)_g &= \sum_{i,j} \frac{(\phi_{2,i}^*, f)_g}{\phi_{\xi\xi z}(1; z_{2,i})} \frac{(\phi_{2,j}^*, h)_g}{\phi_{\xi\xi z}(1; z_{2,j})} \frac{\phi_\xi(1; z_{2,i}) \phi_\xi(1; z_{j,2})}{z_{2,i} + z_{2,j}}.
\end{aligned}$$

We now define the weights b_j, a_j generating the measures $\mathrm{d}\beta = \sum_j B_j \delta_{z_j}$, $d\alpha = \sum_j A_j \delta_{z_j}$ respectively, as residues of Ws or Zs:

$$\begin{aligned}
&\mathit{Type}\, 00 : A_i = \frac{\phi_\xi(1; z_{0,i})}{\phi_z(1; z_{0,i})}, \qquad B_j = \frac{\phi_\xi(1; z_{0,j})}{\phi_z(1; z_{0,j})},\\
&\mathit{Type}\, 01 : A_i = \frac{\phi_{\xi\xi}(1; z_{0,i})}{\phi_z(1, z_{0,i})}, \qquad B_j = -\frac{\phi(1; z_{1,j})}{\phi_{\xi z}(1; z_{1,j})},\\
&\mathit{Type}\, 12 : A_i = \frac{\phi_{\xi\xi}(1, z_{1,i})}{\phi_{\xi z}(1, z_{1,i})}, \qquad B_j = -\frac{\phi(1, z_{2,j})}{\phi_{\xi\xi z}(1, z_{2,j})},\\
&\mathit{Type}\, 22 : A_i = -\frac{\phi_\xi(1, z_{2,i})}{\phi_{\xi\xi z}(1, z_{2,i})}, \qquad B_j = -\frac{\phi_\xi(1, z_{2,j})}{\phi_{\xi\xi z}(1, z_{2,j})},
\end{aligned}$$

and thus indeed

$$(f, h)_g = \sum_{i,j} ((-1)^a \phi^*_{a,i}, f)_g ((-1)^b \phi^*_{b,j}, h)_g \frac{A_i B_j}{z_{a,i} + z_{b,j}}.$$

Thus, in view of Remark 3.5, we obtain

$$(f, h)_g = \sum_{i,j} \hat{f}_1(z_{a,i}) \hat{h}_b(z_{b,j}) \frac{A_i B_j}{z_{a,i} + z_{2,j}} = \int \frac{\hat{f}_a(x)\hat{h}_b(y)}{x+y} d\alpha(x) d\beta(y) = \langle \hat{f}_a | \hat{h}_b \rangle.$$

□

Remark 3.7. *Expanding an arbitrary* $f \in L^2_g[0,1]$

$$f = \sum_i \frac{(\phi^*_{a,i}, f)_g}{(\phi^*_{a,i}, \phi_{a,i})_g} \phi_{a,i},$$

allows one to conclude that

$$\delta(\xi, \xi') := \sum_i \frac{\phi_{a,i}(\xi)\phi^*_{a,i}(\xi')}{(\phi^*_{a,i}, \phi_{a,i})_g} \tag{3-40}$$

plays a role of the Dirac delta on $L^2_g[0,1]$.

Consequently, it is elementary to find the inverse Fourier transforms

Lemma 3.4. *Consider the BVP of type* a. *Let* $\{z_{a,i}\}$ *be the corresponding spectrum and let* $d\nu_a = \sum_i \delta_{z_{a,i}}$ *be an associated measure. Then the inverse generalized Fourier transform of type* a *is given by*

$$(-1)^a \int \hat{f}_a(z) \frac{\phi(\xi; z)}{(\phi^*_a(\bullet, z), \phi(\bullet, z))} d\nu_a(z) \tag{3-41}$$

Proof. This is a direct computation:

$$\begin{aligned}
&(-1)^a \int \hat{f}_a(z) \frac{\phi(\xi; z)}{(\phi^*_a(\bullet, z), \phi(\bullet, z))} d\nu_a(z) = (-1)^a \sum_i \hat{f}_a(z_{a,i}) \frac{\phi_{a,i}(\xi)}{(\phi^*_{a,i}, \phi_{a,i})_g} =\\
&\sum_i (\phi^*_{a,i}, f)_g \frac{\phi_{a,i}(\xi)}{(\phi^*_{a,i}, \phi_{a,i})_g} = \sum_i (\phi_{a,i}, f)_g \frac{\phi_{a,i}(\xi)}{(\phi^*_{a,i}, \phi_{a,i})_g} =\\
&\int_0^1 \sum_i \frac{\phi_{a,i}(\xi)}{(\phi^*_{a,i}, \phi_{a,i})_g} \phi_{a,i}(\xi') f(\xi') g(\xi') d\xi' = \int_0^1 \delta(\xi, \xi') f(\xi') g(\xi') d\xi'.
\end{aligned}$$

□

There are in general two measures associated with each type of the BVP, one generated by W the other by Z. One can use either one of the them instead of the measure $d\nu$. We give as an example the relevant statement for the the case of the BVP of type 2, both the degenerate as well as the nondegenerate one.

Lemma 3.5. *The inverse generalized Fourier transform of type 2 is given by*

$$\int \hat{f}(z)\frac{\phi(\xi;z)}{\phi_\xi(1;z)}d\beta(z) \tag{3-42}$$

Proof. From the definition of $d\beta$ given in Theorem 3.3 we see that

$$\int \hat{f}(z)\frac{\phi(\xi;z)}{\phi_\xi(1;z)}d\beta(z) = \sum_i \hat{f}(z_{2,i})\frac{\phi_{2,i}(\xi)}{\phi_\xi(1;z_{2,i})}b_i = \sum_i \hat{f}(z_{2,i})\frac{\phi_{2,i}(\xi)}{\phi_\xi(1;z_{2,i})}(-\frac{\phi_\xi(1,z_{2,i})}{\phi_{\xi\xi z}(1,z_{2,i})}) =$$
$$\sum_i (\phi_{2,i},f)_g\frac{\phi_{2,i}(\xi)}{\phi_\xi(1;z_{2,i})}(-\frac{\phi_\xi(1,z_{2,i})}{\phi_{\xi\xi z}(1,z_{2,i})} = \sum_i (\phi_{2,i},f)_g\frac{\phi_{2,i}(\xi)}{(\phi^*_{2,i},\phi_{2,i})_g},$$

where in the last two steps we used Remark 3.5 and equation (3-30) respectively. Thus

$$\int \hat{f}(z)\frac{\phi(\xi;z)}{\phi_\xi(1;z)}d\beta(z) = \int_0^1 \sum_i \frac{\phi_{2,i}(\xi)\phi^*_{2,i}(\xi')}{(\phi^*_{2,i},\phi_{2,i})_g}f(\xi')g(\xi')d\xi'.$$

Finally, using equation (3-40) we obtain the claim. □

We now consider an example of generalized Fourier transforms relevant for the remainder of this paper.

Example 3.2. Biorthogonal polynomials *We consider a sequence $\chi_j := \chi_{(\xi_{n-j}-\epsilon,\xi_{n-j}+\epsilon)}$ of indicator functions enclosing points ξ_{n-j} with ϵ small enough to ensure non overlapping supports. Consider now the generalized Fourier transforms for a, b as in Theorem 3.4:*

$$\hat{\chi}_{a,j}(x) = (-1)^a\phi^*_a(\xi_j;x)m_{n-j}, \qquad \hat{\chi}_{b,j}(y) = (-1)^b\phi^*_b(\xi_j;y)m_{n-j}.$$

Then, clearly, $<\hat{\chi}_{a,i}|\hat{\chi}_{b,j}>=0$, $i \neq j$. Also, both $\hat{\chi}_{a,i}(x)$ and $\hat{\chi}_{b,j}(y)$ are polynomials in x, y respectively, whose degrees are $\deg \hat{\chi}_{a,j}(x) = \deg \hat{\chi}_{b,j}(y) = j$ by (3-12).

4 Total positivity of bimoment matrices

As one can see from previous sections, the kernel $K(x,y) = \frac{1}{x+y}$, $x, y > 0$, which we will refer to as the Cauchy kernel, plays a significant, albeit mysterious, role. We now turn to explaining the role of this kernel. We recall, following [21], the definition of the totally positive kernel.

Definition 4.1. *A real function $K(x,y)$ of two variables ranging over linearly ordered sets $\mathcal{X}$ and $\mathcal{Y}$, respectively, is said to be totally positive (TP) if for all*

$$x_1 < x_2 < \cdots < x_m, \quad y_1 < y_2 < \cdots < y_m \quad x_i \in \mathcal{X}, y_j \in \mathcal{Y}, m \in \mathbb{N} \tag{4-1}$$

we have

$$\det \begin{bmatrix} K(x_1,y_1) & K(x_1,y_2) & \cdots K(x_1,y_m) \\ K(x_2,y_1) & K(x_2,y_2) & \cdots K(x_2,y_m) \\ \vdots & \vdots & \vdots \\ K(x_m,y_1) & K(x_m,y_2) & \cdots K(x_m,y_m) \end{bmatrix} > 0 \tag{4-2}$$

We will also use a discrete version of the same concept.

Definition 4.2. *A matrix $A := [a_{ij}]$, $i,j = 0,1,\cdots n$ is said to be totally positive (TP) if all its minors are strictly positive. A matrix $A := [a_{ij}]$, $i,j = 0,1,\cdots n$ is said to be totally nonnegative (TN) if all its minors are nonnegative. A TN matrix A is said to be oscillatory if some positive integer power of A is TP.*

Since we will be working with matrices of infinite size we introduce a concept of the principal truncation.

Definition 4.3. *A finite $n+1$ by $n+1$ matrix $B := [b_{i,j}], i,j = 0,1,\cdots n$ is said to be the principal truncation of an infinite matrix $A := [a_{ij}]$, $i,j = 0,1,\cdots$ if $b_{i,j} = a_{i,j}, i,j = 0,1,\cdots n$. In such a case B will be denoted $A[n]$.*

Finally,

Definition 4.4. *An infinite matrix $A := [a_{ij}]$, $i,j = 0,1,\cdots$ is said to be TP (TN) if $A[n]$ is TP (TN) for every $n = 0,1,\cdots$.*

Definition 4.5. *Basic Setup*

Let $K(x,y)$ be a **totally positive kernel** *on $\mathbb{R}_+ \times \mathbb{R}_+$ and let $\mathrm{d}\alpha, \mathrm{d}\beta$ be two Stieltjes measures on R_+. We make two simplifying assumptions to avoid degenerate cases:*

1. *0 is not an atom of either of the measures (i.e. $\{0\}$ has zero measure).*
2. *α and β have infinitely many points of increase.*

We furthermore assume:

3. *the polynomials are dense in the corresponding Hilbert spaces $H_\alpha := L^2(\mathbb{R}_+, \mathrm{d}\alpha)$, $H_\beta := L^2(\mathbb{R}_+, \mathrm{d}\beta)$,*
4. *the map*
$$K : H_\beta \to H_\alpha, \qquad Kq(x) := \int K(x,y)q(y)\mathrm{d}\beta(y) \tag{4-3}$$
is bounded, injective and has a dense range in H_α.

Under these assumptions K provides a non-degenerate pairing between H_β and H_α:

$$\langle a|b\rangle = \iint a(x)b(y)K(x,y)\mathrm{d}\alpha\mathrm{d}\beta, \quad a\in H_\alpha, b\in H_\beta. \tag{4-4}$$

Now, let us consider the matrix of generalized bimoments

$$I_{ij} := \int\int x^i y^j K(x,y)\mathrm{d}\alpha(x)\mathrm{d}\beta(y)\ . \tag{4-5}$$

We have our preliminary result

Theorem 4.1. *The semiinfinite matrix I is TP.*

Proof. According to a theorem of Fekete, (see Chapter 2, Theorem 3.3 in [21]), we only need to consider minors of consecutive rows/columns. Writing out the determinant,

$$\Delta_n^{ab} := \det[I_{a+i,b+j}]_{0\le i,j\le n-1}$$

we find

$$\Delta_n^{ab} = \sum_{\sigma\in S_n}\epsilon(\sigma)\iint\prod_{j=1}^n x_j^a y_j^b\prod_{j=1}^n x_j^{\sigma_j-1}y_j^{j-1}K(x_j,y_j)\mathrm{d}^n\alpha(X)\mathrm{d}^n\beta(Y) =$$
$$\iint C(X)^a C(Y)^b\Delta(X)\prod_{j=1}^n y_j^{j-1}\prod_{j=1}^n K(x_j,y_j)\mathrm{d}^n\alpha\mathrm{d}^n\beta.$$

Since our intervals are subsets of $\mathbb{R}_+$ we can absorb the powers of $C(X), C(Y)$ into the measures to simplify the notation. Moreover, the function $S(X,Y) := \prod_{j=1}^n K(x_j,y_j)$ enjoys the following simple property

$$S(X,Y_\sigma) = S(X_{\sigma^{-1}},Y)$$

for any $\sigma\in S_n$. Finally, the product measures $\mathrm{d}^n\alpha = \mathrm{d}^n\alpha(X), \mathrm{d}^n\beta = \mathrm{d}^n\beta(Y)$ are clearly permutation invariant.

Thus, without any loss of generality, we only need to show that

$$D_n := \iint\Delta(X)\prod_{j=1}^n y_j^{j-1}S(X,Y)\mathrm{d}^n\alpha\mathrm{d}^n\beta\ > 0,$$

which is tantamount to showing positivity for $a=b=0$. First, we symmetrize D_n with respect to the variables X; this produces

$$D_n = \frac{1}{n!}\sum_{\sigma\in S_n}\iint\Delta(X_\sigma)\prod_{j=1}^n y_j^{j-1}S(X_\sigma,Y)\mathrm{d}^n\alpha\mathrm{d}^n\beta = \frac{1}{n!}\sum_{\sigma\in S_n}\iint\Delta(X)\epsilon(\sigma)\prod_{j=1}^n y_j^{j-1}S(X,Y_{\sigma^{-1}})\mathrm{d}^n\alpha\mathrm{d}^n\beta =$$
$$\frac{1}{n!}\sum_{\sigma\in S_n}\iint\Delta(X)\epsilon(\sigma)\prod_{j=1}^n y_{\sigma_j}^{j-1}S(X,Y)\mathrm{d}^n\alpha\mathrm{d}^n\beta = \frac{1}{n!}\iint\Delta(X)\Delta(Y)S(X,Y)\mathrm{d}^n\alpha\mathrm{d}^n\beta.$$

Subsequent symmetrization over the Y variables does not change the value of the integral and we obtain (after restoring the definition of $S(X,Y)$)

$$D_n = \frac{1}{(n!)^2}\sum_{\sigma\in S_n}\epsilon(\sigma)\iint \Delta(X)\Delta(Y)\prod_{j=1}^n K(x_j, y_{\sigma_j})\mathrm{d}^n\alpha\mathrm{d}^n\beta = \frac{1}{(n!)^2}\iint \Delta(X)\Delta(Y)\det[K(x_i,y_j)]_{i,j\le n}\mathrm{d}^n\alpha\mathrm{d}^n\beta.$$

Finally, since $\Delta(X)\Delta(Y)\det[K(x_i,y_j)]_{i,j\le n}\mathrm{d}^n\alpha\mathrm{d}^n\beta$ is permutation invariant, it suffices to integrate over the region $0<x_1<x_2<\cdots<x_n \times 0<y_1<y_2<\cdots<y_n$, and, as a result

$$D_n = \iint_{\substack{0<x_1<x_2<\cdots<x_n\\ 0<y_1<y_2<\cdots<y_n}} \Delta(X)\Delta(Y)\det[K(x_i,y_j)]_{i,j\le n}\mathrm{d}^n\alpha\mathrm{d}^n\beta. \tag{4-6}$$

Due to the total positivity of the kernel $K(x,y)$ the integrand is a positive function of all variables and so the integral must be strictly positive. □

To simplify future computations we define

$$[x] := (1, x, x^2, \dots)^T \tag{4-7}$$

so that the matrix of generalized bimoments (4-5) is simply given by:

$$I = \langle [x] | [y]^T \rangle. \tag{4-8}$$

Observe that multiplying the measure $\mathrm{d}\alpha(x)$ by x^i or, multiplying $\mathrm{d}\beta(y)$ by y^j, is tantamount to multiplying I on the left, respectively on the right, by Λ^i, respectively by $(\Lambda^T)^j$, which gives us a whole family of bimoment matrices associated with the same $K(x,y)$ but different measures. Thus we have

Corollary 4.1. *For any nonnegative integers i,j the matrix of generalized bimoments $\Lambda^i I (\Lambda^T)^j$ is TP.*

We conclude this section with a few comments about the scope of Theorem 4.1.

Remark 4.1. *Provided that the negative moments are well defined, the theorem then applies to the doubly infinite matrix $I_{i,j}$, $i,j\in\mathbb{Z}$.*

Remark 4.2. *If the intervals are $\mathbb{R}$ and $K(x,y)=\mathrm{e}^{xy}$ then the proof above fails because we cannot re-define the measures by multiplying by powers of the variables, since they become then signed measures, so in general the matrix of bimoments is* **not** *totally positive. Nevertheless the proof above shows (with $a=b=0$ or $a,b\in 2\mathbb{Z}$) that the matrix of bimoments is positive definite and –in particular– the biorthogonal polynomials always exist, which is known and proved in [14].*

4.1 Biorthogonal polynomials

Due to the total positivity of the matrix of bimoments in our setting, there exist uniquely defined two sequences of monic polynomials

$$\widetilde{p}_n(x) = x^n + \dots \ , \ \ \widetilde{q}_n(y) = y^n + \dots$$

such that

$$\iint \widetilde{p}_n(x)\widetilde{q}_m(y)K(x,y)\mathrm{d}\alpha(x)\mathrm{d}\beta(y) = h_n\delta_{mn} \ .$$

Standard considerations (Cramer's Rule) show that they are provided by the following formulæ

$$\widetilde{p}_n(x) = \frac{1}{D_n}\det\left[\begin{array}{ccc|c} I_{00} & \dots & I_{0n-1} & 1 \\ \vdots & & \vdots & \vdots \\ I_{n0} & \dots & I_{nn-1} & x^n \end{array}\right] \qquad \widetilde{q}_n(y) = \frac{1}{D_n}\det\left[\begin{array}{ccc} I_{00} & \dots & I_{0n} \\ \vdots & & \vdots \\ I_{n-10} & \dots & I_{n-1n} \\ \hline 1 & \dots & y^n \end{array}\right] \tag{4-9}$$

$$h_n = \frac{D_{n+1}}{D_n} > 0, \tag{4-10}$$

where $D_j > 0$ by equation (4-6). For convenience we re-define the sequence in such a way that they are also **normalized** (instead of monic), by dividing them by the square root of h_n;

$$p_n(x) = \frac{1}{\sqrt{D_nD_{n+1}}}\det\left[\begin{array}{ccc|c} I_{00} & \dots & I_{0n-1} & 1 \\ \vdots & & \vdots & \vdots \\ I_{n0} & \dots & I_{nn-1} & x^n \end{array}\right], \tag{4-11}$$

$$q_n(y) = \frac{1}{\sqrt{D_nD_{n+1}}}\det\left[\begin{array}{ccc} I_{00} & \dots & I_{0n} \\ \vdots & & \vdots \\ I_{n-10} & \dots & I_{n-1n} \\ \hline 1 & \dots & y^n \end{array}\right], \tag{4-12}$$

$$\langle p_n|q_m\rangle = \delta_{nm}. \tag{4-13}$$

We note also that the BOPs can be obtained by triangular transformations of $[x], [y]$

$$\mathbf{p} = S_p[x] \ , \quad \mathbf{q} = S_q[y] \tag{4-14}$$

where $S_{p,q}$ are (formally) invertible lower triangular matrices such that $S_p^{-1}(S_q^{-1})^T = I$, where, we recall, I is the generalized bimoment matrix. Moreover, our BOPs satisfy, by construction, the recursion relations:

$$xp_i(x) = X_{i,i+1}p_{i+1}(x) + X_{i,i}p_i(x) + \cdots X_{i,0}p_0(x),$$
$$yq_i(y) = Y_{i,i+1}q_{i+1}(y) + Y_{i,i}q_i(y) + \cdots Y_{i,0}q_0(y),$$

which will be abbreviated as

$$x\mathbf{p}(x) = \mathbf{X}\mathbf{p}(x)\ , \quad y\mathbf{q}(y)^T = \mathbf{q}(y)\mathbf{Y}^T, \tag{4-15}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are Hessenberg matrices with positive entries on the supradiagonal, and $\mathbf{p}(x)\,\mathbf{q}(y)$ are infinite column vectors $\mathbf{p}(x)^T := (p_0(x), p_1(x), p_2(x), \dots)^t$, $\mathbf{q}(y)^T := (q_0(y), q_1(y), q_2(y), \dots)^T$ respectively.

The biorthogonality can now be written as

$$\langle \mathbf{p}|\mathbf{q}^T\rangle = \mathbf{1}\ . \tag{4-16}$$

Moreover

$$\langle x\mathbf{p}|\mathbf{q}^T\rangle = \mathbf{X}\ , \qquad \langle \mathbf{p}|y\mathbf{q}^T\rangle = \mathbf{Y}^T \tag{4-17}$$

Remark 4.3. *The significance of the last two formulas lies in the fact that the operator of multiplication is no longer symmetric with respect to the pairing $\langle\bullet|\bullet\rangle$ and as a result the matrices $\mathbf{X}$ and $\mathbf{Y}^T$ are distinct.*

4.2 Simplicity of the zeroes

We recall the definition of a Chebyshev system. We refer to [22] and [23] for more information.

Definition 4.6. *We call a system of continuous functions $\{u_i(x)|i = 0\cdots n\}$ defined on a subset U of $\mathbb{R}$ a Chebyshev system of order n on U if any nontrivial linear combination $\sum_{i=0}^n a_i u_i$, $\sum_{i=0}^n a_i^2 \neq 0$ has no more than n zeros on U.*

Another closely related concept is that of a Markov sequence (see [23], p.181).

Definition 4.7. *A sequence of continuous functions*

$$u_0, u_1, u_2, \dots$$

is a Markov sequence on U if for every n the functions $\{u_i(x)|i = 0\cdots n\}$ form a Chebyshev system of order n on U.

The following theorem is a convenient restatement of Lemma 2 in [23], p.137.

Theorem 4.2. *Given a system of continuous functions $\{u_i(x)|i = 0\cdots n\}$ let us define the vector field*

$$\mathbf{u}(x) = \begin{bmatrix} u_0(x) \\ u_1(x) \\ \vdots \\ u_n(x) \end{bmatrix}, \qquad x \in U. \tag{4-18}$$

Then $\{u_i(x)|i = 0\cdots n\}$ is a Chebyshev system of order n on U iff the top exterior power

$$\mathbf{u}(x_0) \wedge \mathbf{u}(x_1) \wedge \cdots \mathbf{u}(x_n) \neq 0 \tag{4-19}$$

for all $x_0 < x_1 < \cdots < x_n$ in U. Furthermore, for $\{u_i(x)|i = 0\cdots\}$, if we denote the truncation of $\mathbf{u}(x)$ to the first $n+1$ components by $\mathbf{u}_n(x)$, then $\{u_i(x)|i = 0\cdots\}$ is a Markov system iff the top exterior power

$$\mathbf{u}_n(x_0) \wedge \mathbf{u}_n(x_1) \wedge \cdots \mathbf{u}_n(x_n) \neq 0 \tag{4-20}$$

for all $x_0 < x_1 < \cdots < x_n$ in U and all $n \in \mathbb{N}$.

The following well known theorem is now immediate

Theorem 4.3. *Suppose $\{u_i(x)|i = 0\cdots n\}$ is a Chebyshev system of order n on U, and suppose we are given n distinct points $x_1, \cdots x_n$ in U. Then, up to a multiplicative factor, the only generalized polynomial $P(x) = \sum_{i=0}^n a_i u_i(x)$, which vanishes precisely at $x_1, \cdots x_n$ in U is given by*

$$P(x) = \mathbf{u}(x) \wedge \mathbf{u}(x_1) \wedge \cdots \mathbf{u}(x_n) \tag{4-21}$$

Theorem 4.4. *Denote by $u_i(x) = \int K(x,y)y^i d\beta(y)$, $i = 0\cdots n$. Then $\{u_i(x)|i = 0\cdots n\}$ is a Chebyshev system of order n on $\mathbb{R}_+$. Moreover, $P(x)$ as defined in Theorem 4.3 changes sign each time x passes through any of the zeros x_j.*

Proof. It is instructive to look at the computation. Let $x_0 < x_1 < \cdots x_n$, then using multi-linearity of the exterior product,

$$\begin{aligned}
&P(x_0) = \mathbf{u}(x_0) \wedge \mathbf{u}(x_1) \wedge \cdots \mathbf{u}(x_n) = \\
&\int K(x_0,y_0)K(x_1,y_1)\cdots K(x_n,y_n)[y_0]_n \wedge [y_1]_n \wedge \cdots \wedge [y_n]_n d\beta(y_0)\cdots d\beta(y_n) = \\
&\frac{1}{n!}\int \det[K(x_i,y_j)]_{i,j=0}^n \Delta(Y) d\beta(y_0)\cdots d\beta(y_n) = \\
&\int_{y_0<y_1<\cdots y_n} \det[K(x_i,y_j)]_{i,j=0}^n \Delta(Y) d\beta(y_0)\cdots d\beta(y_n),
\end{aligned}$$

where

$$[y]_n = \begin{bmatrix} y^0 \\ y^1 \\ \vdots \\ y^n \end{bmatrix}. \tag{4-22}$$

Thus $P(x_0) > 0$. The rest of the proof is the argument about the sign of the integrand. To see how sign changes we observe that the sign of P depends only on the ordering of $x, x_1, x_2, \cdots x_n$, in view of the total positivity of the kernel. In other words, the sign of P is $sgn(\pi)$ where π is the permutation rearranging $x, x_1, x_2, \cdots x_n$ in in an increasing sequence. □

Corollary 4.2. *Let $\{f_i(x) := \int K(x,y)q_i(y)d\beta(y), |i = 0\cdots\}$. Then $\{f_i(x)|i = 0\cdots n\}$ is a Markov sequence on $\mathbb{R}_+$,* *Proof.* Indeed, Theorem 4.2 implies that the group $GL(n+1)$ acts on the set of Chebyshev systems of order n. It suffices now to observe that q_j are obtained from $\{1, y, \cdots, y^n\}$ by an invertible transformation. □

Remark 4.4. *Observe that* $\{f_i(x) | i = 0 \cdots n\}$ *is a Markov sequence regardless of biorthogonality.*

Biorthogonality enters however in the main theorem

Theorem 4.5. *The zeroes of* p_n, q_n *are all simple and positive. They fall within the convex hull of the support of the measure* $\mathrm{d}\alpha$ *(for* p_n*'s) and* $\mathrm{d}\beta$ *(for the* q_n*'s).*

Proof. We give first a proof for p_n. The theorem is trivial for $n = 0$. For $1 \leq n$, let us suppose p_n has $r < n$ zeros of odd order in the convex full of *supp*$(\mathrm{d}\alpha)$. In full analogy with the classical case, $1 \leq r$, since

$$\int p_n(x) f_0(x) d\alpha(x) = \iint p_n(x) K(x, y) d\alpha(x) d\beta(y) = 0$$

by biorthogonality, forcing, in view of positivity of $K(x, y)$, $p_n(x)$ to change sign in the convex hull of *supp*$(\mathrm{d}\alpha)$. In the general case, denote the zeros by $x_1 < x_2 < \cdots x_r$. Using a Chebyshev system $f_i(x), i = 0, \cdots r$ on $\mathbb{R}_+$ we can construct a unique, up to a multiplicative constant, generalized polynomial which vanishes exactly at those points, namely

$$R(x) = F(x) \wedge F(x_1) \wedge F(x_2) \wedge \cdots \wedge F(x_r) \tag{4-23}$$

where

$$F(x) = \begin{bmatrix} f_0(x) \\ f_1(x) \\ \vdots \\ f_r(x) \end{bmatrix}, \qquad x \in \mathbb{R}.$$

It follows then directly from biorthogonaliy that

$$\int p_n(x) F(x) \wedge F(x_1) \wedge F(x_2) \wedge \cdots \wedge F(x_r) d\alpha(x) = 0 \tag{4-24}$$

On the other hand, $R(x)$ is proportional to $P(x)$ in Theorem 4.3 which, by Theorem 4.4, changes sign at each of its zeroes, so the product $p_n(x)R(x)$ is nonzero and of fixed sign over $\mathbb{R}_+ \setminus \{x_1, x_2, \cdots, x_r\}$. Consequently, the integral is nonzero, since α is assumed to have infinitely many points of increase. Thus, in view of the contradiction, $r \geq n$, hence $r = n$, for p_n is a polynomial of degree n. The case of q_n follows by observing that the adjoint K^* is also a TP kernel, and hence it suffices to switch α with β throughout the argument given above. □

Lemma 4.1. *In the notation of Corollary 4.2* $f_n(x)$ *has* n *zeros and* n *sign changes in the convex hull of supp*$(\mathrm{d}\alpha)$.

Proof. Clearly, since $\{u_i(x)|i = 0 \cdots n\}$ is a Chebyshev system of order n on $\mathbb{R}_+$, the number of zeros of f_n cannot be greater than n. Again, from

$$\int f_n(x)p_0(x)\mathrm{d}\alpha(x) = 0,$$

we conclude that f_n changes sign at least once within the convex hull of $supp(\mathrm{d}\alpha)$. Let then $x_1 < x_2 < \cdots x_r$, $1 \le r \le n$ be all zeros of f_n within the convex hull of $supp(\mathrm{d}\alpha)$ at which f_n changes its sign. Thus, on one hand,

$$\int \epsilon \prod_{i=1}^{r}(x - x_i)f_n(x)\mathrm{d}\alpha(x) > 0, \qquad \epsilon = \pm,$$

while, on the other hand, using biortogonality we get

$$\int \epsilon \prod_{i=1}^{r}(x - x_i)f_n(x)\mathrm{d}\alpha(x) = 0, \qquad \epsilon = \pm,$$

which shows that $r = n$. □

In view of Theorem 4.3 the statement about the zeros of f_n has the following corollary

Corollary 4.3. Heine-like representation for f_n

$$f_n(x) = Cu(x) \wedge u(x_1) \wedge u(x_2) \cdots \wedge u(x_n) \tag{4-25}$$

where x_j are the zeros of f_n.

5 Cauchy BOPs

From now on we restrict our attention to the particular case of the totally positive kernel

$$K(x, y) = \frac{1}{x + y} \tag{5-1}$$

and we call this case "Cauchy kernel" and correspondingly "Cauchy BOPs" because of the appearance of Cauchy matrices. Thus, from this point onward, we will be studying the general properties of BOPs for the pairing

$$\iint p_n(x)q_m(y)\frac{\mathrm{d}\alpha(x)\mathrm{d}\beta(y)}{x + y} = \langle p_n|q_m\rangle \ . \tag{5-2}$$

Until further notice, we do not assume anything about the relationship between the two measures $\mathrm{d}\alpha, \mathrm{d}\beta$, other than what is in the basic setup of Definition 4.5.

5.1 Rank One Shift Condition

It follows immediately from equation (5-1) that

$$I_{i+1,j} + I_{i,j+1} = \langle x^{i+1}|y^j\rangle + \langle x^i|y^{j+1}\rangle = \int x^i \mathrm{d}\alpha \int y^j \mathrm{d}\beta \ , \tag{5-3}$$

which, with the help of the shift matrix Λ and the matrix of generalized bimoments I, can be written as:

$$\begin{aligned}
&\Lambda I + I\Lambda^T = \boldsymbol{\alpha}\boldsymbol{\beta}^T,\\
&\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots)^T \ , \quad \alpha_j = \int x^j \mathrm{d}\alpha(x) > 0,\\
&\boldsymbol{\beta} = (\beta_0, \beta_1, \dots)^T \ , \quad \beta_j = \int y^j \mathrm{d}\beta(y) > 0.
\end{aligned}$$

Moreover, by linearity and equation (4-17), we have

$$\mathbf{X} + \mathbf{Y}^T = \boldsymbol{\pi}\boldsymbol{\eta}^T \ , \qquad \boldsymbol{\pi} := \int \mathbf{p}\mathrm{d}\alpha \ , \quad \boldsymbol{\eta} := \int \mathbf{q}\mathrm{d}\beta \tag{5-4}$$

which connects the multiplication operators in H_α and H_β. Before we elaborate on the nature of this connection we need to clarify one aspect of equation (5-4).

Remark 5.1. *One needs to exercise a great deal of caution using the matrix relation given by equation (5-4). Its only rigorous meaning is in action on vectors with finitely many nonzero entries or, equivalently, this equation holds for all principal truncations.*

Proposition 5.1. *The vectors $\boldsymbol{\pi}, \boldsymbol{\eta}$ are strictly positive (have nonvanishing positive coefficients).*

Proof. We prove the assertion only for $\boldsymbol{\pi}$, the one for $\boldsymbol{\eta}$ being obtained by interchanging the roles of $\mathrm{d}\alpha$ and $\mathrm{d}\beta$.

From the expressions (4-12) for $p_n(x)$ we immediately have

$$\pi_n = \sqrt{\frac{1}{D_n D_{n+1}}} \det \left[\begin{array}{ccc|c} I_{00} & \dots & I_{0n-1} & \alpha_0 \\ \vdots & & \vdots & \vdots \\ I_{n0} & \dots & I_{nn-1} & \alpha_n \end{array}\right]. \tag{5-5}$$

Since we know that $D_n > 0$ we need to prove the positivity of the other determinant. Determinants of this type were studied in Lemma 4.10 in [9].

We nevertheless give a complete proof of positivity. First, we observe that

$$\begin{aligned}
\pi_n \sqrt{D_{n+1} D_n} \quad &= \sum_{\sigma \in S_{n+1}} \epsilon(\sigma) \int \prod_{j=1}^{n+1} x_j^{\sigma_j - 1} \prod_{j=1}^{n} y_j^{j-1} \frac{\mathrm{d}^{n+1}\alpha \mathrm{d}^n\beta}{\prod_{j=1}^n (x_j + y_j)} = \\
&= \int \Delta(X_1^{n+1}) \prod_{j=1}^{n} y_j^{j-1} \frac{\mathrm{d}^{n+1}\alpha \mathrm{d}^n\beta}{\prod_{j=1}^n (x_j + y_j)}.
\end{aligned} \tag{5-6}$$

Here the symbol X_1^{n+1} is to remind that the vector consists of $n+1$ entries (whereas Y consists of n entries) and that the Vandermonde determinant is taken accordingly. Note also that the variable x_{n+1} never appears in the product in the denominator. Symmetrizing the integral in the x_j's with respect to labels $j = 1, \ldots, n$, but leaving x_{n+1} fixed, gives

$$\pi_n \sqrt{D_{n+1} D_n} = \frac{1}{n!} \int \Delta(X_1^{n+1}) \Delta(Y) \frac{\mathrm{d}^{n+1}\alpha \mathrm{d}^n \beta}{\prod_{j=1}^n (x_j + y_j)}. \tag{5-7}$$

Symmetrizing now with respect to the whole set $x_1, \ldots, x_{n+1}$ we obtain

$$\pi_n \sqrt{D_{n+1} D_n} = \frac{1}{n!(n+1)!} \int \Delta(X_1^{n+1}) \Delta(Y) \det \left[\begin{array}{ccc} K(x_1, y_1) & \ldots & K(x_{n+1}, y_1) \\ \vdots & & \vdots \\ K(x_1, y_n) & \ldots & K(x_{n+1}, y_n) \\ 1 & \ldots & 1 \end{array} \right] \mathrm{d}^{n+1}\alpha \mathrm{d}^n \beta \tag{5-8}$$

Moreover, since the integrand is permutation invariant, it suffices to integrate over the region $0 < x_1 < x_2 < \cdots < x_n < x_{n+1} \times 0 < y_1 < y_2 < \cdots < y_n$, and, as a result

$$\pi_n \sqrt{D_{n+1} D_n} =$$

$$\iint_{0<x_1<x_2<\cdots<x_{n+1} 0<y_1<y_2<\cdots<y_n} \Delta(X_1^{n+1}) \Delta(Y) \det \left[\begin{array}{ccc} K(x_1, y_1) & \ldots & K(x_{n+1}, y_1) \\ \vdots & & \vdots \\ K(x_1, y_n) & \ldots & K(x_{n+1}, y_n) \\ 1 & \ldots & 1 \end{array} \right] \mathrm{d}^{n+1}\alpha \mathrm{d}^n \beta. \tag{5-9}$$

We thus need to prove that the determinant containing the Cauchy kernel $\frac{1}{x+y}$ is positive for $0 < x_1 < x_2 < \cdots < x_{n+1}$ and $0 < y_1 < y_2 < \cdots < y_n$. It is not difficult to prove that

$$\det \left[\begin{array}{ccc} \frac{1}{x_1+y_1} & \cdots & \frac{1}{x_{n+1}+y_1} \\ \vdots & & \vdots \\ \frac{1}{x_1+y_n} & \ldots & \frac{1}{x_{n+1}+y_n} \\ 1 & \ldots & 1 \end{array} \right] = \frac{\Delta(X_1^{n+1}) \Delta(Y)}{\prod_{j=1}^{n+1} \prod_{k=1}^n (x_j + y_k)} \tag{5-10}$$

and this function is clearly positive in the above range. □

5.2 Interlacing properties of the zeroes

From (4-8), (4-14) and 4-15 the following factorizations are valid for all principal truncations:

$$I = S_p^{-1} (S_q^{-1})^T \ , \quad \mathbf{X} = S_p \Lambda (S_p)^{-1} \ , \quad \mathbf{Y} = S_q \Lambda S_q^{-1} \ .$$

Moreover, since I is TP, the triangular matrices S_p^{-1} and S_q^{-1} are totally nonnegative (TN) [24] and have the same diagonal entries: the nth diagonal entry being $\sqrt{D_n / D_{n-1}}$. Furthermore, one can amplify the statement about S_p^{-1} and S_q^{-1} using another result of Cryer ([25]) which implies that both triangular matrices are in fact triangular TP matrices (all notrivial in the sense defined in [25] minors are strictly positive). This has the immediate consequence

Lemma 5.1. *All principal truncations* $\mathbf{X}[n], \mathbf{Y}[n]$ *are invertible.*

Proof. From the factorization $\mathbf{X} = S_p\Lambda(S_p)^{-1}$ we conclude that it suffices to prove the claim for $\Lambda S_p^{-1}[n]$ which in matrix form reads:

$$\begin{bmatrix} (S_p^{-1})_{10} & (S_p^{-1})_{11} & & & \\ (S_p^{-1})_{20} & (S_p^{-1})_{21} & (S_p^{-1})_{22} & & \Large 0 \\ & & & \ddots & \\ \vdots & \vdots & & & (S_p^{-1})_{n+1,n+1} \\ (S_p^{-1})_{n+1,0} & (S_p^{-1})_{n+1,1} & \dots & & (S_p^{-1})_{n+1,n} \end{bmatrix}.$$

However, the determinant of this matrix is strictly positive, because S_p^{-1} is a triangular TP.

□

Remark 5.2. *This lemma is not automatic, since* $\Lambda[n]$ *is not invertible.*

We now state the main theorem of this section.

Theorem 5.1. $\mathbf{X}$ *and* $\mathbf{Y}$ *are TN.*

Proof. We need to prove the theorem for every principal truncation. Let $n \geq 0$ be fixed. We will suppress the dependence on n, for example $\mathbf{X}$ in the body of the proof means $\mathbf{X}[n]$ etc. First, we claim that $\mathbf{X}$ and $\mathbf{Y}$ admit the L-U factorization: $\mathbf{X} = \mathbf{X}_-\mathbf{X}_+,\ \mathbf{Y} = \mathbf{Y}_-\mathbf{Y}_+$, where A_+ denotes the upper triangular factor and A_- is the unipotent lower triangular factor in the Gauss factorization of a matrix A. Indeed, $\mathbf{X}_+ = (\Lambda S_p^{-1})_+,\ \mathbf{Y}_+ = (\Lambda S_q^{-1})_+$ are upper triangular components of TN matrices ΛS_p^{-1} and ΛS_q^{-1} and thus are totally nonnegative invertible bi-diagonal matrices by Lemma 5.1.

From $\mathbf{X} + \mathbf{Y}^T = \boldsymbol{\pi}\boldsymbol{\eta}^T$ we then obtain

$$(\mathbf{Y}_+^T)^{-1}\mathbf{X}_- + \mathbf{Y}_-\mathbf{X}_+^{-1} = \left((\mathbf{Y}_+^T)^{-1}\boldsymbol{\pi}\right)\left(\boldsymbol{\eta}^T\mathbf{X}_+^{-1}\right) := \boldsymbol{\rho}\boldsymbol{\mu}^T\ .$$

We need to show that vectors $\boldsymbol{\rho}$, $\boldsymbol{\mu}$ have positive entries. For this, notice that

$$\boldsymbol{\rho} = ((Y_+)^T)^{-1}S_p\boldsymbol{\alpha} = (((\Lambda S_q^{-1})_+)^T)^{-1}S_p\boldsymbol{\alpha}\ ,$$
$$\boldsymbol{\mu} = ((X_+)^T)^{-1}S_q\boldsymbol{\beta} = (((\Lambda S_p^{-1})_+)^T)^{-1}S_q\boldsymbol{\beta}.$$

Now, it is easy to check that if the matrix of generalized bimoments I is replaced by $I\Lambda^T$ (see Corollary 4.1) then $S_p \to (((\Lambda S_q^{-1})_+)^T)^{-1}S_p$, while $\boldsymbol{\alpha}$ is unchanged, which implies that $\boldsymbol{\rho}$ is a new $\boldsymbol{\pi}$ in the notation of Proposition 5.1 and hence positive by the same Proposition. Likewise, considering the matrix of generalized bimoments ΛI, for which $\boldsymbol{\beta}$ is unchanged, $S_q \to (((\Lambda S_p^{-1})_+)^T)^{-1}S_q$ and $\boldsymbol{\mu}$ is a new $\boldsymbol{\eta}$ in the notation of Proposition 5.1 implying the claim.

Thus

$$\boldsymbol{\rho} = D_\rho\mathbf{1}\ , \boldsymbol{\mu} = D_\mu\mathbf{1},$$

where D_ρ , D_μ are diagonal matrices with positive entries and $\mathbf{1}$ is a vector of 1s.

We have

$$D_\rho^{-1}(\mathbf{Y}_+^T)^{-1}\mathbf{X}_- D_\mu^{-1} + D_\rho^{-1}\mathbf{Y}_-\mathbf{X}_+^{-1}D_\mu^{-1} = \mathbf{1}\ \mathbf{1}^T \ .$$

The first (resp. second) term on the left that we can call $\tilde{\mathbf{X}}$ (resp. $\tilde{\mathbf{Y}}^T$) is a lower (resp. upper) triangular matrix with positive diagonal entries . The equality above then implies that (i) $\tilde{X}_{ij} = \tilde{Y}_{ij} = 1$ for all $i > j$ and (ii) $\tilde{X}_{ii} + \tilde{Y}_{ii} = 1$ for all i. In particular, both $\tilde{X}_{ii}$ and $\tilde{Y}_{ii}$ are positive numbers strictly less then 1.

This means that $\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}$ admits factorizations

$$\tilde{\mathbf{X}} = (Id - \Lambda^T)^{-1} L_X \ , \ \tilde{\mathbf{Y}} = (Id - \Lambda^T)^{-1} L_Y \ ,$$

where

$$L_X = \sum_{i=0}^{\infty} \tilde{X}_{ii} E_{ii} + (1 - \tilde{X}_{ii}) E_{i+1\ i} \ , L_Y = \sum_{i=0}^{\infty} \tilde{Y}_{ii} E_{ii} + (1 - \tilde{Y}_{ii}) E_{i+1\ i} \ .$$

Since all entries of bi-diagonal matrices L_X, L_Y are positive, these matrices are totally nonnegative and so are

$$\mathbf{X} = \mathbf{Y}_+^T (Id - \Lambda^T)^{-1} L_X \mathbf{X}_+ \ , \quad \mathbf{Y} = \mathbf{X}_+^T (Id - \Lambda^T)^{-1} L_Y \mathbf{Y}_+ \ . \tag{5-11}$$

□

Corollary 5.1. $\mathbf{X}$ *and* $\mathbf{Y}$ *are oscillatory matrices.*

Proof. We give a proof for $\mathbf{X}$. The factorization (5-11) we have just obtained shows that $\mathbf{X}$ is the product of an invertible lower-triangular TN matrix $\mathbf{Y}_+^T(Id - \Lambda^T)^{-1}$ and a tri-diagonal matrix $J = L_X \mathbf{X}_+$. Note that L_X has all positive values on the main diagonal and the first sub-diagonal. Entries on the first super-diagonal of $\mathbf{X}_+$ coincide with corresponding entries of $\mathbf{X}$ and thus are strictly positive by construction. Moreover, leading principal minors of $\mathbf{X}$ are strictly positive (see the proof of Lemma 5.1), which implies that all diagonal entries of $\mathbf{X}_+$ are strictly positive too. Thus J is a tri-diagonal matrix with all non-trivial entries strictly positive.

Since diagonal entries of $\mathbf{Y}_+^T(Id - \Lambda^T)^{-1}$ are strictly positive and all other entries are non-negative, every zero entry of $\mathbf{X}$ implies that the corresponding entry of J is zero. In view of that all entries on the first super- and sub-diagonals of $\mathbf{X}$ must be strictly positive, which, by a fundamental criterion of Gantmacher and Krein (Theorem 10, II, [23]), ensures that $\mathbf{X}$ is oscillatory.

□

Thus interlacing properties for zeros of polynomials p_n, q_n, as well as other properties of Sturm sequences, follow then from Gantmacher-Krein theorems on spectral properties of oscillatory matrices (see II, Theorem 13, in [23]). We summarize the most important properties implied by Gantmacher-Krein theory.

Theorem 5.2. *The sequences of BOPs $\{q_n\}$ and $\{p_n\}$ are Sturm sequences. Moreover,*

1. *their respective zeros are positive and simple,*
2. *the roots of adjacent polynomials in the sequences are interlaced,*
3. *the following alternative representations of the biorthogonal polynomials hold*

$$p_n(x) = \sqrt{\frac{D_n}{D_{n+1}}} \det(x - X[n-1]), \quad 1 \le n,$$
$$q_n(y) = \sqrt{\frac{D_n}{D_{n+1}}} \det(y - Y[n-1]), \quad 1 \le n.$$

Remark 5.3. *The fact that the roots are positive and simple follows indeed from the fact that X and Y are oscillatory. Theorem (4.5), however, indicates that this property is true even for a more general case when the totally positive kernel $K(x,y)$ is not necessarily the Cauchy kernel.*

6 Four-term recurrence relations and Christoffel Darboux identities

We establish in this section a basic form of recurrence relations and an analog of classical Christoffel-Darboux identities satisfied by $\{q_n\}$ and $\{p_n\}$. First, we introduce the following notation for semi-infinite, finite-band matrices.

Definition 6.1. *Given two integers $a \le b$, a semi-infinite matrix A is said to have the support in $[a,b]$ if*

$$j - i < a \text{ or } j - i > b \text{ imply } A_{ij} = 0 \tag{6-1}$$

The set of all matrices with supports in $[a,b]$ is denoted $M_{[a,b]}$.

The content of this section relies heavily on the relation (5-4) which we recall for convenience:

$$\mathbf{X} + \mathbf{Y}^T = \boldsymbol{\pi}\boldsymbol{\eta}^T = D_\pi \mathbf{1}\mathbf{1}^T D_\eta$$

where D_π, D_η respectively, are diagonal matrices of averages of $\mathbf{p}$ and $\mathbf{q}$. Since the vector $\mathbf{1}$ is a null vector of $\Lambda - Id$ we obtain

Proposition 6.1. $\mathbf{X}$ *and* $\mathbf{Y}$ *satisfy:*

1. $(\Lambda - Id)D_\pi^{-1}\mathbf{X} + (\Lambda - Id)D_\pi^{-1}\mathbf{Y}^T = 0.$
2. $A := (\Lambda - Id)D_\pi^{-1}\mathbf{X} \in M_{[-1,2]}.$

3. $\mathbf{X}D_\eta^{-1}(\Lambda^T - Id) + \mathbf{Y}^T D_\eta^{-1}(\Lambda^T - Id) = 0$.

4. $\widehat{A} := \mathbf{X}D_\eta^{-1}(\Lambda^T - Id) \in M_{[-2,1]}$.

It is easy to check that the bordering (maximally away from the diagonal) elements in $(\Lambda - Id)D_\pi^{-1}\mathbf{X}$ and $\mathbf{X}D_\eta^{-1}(\Lambda^T - Id)$ are nonzero. Thus

Corollary 6.1. *$\mathbf{p}$ and $\mathbf{q}$ satisfy four-term recurrence relations.*

Proof. We give the proof for $\mathbf{p}$. Indeed, from

$$x\mathbf{p} = \mathbf{X}\mathbf{p},$$

it follows that

$$x(\Lambda - Id)D_\pi^{-1}\mathbf{p} = (\Lambda - Id)D_\pi^{-1}\mathbf{X}\mathbf{p},$$

hence the claim, since $(\Lambda - Id)D_\pi^{-1} \in M_{[0,1]}$ and $(\Lambda - Id)D_\pi^{-1}\mathbf{X} \in M_{[-1,2]}$. □

Theorem 6.1 (Christoffel-Darboux Identities for $\mathbf{q}$ and $\mathbf{p}$)**.** *Let*

$$L := (\Lambda - Id)D_\pi^{-1}, \qquad \widehat{L} := D_\eta^{-1}(\Lambda^T - Id)$$

respectively, denote the multipliers used in Proposition 6.1. Moreover, let us define

$$\widetilde{\mathbf{q}}(y) = L^{-1}\mathbf{q}(y), \quad \widehat{\mathbf{p}}(x) = \widehat{L}^{-1}\mathbf{p}(x).$$

Then

$$(x+y)\sum_{j=0}^{n-1} q_j(y)p_j(x) \quad = \widetilde{\mathbf{q}}^T(y)[\Pi, L(x - \mathbf{X})]\mathbf{p}(x) \tag{6-2a}$$

$$(x+y)\sum_{j=0}^{n-1} q_j(y)p_j(x) \quad = \mathbf{q}^T(y)[\Pi, (y - \mathbf{Y}^T)\widehat{L}]\widehat{\mathbf{p}}(x) \tag{6-2b}$$

where $\Pi := \Pi_n$ is the diagonal matrix $diag(1, 1, \dots, 1, 0, \dots)$ with n ones (the entries are labeled from 0 to $n-1$). The explicit form of the commutators is:

$$[\Pi, L(x - \mathbf{X})] = -A_{n-2,n}E_{n-2,n} - A_{n-1,n+1}E_{n-1,n+1} - (A_{n-1,n} - \frac{x}{\pi_n})E_{n-1,n} - A_{n,n-2}E_{n,n-2} + A_{n,n-1}E_{n,n-1}, \tag{6-3}$$

$$[\Pi, (y - \mathbf{Y}^T)\widehat{L}] = \widehat{A}_{n-1,n}E_{n-1,n} - (\frac{y}{\eta_n} + \widehat{A}_{n,n-1})E_{n,n-1} - \widehat{A}_{n,n-2}E_{n,n-2} - \widehat{A}_{n+1,n-1}E_{n+1,n-1}, \tag{6-4}$$

where $A_{i,j}$, $\widehat{A}_{i,j}$ respectively, denote the (i,j)th entries of A, $\widehat{A}$, occurring in Proposition 6.1.

Proof. We give the proof of equation (6-2b). Since $(y - \mathbf{Y})\mathbf{q} = 0$, it suffices to prove that the left hand side equals $\mathbf{q}^T \Pi(y - Y^T)\widehat{L}\widehat{\mathbf{p}}(x)$. ¿From the definition of $\widehat{\mathbf{p}}$ and equation (4-15) we obtain

$$(x+y)\mathbf{q}^T(y)\Pi\mathbf{p}(x) = \mathbf{q}^T(y)\Pi y\widehat{L}\widehat{\mathbf{p}}(x) + \mathbf{q}^T(y)\Pi\mathbf{X}\mathbf{p}(x) = \mathbf{q}^T(y)\Pi y\widehat{L}\widehat{\mathbf{p}}(x) + \mathbf{q}^T(y)\Pi\mathbf{X}\widehat{L}\widehat{\mathbf{p}}(x),$$

which, after switching $\mathbf{X}\widehat{L}$ with $-\mathbf{Y}^T\widehat{L}$ in view of Proposition 6.1, gives equation (6-2b). To get either one of the commutator equations (6-3), (6-4), one needs to perform an elementary computation using, as appropriate, the definitions of either A or $\widehat{A}$. □

Remark 6.1. *The theory from this point onward could be developed using* $\widehat{\mathbf{p}}$*, or using* $\widetilde{\mathbf{q}}$*. We choose to work with* $\widehat{\mathbf{p}}$*.*

We establish now basic properties of $\widehat{\mathbf{p}}$ and its biorthogonal partner $\widehat{\mathbf{q}}$ defined below.

Proposition 6.2. *The sequences of polynomials*

$$\widehat{\mathbf{p}} = \widehat{L}^{-1}\mathbf{p}\ , \quad \widehat{\mathbf{q}}^T = \mathbf{q}^T\widehat{L} \tag{6-5}$$

are characterized by the following properties

1. $\deg \widehat{q}_n = n+1$, $\deg \widehat{p}_n = n$;
2. $\int \widehat{q}_n \mathrm{d}\beta = 0$;
3. $\iint \widehat{p}_n(x)\widehat{q}_m(y)\frac{\mathrm{d}\alpha \mathrm{d}\beta}{x+y} = \delta_{mn}$;
4. $\widehat{q}_n(y) = \frac{1}{\eta_{n+1}}\sqrt{\frac{D_{n+1}}{D_{n+2}}}y^{n+1} + \mathcal{O}(y^n)$;

In addition

a. $\widehat{\mathbf{q}}$ *and* $\widehat{\mathbf{p}}$ *satisfy the intertwining relations with* $\mathbf{q}$ *and* $\mathbf{p}$

$$\begin{aligned} y\widehat{\mathbf{q}}^T &= -\mathbf{q}^T\widehat{A}, \\ x\mathbf{p} &= \widehat{A}\widehat{\mathbf{p}}; \end{aligned} \tag{6-6}$$

b. $\widehat{\mathbf{q}}$ *and* $\widehat{\mathbf{p}}$ *admit the determinantal representations:*

$$\widehat{q}_n(y) = \frac{1}{\eta_n\eta_{n+1}\sqrt{D_nD_{n+2}}}\det\begin{bmatrix} I_{00} & \dots & I_{0\,n+1} \\ \vdots & & \vdots \\ I_{n-1\,0} & \dots & I_{n-1\,n+1} \\ \beta_0 & \dots & \beta_{n+1} \\ 1 & \dots & y^{n+1} \end{bmatrix} \tag{6-7}$$

$$\widehat{p}_n(x) = \frac{1}{D_{n+1}}\det\begin{bmatrix} I_{00} & \dots & I_{0\,n} & 1 \\ \vdots & & & \vdots \\ I_{n-1\,0} & \dots & I_{n-1\,n} & x^{n-1} \\ I_{n0} & \dots & I_{n\,n} & x^n \\ \beta_0 & \dots & \beta_n & 0 \end{bmatrix} \tag{6-8}$$

c. $\beta_0 \iint \widehat{p}_n(x) y^j \frac{\mathrm{d}\alpha \mathrm{d}\beta}{x+y} = \beta_j \iint \widehat{p}_n(x) \frac{\mathrm{d}\alpha \mathrm{d}\beta}{x+y}$, $j \leq n$.

Proof. Assertions (1), (2) and (4) follow directly from the shape of the matrix $\widehat{L}$. Assertion (3) follows from $\langle \mathbf{p}, \mathbf{q}^t \rangle = \mathbf{1}$ by multiplying it by $\widehat{L}$ on the right and by $\widehat{L}^{-1}$ on the left. Assertion (c) follows from assertions (1), (2) and (3); indeed from (2) and (3), it follows that the polynomial $\widehat{p}_n$ is biorthogonal to all polynomials of degree $\leq n$ with zero dβ–average and $\{\beta_0 y^j - \beta_j : 0 \leq j \leq n\}$ is a basis for such polynomials.

The intertwining relations follow from the definitions of the matrices $\widehat{L}, \widehat{A}$ and of the polynomials $\widehat{\mathbf{p}}, \widehat{\mathbf{q}}$.

The determinantal expression for $\widehat{q}_n$ follows by inspection since the proposed expression has the defining properties (1) and (2) and is biorthogonal to all powers $1, x, \ldots, x^{n-1}$. The normalization is found by comparing the leading coefficients of $\widehat{q}_n = \frac{1}{\eta_{n+1}} q_{n+1} + \mathcal{O}(y^n)$. The determinantal expression for $\widehat{p}_n(x)$ follows again by inspection; indeed if $F(x)$ is the determinant in (6-8) then

$$\langle F(x) | y^j \rangle = \det \begin{bmatrix} I_{00} & \dots & I_{0\,n} & I_{0j} \\ \vdots & & & \vdots \\ I_{n-1\,0} & \dots & I_{n-1\,n} & I_{n-1\,j} \\ I_{n0} & \dots & I_{n\,n} & I_{n\,j} \\ \beta_0 & \dots & \beta_n & 0 \end{bmatrix} = -\beta_j D_{n+1} = \frac{\beta_j}{\beta_0} \langle F(x) | 1 \rangle. \tag{6-9}$$

where the determinants are computed by expansion along the last row. The proportionality constant is again found by comparison. ☐

One easily establishes a counterpart to Theorem 6.1 valid for $\widehat{\mathbf{q}}$ and $\widehat{\mathbf{p}}$.

Proposition 6.3 (Christoffel–Darboux identities for $\widehat{\mathbf{q}}$ and $\widehat{\mathbf{p}}$)**.** *We have*

$$(x+y) \sum_{j=0}^{n-1} \widehat{q}_j(y) \widehat{p}_j(x) = \mathbf{q}^T(y) [(x - X)\widehat{L}, \Pi] \widehat{\mathbf{p}}(x) = \mathbf{q}^T(y) [\Pi, (-x - Y^T)\widehat{L}] \widehat{\mathbf{p}}(x). \tag{6-10}$$

Remark 6.2. *Observe that the commutators occurring in both theorems have identical structure; they only differ in the variable y in Theorem 6.1 being now replaced by $-x$. We will denote by $\mathbb{A}(x)$ the commutator $[\Pi, (-x - \mathbf{Y}^T)\widehat{L}]$ and by $\mathbb{A}_n(x)$ its nontrivial 3×3 block. Thus the nontrivial block in Proposition 6.3 reads:*

$$\mathbb{A}_n(x) = \left[\begin{array}{cc|c} 0 & 0 & \widehat{A}_{n-1,n} \\ \hline -\widehat{A}_{n,n-2} & \frac{x}{\eta_{n+1}} - \widehat{A}_{n,n-1} & 0 \\ 0 & -\widehat{A}_{n+1,n-1} & 0 \end{array} \right] \tag{6-11}$$

while the block appearing in Theorem 6.1 is simply $\mathbb{A}_n(-y)$.

With this notation in place we can present the Christoffel-Darboux identities in a unified way.

Corollary 6.2 (Christoffel–Darboux identities for $\mathbf{q}, \mathbf{p}$, and $\widehat{\mathbf{q}}, \widehat{\mathbf{p}}$)**.** *The biorthogonal polynomials* $\mathbf{q}, \mathbf{p}$, *and* $\widehat{\mathbf{q}}, \widehat{\mathbf{p}}$ *satisfy*

$$(x+y)\sum_{j=0}^{n-1} q_j(y)p_j(x) = \mathbf{q}^T(y)\mathbb{A}(-y)\widehat{\mathbf{p}}(x), \tag{6-12}$$

$$(x+y)\sum_{j=0}^{n-1} \widehat{q}_j(y)\widehat{p}_j(x) = \mathbf{q}^T(y)\mathbb{A}(x)\widehat{\mathbf{p}}(x). \tag{6-13}$$

7 Approximation problems and perfect duality

We will associate a chain of Markov functions with measures $\mathrm{d}\alpha$ and $\mathrm{d}\beta$. These are defined as Stieltjes' transforms of the corresponding measures. They are abstract analogs of Weyl functions discussed in earlier sections (see Definition 3.6).

Definition 7.1. *Define*

$$W_\beta(z) = \int \frac{1}{z-y}\mathrm{d}\beta(y), \qquad W_{\alpha^*}(z) = \int \frac{1}{z+x}\mathrm{d}\alpha(x),$$
$$W_{\alpha^*\beta}(z) = -\iint \frac{1}{(z+x)(x+y)}\mathrm{d}\alpha(x)\mathrm{d}\beta(y), \qquad W_{\beta\alpha^*}(z) = \iint \frac{1}{(z-y)(y+x)}\mathrm{d}\alpha(x)\mathrm{d}\beta(y). \tag{7-1}$$

We recall now an important notion of a Nikishin system associated with two measures (see [22], p. 142, called there a MT system of order 2).

Definition 7.2. *Given two measures* $\mathrm{d}\mu_1$ *and* $\mathrm{d}\mu_2$ *with disjoint supports* Δ_1, Δ_2 *respectively, a Nikishin system of order* 2 *is a pair of functions*

$$f_1(z) = \int_{\Delta_1} \frac{\mathrm{d}\mu_1(x_1)}{z-x_1} \tag{7-2}$$

$$f_2(z) = \int_{\Delta_1} \frac{\mathrm{d}\mu_1(x_1)}{z-x_1}\int_{\Delta_2} \frac{\mathrm{d}\mu_2(x_2)}{x_1-x_2}. \tag{7-3}$$

Remark 7.1. *The definition of a Nikishin system depends on the order in which one "folds" measures. If one starts from* $\mathrm{d}\mu_2$ *, rather than* $\mathrm{d}\mu_1$ *one obtains a priory a different system. As we show below the relation between these two Nikishin systems is in fact of central importance to the theory we are developing.*

The following elementary observation provides the proper framework for our discussion.

Lemma 7.1. *Let* $\mathrm{d}\alpha^*$ *denote the measure obtained from* $\mathrm{d}\alpha$ *by reflecting the support of* $\mathrm{d}\alpha$ *with respect to the origin. Then* $W_\beta, W_{\beta\alpha^*}$ *and* $W_{\alpha^*}, W_{\alpha^*\beta}$ *are Nikishin systems associated with measures* $\mathrm{d}\beta$ *and* $\mathrm{d}\alpha^*$ *with no predetermined ordering of measures.*

The relation between these two Nikishin systems can now be readily obtained.

Lemma 7.2.

$$W_\beta(z)W_{\alpha^*}(z) = W_{\beta\alpha^*}(z) + W_{\alpha^*\beta}(z). \tag{7-4}$$

Proof. Elementary computation gives:

$$W_\beta(z)W_{\alpha^*}(z) = \iint \frac{1}{(z-y)(z+x)} \mathrm{d}\alpha(x)\mathrm{d}\beta(y) = \iint \frac{1}{(x+y)}\Big[\frac{1}{z-y} - \frac{1}{z+x}\Big]\mathrm{d}\alpha(x)\mathrm{d}\beta(y),$$

which implies the claim. □

Remark 7.2. *Equation (7-4) was introduced in [9] for the DP peakons (see Lemma 4.7 there). This equation represents a generalization of the formula in Lemma 3.3 of the present paper. Observe that this formula is valid for any Nikishin system of order* 2*.*

We formulate now the main approximation problem, modeled after that of [9].

Definition 7.3. *Let $n \geq 1$. Given two Nikishin systems $W_\beta, W_{\beta\alpha^*}$ and $W_{\alpha^*}, W_{\alpha^*\beta}$ we seek polynomials $Q(z), degQ = n$, $P_\beta(z), degP_\beta = n-1$ and $P_{\beta\alpha^*}(z), degP_{\beta\alpha^*} = n-1$, which satisfy Padé-like approximation conditions as $z \to \infty$, $z \in \mathbb{C}_\pm$:*

$$Q(z)W_\beta(z) - P_\beta(z) = \mathcal{O}(\frac{1}{z}), \tag{7-5a}$$

$$Q(z)W_{\beta\alpha^*}(z) - P_{\beta\alpha^*}(z) = \mathcal{O}(\frac{1}{z}), \tag{7-5b}$$

$$Q(z)W_{\alpha^*\beta}(z) - P_\beta(z)W_{\alpha^*}(z) + P_{\beta\alpha^*}(z) = \mathcal{O}(\frac{1}{z^{n+1}}) \tag{7-5c}$$

Remark 7.3. *In the case that both measures have compact support we can remove the condition that $z \in \mathbb{C}_\pm$ since all the functions involved are then holomorphic around $z = \infty$.*

Remark 7.4. *In the terminology used, for example in [26], the triplets of polynomials $Q, P_\beta, P_{\beta\alpha^*}$ provide a Hermite-Padé approximation of type I to the Nikishin system $W_\beta, W_{\beta\alpha^*}$ and, simultaneously, a Hermite-Padé approximation of type II to the Nikishin system $W_{\alpha^*}, W_{\alpha^*\beta}$.*

Definition 7.4. *We call the right hand sides of approximation problems (7-5) $R_\beta, R_{\beta\alpha^*}$ and $R_{\alpha^*\beta}$ respectively, referring to them as remainders.*

The relation of the approximation problem (7-5) to the theory of biorthogonal polynomials $\mathbf{q}$ and $\mathbf{p}$ is the subject of the next theorem.

Theorem 7.1. *Let $q_n(y)$ be defined as in (4-12), and let us set $Q(z) = q_n(z)$ Then $Q(z)$ is the unique, up to a multiplicative constant, solution of the approximation problem (7-5). Moreover, $P_\beta, P_{\beta\alpha^*}$ and all the remainders $R_\beta, R_{\beta\alpha^*}$ and $R_{\alpha^*\beta}$ are uniquely determined from Q with the help of the formulas:*

$$P_\beta(z) = \int \frac{Q(z) - Q(y)}{z - y} \mathrm{d}\beta(y), \tag{7-6a}$$

$$P_{\beta\alpha^*}(z) = \iint \frac{Q(z) - Q(y)}{(z - y)(x + y)} \mathrm{d}\alpha(x)\mathrm{d}\beta(y), \tag{7-6b}$$

$$R_\beta(z) = \int \frac{Q(y)}{z - y} \mathrm{d}\beta(y), \tag{7-6c}$$

$$R_{\beta\alpha^*}(z) = \iint \frac{Q(y)}{(z - y)(x + y)} \mathrm{d}\alpha(x)\mathrm{d}\beta(y), \tag{7-6d}$$

$$R_{\alpha^*\beta}(z) = -\iint \frac{Q(y)}{(z + x)(x + y)} \mathrm{d}\alpha(x)\mathrm{d}\beta(y) = \int \frac{R_\beta(x)}{z - x} \mathrm{d}\alpha^*(x). \tag{7-6e}$$

Proof. We start with the first approximation problem involving $Q(z)W_\beta(z)$. Writing explicitly its first term we get:

$$\int \frac{Q(z)}{z - y} \mathrm{d}\beta(y) = \int \frac{Q(z) - Q(y)}{z - y} \mathrm{d}\beta(y) + \int \frac{Q(y)}{z - y} \mathrm{d}\beta(y).$$

Since $\int \frac{Q(z) - Q(y)}{z - y} \mathrm{d}\beta(y)$ is a polynomial in z of degree $n - 1$, while $\int \frac{Q(y)}{z - y} \mathrm{d}\beta(y) = \mathcal{O}(\frac{1}{z})$, we get the first and the third formulas. The second and fourth formulas are obtained in an analogous way from the second approximation problem. Furthermore, to get the last formula we compute P_β and $P_{\beta\alpha^*}$ from the first two approximation problems and substitute into the third approximation problem, using on the way Lemma 7.2, to obtain:

$$R_\beta W_{\alpha^*} - R_{\beta\alpha^*} = R_{\alpha^*\beta}.$$

Substituting explicit formulas for R_β and $R_{\beta\alpha^*}$ gives the final formula. To see that $Q(z)$ is proportional to $q_n(z)$ we rewrite $-R_{\alpha^*\beta}$ in the following way:

$$\iint \frac{Q(y)}{(z + x)(x + y)} \mathrm{d}\alpha(x)\mathrm{d}\beta(y) = \iint \frac{Q(y)}{(x + y)} \left[\frac{1}{z + x} - \frac{1 - (-(\frac{x}{z}))^n}{z + x} \right] \mathrm{d}\alpha(x)\mathrm{d}\beta(y) +$$
$$\iint \sum_{j=0}^{n-1} \frac{(-x)^j}{z^{j+1}} \frac{Q(y)}{(x + y)(z + x)} \mathrm{d}\alpha\mathrm{d}\beta = \iint \frac{Q(y)}{(x + y)} \left[\frac{(\frac{-x}{z})^n}{z + x} \right] \mathrm{d}\alpha(x)\mathrm{d}\beta(y) + \iint \sum_{j=0}^{n-1} \frac{(-x)^j}{z^{j+1}} \frac{Q(y)}{(x + y)(z + x)} \mathrm{d}\alpha\mathrm{d}\beta$$

To finish the argument we observe that the first term is already $\mathcal{O}(\frac{1}{z^{n+1}})$, hence the second term must vanish. This gives:

$$\iint \frac{x^j Q(y)}{x + y} \mathrm{d}\alpha(x)\mathrm{d}\beta(y) = 0, \qquad 0 \le j \le n - 1,$$

which characterizes uniquely (up to a multiplicative constant) the polynomial q_n. □

Remark 7.5. *In the body of the proof we used an equivalent form of the third approximation condition, namely*

$$R_\beta W_{\alpha^*}(z) - R_{\beta\alpha^*}(z) = R_{\alpha^*\beta}(z) = \mathcal{O}(\frac{1}{z^{n+1}}). \tag{7-7}$$

By symmetry, we can consider the Nikishin systems associated with measures α and β^* with the corresponding Markov functions $W_\alpha, W_{\alpha\beta^*}$ and $W_{\beta^*}, W_{\beta^*\alpha}$. We then have an obvious interpretation of the polynomials p_n.

Theorem 7.2. *Let $p_n(x)$ be defined as in (4-12), and let us set $Q(z) = p_n(z)$. Then $Q(z)$ is the unique, up to a multiplicative constant, solution of the approximation problem for $z \to \infty, z \in \mathbb{C}_\pm$:*

$$Q(z)W_\alpha(z) - P_\alpha(z) = \mathcal{O}(\frac{1}{z}), \tag{7-8a}$$

$$Q(z)W_{\alpha\beta^*}(z) - P_{\alpha\beta^*}(z) = \mathcal{O}(\frac{1}{z}), \tag{7-8b}$$

$$Q(z)W_{\beta^*\alpha}(z) - P_\alpha(z)W_{\beta^*}(z) + P_{\alpha\beta^*}(z) = \mathcal{O}(\frac{1}{z^{n+1}}), \tag{7-8c}$$

where $P_\alpha, P_{\alpha\beta^*}$ are given by formulas of Theorem 7.1 after switching α with β.

Clearly, one does not need to go to four different types of Nikishin systems in order to characterize q_n and p_n. The following corollary is an alternative characterization of biorthogonal polynomials which uses only the first pair of Nikishin systems.

Corollary 7.1. *Consider the Nikishin systems $W_\beta, W_{\beta\alpha^*}$ and $W_{\alpha^*}, W_{\alpha^*\beta}$. Then the pair of biorthogonal polynomials $\{q_n, p_n\}$ solves:*

1. *$Q(z) = q_n(z)$ solves Hermite-Padé approximations given by equations (7-5),*

$$Q(z)W_\beta(z) - P_\beta(z) = \mathcal{O}(\frac{1}{z}),$$

$$Q(z)W_{\beta\alpha^*}(z) - P_{\beta\alpha^*}(z) = \mathcal{O}(\frac{1}{z}),$$

$$Q(z)W_{\alpha^*\beta}(z) - P_\beta(z)W_{\alpha^*}(z) + P_{\beta\alpha^*}(z) = \mathcal{O}(\frac{1}{z^{n+1}})$$

2. *$Q(z) = p_n(-z)$ solves switched (Type I with Type II) Hermite-Padé approximations*

$$Q(z)W_{\alpha^*}(z) - P_{\alpha^*}(z) = \mathcal{O}(\frac{1}{z}), \tag{7-10a}$$

$$Q(z)W_{\alpha^*\beta}(z) - P_{\alpha^*\beta}(z) = \mathcal{O}(\frac{1}{z}), \tag{7-10b}$$

$$Q(z)W_{\beta\alpha^*}(z) - P_{\alpha^*}(z)W_\beta(z) + P_{\alpha^*\beta}(z) = \mathcal{O}(\frac{1}{z^{n+1}}) \tag{7-10c}$$

We finish this section with a few results which pave the way to the Riemann-Hilbert problem approach to biortogonal polynomials $\{q_n, p_n\}$ which will be presented in the next section.

Definition 7.5. *We define the auxiliary vectors in addition to the main polynomial vectors $\mathbf{q}_0(w) := \mathbf{q}(w)$ and $\mathbf{p}_0(z) := \mathbf{p}(z)$, as*

$$\mathbf{q}_1(w) := \int \mathbf{q}(y)\frac{\mathrm{d}\beta(y)}{w-y},$$

$$\mathbf{q}_2(w) := \int \frac{\mathbf{q}_1(x)}{w-x} \mathrm{d}\alpha^*(x), \tag{7-11}$$

$$\mathbf{p}_1(z) := \int \frac{\mathbf{p}(x)\mathrm{d}\alpha(x)}{z-x}, \tag{7-12}$$

$$\mathbf{p}_2(z) := \int \frac{\mathbf{p}_1(y)}{z-y} \mathrm{d}\beta^*(y). \tag{7-13}$$

Moreover,

$$\widehat{\mathbf{p}}_1(z) := \widehat{L}^{-1}(\mathbf{p}_1(z) + \frac{1}{\beta_0}\langle \mathbf{p}|1\rangle) = \widehat{L}^{-1}\mathbf{p}_1(z) - \mathbf{1}, \tag{7-14}$$

$$\widehat{\mathbf{p}}_2(z) := \int \frac{\widehat{\mathbf{p}}_1(y)}{z-y} \mathrm{d}\beta^*(y). \tag{7-15}$$

Here $\mathbf{1}$ *is the vector of ones.*[4].

Remark 7.6. *Note that the definition above unifies the approximants and their respective remainders (see Theorem 7.1), thus, for example,* $\mathbf{q}_1(w) = \mathbf{R}_\beta(w), \mathbf{q}_2(w) = \mathbf{R}_{\alpha^*\beta}(w)$ *etc. The definition of "hatted" quantities is justified below.*

Theorem 7.3 (Extended Christoffel-Darboux Identities)**.** *Let* $a, b = 0, \dots 2$*. Then*

$$(w+z)\mathbf{q}_a^T(w)\Pi\mathbf{p}_b(z) = \mathbf{q}_a^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_b(z) - \mathbb{F}(w,z)_{ab} \tag{7-16}$$

where

$$\mathbb{F}(w,z) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & W_{\beta^*}(z) + W_\beta(w) \\ 1 & W_\alpha(z) + W_{\alpha^*}(w) & W_{\alpha^*}(w)W_{\beta^*}(z) + W_{\alpha^*\beta}(w) + W_{\beta^*\alpha}(z) \end{bmatrix}. \tag{7-17}$$

Proof. The proof goes by repeated applications of the Christoffel-Darboux Identities given by Theorem 6.1 and Padè approximation conditions 7-5. We observe that all quantities with labels $a = 1, 2$ have asymptotic expansions around ∞ in the open half-planes $\mathbb{C}_\pm$ (they are holomorphic expansions in the case of compactly supported measures $\mathrm{d}\alpha, \mathrm{d}\beta$). We will subsequently call the part of the expansion corresponding to negative powers of z or w, of a function $f(z,w)$ the *regular* part of f and denote it $(f(z,w))_{-,z}$, $(f(z,w))_{-,w}$ respectively. In all cases the *regular* parts are obtained by subtracting certain polynomial expressions from functions holomorphic in $\mathbb{C}_\pm$ and as such the *regular* parts are holomorphic in these half-planes with vanishing limits at ∞ approach from within the respective half-planes.

We the indicate the main steps in computations for each entry, denoted below by (a, b).

(0,1):

With the help of the first approximation condition, we have

$$\mathbf{q}_1^T(w)\Pi\mathbf{p}_0(z) = \left(\int \frac{\mathbf{q}_0^T(w)\Pi\mathbf{p}_0(z)}{w-y} \mathrm{d}\beta(y)\right)_{-,w}.$$

[4]The formula $\beta_0^{-1} < \widehat{\mathbf{p}}_n, 1 >= -1$ follows directly from the determinantal expression in Proposition 6.2

Using the Christoffel-Darboux Identities and the notation of Corollary 6.2 we get

$$\mathbf{q}_1^T(w)\Pi\mathbf{p}_0(z) = \left(\int \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z)}{(w+z)(w-y)}\mathrm{d}\beta(y)\right)_{-,w} =$$
$$\int \frac{\mathbf{q}_0^T(y)\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z)}{(w+z)(w-y)}\mathrm{d}\beta(y) + \left(\int \frac{(\mathbf{q}_0^T(w)-\mathbf{q}_0^T(y))\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z)}{(w+z)(w-y)}\mathrm{d}\beta(y)\right)_{-,w},$$

where we dropped the projection sign in the first term because $\mathbb{A}(-w)$ is a polynomial of degree one. Using now the partial fraction decomposition

$$\frac{1}{(w+z)(w-y)} = \frac{1}{z+y}\left(\frac{1}{w-y} - \frac{1}{w+z}\right),$$

we get that

$$\left(\int \frac{(\mathbf{q}_0^T(w)-\mathbf{q}_0^T(y))\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z)}{(w+z)(w-y)}\mathrm{d}\beta(y)\right)_{-,w} = -\left(\int \frac{(\mathbf{q}_0^T(-z)-\mathbf{q}_0^T(y))[\Pi,(-z-\mathbf{Y}^T)\widehat{L}]\widehat{\mathbf{p}}_0(z)}{(w+z)(z+y)}\mathrm{d}\beta(y)\right)_{-,w}.$$

Observe that $(-z-\mathbf{Y}^T)\widehat{L}\widehat{\mathbf{p}}_0(z)=0$, $\mathbf{q}_0^T(-z)(-z-\mathbf{Y}^T)\widehat{L}=0$ and $\mathbf{q}_0^T(y)(-z-\mathbf{Y}^T)\widehat{L}=-(y+z)\mathbf{q}_0^T(y)\widehat{L}$ so

$$\left(\int \frac{(\mathbf{q}_0^T(w)-\mathbf{q}_0^T(y))\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z)}{(w+z)(w-y)}\mathrm{d}\beta(y)\right)_{-,w} = \left(\int \frac{(\mathbf{q}_0^T(y))(z+\mathbf{Y}^T)\widehat{L}\Pi\widehat{\mathbf{p}}_0(z)}{(w+z)(z+y)}\mathrm{d}\beta(y)\right)_{-,w} =$$
$$\int \frac{\mathbf{q}_0^T(y)\widehat{L}\Pi\widehat{\mathbf{p}}_0(z)}{w+z}\mathrm{d}\beta(y) = 0,$$

because the β averages of $\widehat{\mathbf{q}}$ are zero. Thus

$$(w+z)\mathbf{q}_1^T(w)\Pi\mathbf{p}_0(z) = \mathbf{q}_1^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z).$$

(2,0):

Using the second Padè approximation condition and biorthogonality we easily obtain

$$\mathbf{R}_{\beta\alpha^*}^T(w)\Pi\mathbf{p}_0(z) = \frac{\mathbf{R}_{\beta\alpha^*}^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z)+1}{w+z},$$

Now, substituting this formula into the formula for the third approximation condition, written as in equation (7-7), gives:

$$\mathbf{R}_{\alpha^*\beta}^T(w)\Pi\mathbf{p}_0(z) = \frac{\mathbf{R}_{\alpha^*\beta}^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z)-1}{w+z}.$$

Restoring the collective notation of $\mathbf{q}_a, \mathbf{p}_a$ we obtain :

$$(w+z)\mathbf{q}_2^T(w)\Pi\mathbf{p}_0(z) = \mathbf{q}_2^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z) - 1.$$

(0,1):

To compute $\mathbf{q}_0^T(w)\Pi\mathbf{p}_1(z)$ we use the Padè approximation conditions 7-8, in particular the first condition gives us:

$$\mathbf{q}_0^T(w)\Pi\mathbf{p}_0(z)W_\alpha(z) - \mathbf{q}_0^T(w)\Pi\mathbf{P}_\alpha(z) = \mathbf{q}_0^T(w)\Pi\mathbf{R}_\alpha(z).$$

We observe that this time we have to project on the negative powers of z. Thus the goal is to compute $\left(\mathbf{q}_0^T(w)\Pi\mathbf{p}_0(z)W_\alpha(z)\right)_{-,z}$. We have

$$\left(\int \frac{\mathbf{q}_0^T(w)\Pi\mathbf{p}_0(z)\mathrm{d}\alpha(x)}{z-x}\right)_{-,z} = \left(\int \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_0(z)\mathrm{d}\alpha(x)}{(z-x)(w+z)}\right)_{-,z} =$$
$$\left(\int \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_0(x)\mathrm{d}\alpha(x)}{(z-x)(w+z)}\right)_{-,z} + \left(\int \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)(\widehat{\mathbf{p}}_0(z)-\widehat{\mathbf{p}}_0(x))\mathrm{d}\alpha(x)}{(z-x)(w+z)}\right)_{-,z}.$$

We see that the first term is already *regular* in z. To treat the second term we perform the partial fraction expansion $\frac{1}{(z-x)(w+z)} = \frac{1}{w+x}[\frac{1}{z-x} - \frac{1}{w+z}]$ and observe that the term with $\frac{1}{z-x}$ does not contribute, while the second term

$$-\left(\int \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)(\widehat{\mathbf{p}}_0(z)-\widehat{\mathbf{p}}_0(x))\mathrm{d}\alpha(x)}{(w+x)(w+z)}\right)_{-,z} = -\left(\int \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)(\widehat{\mathbf{p}}_0(-w)-\widehat{\mathbf{p}}_0(x))\mathrm{d}\alpha(x)}{(w+x)(w+z)}\right)_{-,z} =$$
$$\int \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_0(x)\mathrm{d}\alpha(x)}{(w+x)(w+z)}.$$

Thus

$$\mathbf{q}_0^T(w)\Pi\mathbf{p}_1(z) = \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{L}^{-1}\mathbf{p}_1(z)}{w+z} - \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{L}^{-1}\mathbf{p}_1(-w)}{w+z}.$$

In other words,

$$(w+z)\mathbf{q}_0^T(w)\Pi\mathbf{p}_1(z) = \mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{L}^{-1}(\mathbf{p}_1(z)-\mathbf{p}_1(-w)).$$

More explicitly, the second term above can be rewritten as

$$-\mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{L}^{-1}\mathbf{p}_1(-w) = \mathbf{q}_0^T(w)\Pi\int \mathbf{p}(x)\mathrm{d}\alpha(x).$$

On the other hand

$$\mathbf{q}_0^T(w)\mathbb{A}(-w)\iint \frac{\widehat{\mathbf{p}}(x)\mathrm{d}\alpha(x)\mathrm{d}\beta(y)}{\beta_0(x+y)} = \mathbf{q}_0^T(w)\Pi\iint \frac{(w+x)\mathbf{p}(x)\mathrm{d}\alpha(x)\mathrm{d}\beta(y)}{\beta_0(x+y)} =$$
$$\mathbf{q}_0^T(w)\Pi\int \mathbf{p}(x)\mathrm{d}\alpha(x) + \mathbf{q}_0^T(w)\Pi\iint \frac{(w-y)\mathbf{p}(x)\mathrm{d}\alpha(x)\mathrm{d}\beta(y)}{\beta_0(x+y)}.$$

Now the second term $\mathbf{q}_0^T(w)\Pi\iint \frac{(w-y)\mathbf{p}(x)\mathrm{d}\alpha(x)\mathrm{d}\beta(y)}{\beta_0(x+y)} = 0$ because $\mathbf{q}_0^T(w)\Pi\langle\mathbf{p}(x)|\bullet\rangle$ is a projector on polynomials of degree $\le n-1$ and thus $w\mathbf{q}_0^T(w)\Pi\langle\mathbf{p}(x)|1\rangle - \mathbf{q}_0^T(w)\Pi\langle\mathbf{p}(x)|y\rangle = w - w = 0$, hence

$$(w+z)\mathbf{q}_0^T(w)\Pi\mathbf{p}_1(z) = \mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_1(z),$$

where $\widehat{\mathbf{p}}_1(z) = \widehat{L}^{-1}(\mathbf{p}_1(z) + \frac{1}{\beta_0}\langle\mathbf{p}|1\rangle)$ as advertised earlier.

$(1,1)$:

We use again the Padè approximation conditions 7-8, this time multiplying on the left by $\mathbf{q}_1^T(w)\Pi$ and projecting on the negative powers of z , to obtain:

$$\left(\mathbf{q}_1^T(w)\Pi\mathbf{p}_0(z)W_\alpha(z)\right)_{-,z} = \mathbf{q}_1^T(w)\Pi\mathbf{p}_1(z).$$

With the help of the result for the $(0,1)$ entry, after carrying out the projection, we obtain

$$(w+z)\mathbf{q}_1^T(w)\Pi\mathbf{p}_1(z) = \mathbf{q}_1^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_1(z) + \mathbf{q}_1^T(w)\mathbb{A}(-w)\left(\int \frac{\widehat{\mathbf{p}}(x)\mathrm{d}\alpha(x)}{w+x} - \frac{1}{\beta_0}\langle\widehat{\mathbf{p}}|1\rangle\right).$$

We claim that

$$\mathbf{q}_1^T(w)\mathbb{A}(-w)\left(\int \frac{\widehat{\mathbf{p}}(x)\mathrm{d}\alpha(x)}{w+x} - \frac{1}{\beta_0}\langle\widehat{\mathbf{p}}|1\rangle\right) = -1.$$

Indeed, the left hand side of the equation equals:

$$\frac{1}{\beta_0}\mathbf{q}_1^T(w)\Pi\iint \frac{(y-w)\mathbf{p}(x)\mathrm{d}\alpha(x)\mathrm{d}\beta(y)}{x+y} = \frac{1}{\beta_0}\int \frac{\mathbf{q}_0^T(\xi)}{w-\xi}\Pi(\langle\mathbf{p}|y\rangle - w\langle\mathbf{p}|1\rangle)\mathrm{d}\beta(\xi) =$$
$$\frac{1}{\beta_0}\int \frac{\xi-w}{w-\xi}\mathrm{d}\beta(\xi) = -1.$$

Thus

$$(w+z)\mathbf{q}_1^T(w)\Pi\mathbf{p}_1(z) = \mathbf{q}_1^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_1(z) - 1.$$

$(2,1)$:

This time we use projections in both variables, one at a time, and compare the results. First, let us use the projections in z. Thus

$$\mathbf{q}_2^T(w)\Pi\mathbf{p}_1(z) = \left(\mathbf{q}_2^T(w)\Pi\mathbf{p}_0(z)W_\alpha(z)\right)_{-,z}.$$

Carrying out all the projections we obtain an expression of the form:

$$\mathbf{q}_2^T(w)\Pi\mathbf{p}_1(z) = \frac{\mathbf{q}_2^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_1(z)}{w+z} - \frac{W_\alpha(z)+F(w)}{w+z}.$$

Observe that, since $\mathbf{q}_2^T(w)$ is $\mathcal{O}(1/w)$ and the first term on the right is much smaller, $F(w) = \mathcal{O}(1)$. More precisely, by comparing the terms at $1/w$ on both sides, we conclude that in fact, $F(w) = \mathcal{O}(1/w)$. Now, we turn to the projection in w, resulting in an expression of the form:

$$\mathbf{q}_2^T(w)\Pi\mathbf{p}_1(z) = \frac{\mathbf{q}_2^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_1(z)}{w+z} - \frac{W_{\alpha^*}(w)+G(z)}{w+z}.$$

This, and the fact that $F(w) = \mathcal{O}(1/w)$, implies that $F(w) = W_{\alpha^*}(w), G(z) = W_\alpha(z)$. Hence

$$(w+z)\mathbf{q}_2^T(w)\Pi\mathbf{p}_1(z) = \mathbf{q}_2^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_1(z) - (W_\alpha(z)+W_{\alpha^*}(w)).$$

(0, 2):

We use the projection in the z variable and the fact that by the Padè approximation condition (7-7), after exchanging α with β, $\mathbf{p}_2(z) = \mathbf{p}_1(z)W_{\beta^*}(z) - \mathbf{R}_{\alpha\beta^*}(z)$. Using the result for the $(0,1)$ entry we obtain:

$$\mathbf{q}_0^T(w)\Pi\mathbf{p}_2(z) = \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\mathbf{p}_1(z)W_{\beta^*}(z)}{w+z} - \left(\frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\mathbf{p}_0(z)W_{\alpha\beta^*}(z)}{w+z}\right)_{-,z}.$$

Carrying out the projection and reassembling terms according to the definition of $\widehat{\mathbf{p}}_2(z)$ we obtain:

$$\mathbf{q}_0^T(w)\Pi\mathbf{p}_2(z) = \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_2(z)}{w+z} - \frac{\mathbf{q}_0^T(w)\Pi\langle\mathbf{p}_0|1\rangle}{w+z} = \frac{\mathbf{q}_0^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_2(z)}{w+z} - \frac{1}{w+z}.$$

(1, 2):

We use the projection in the z variable and the Padè approximation condition $\mathbf{p}_2(z) = \mathbf{p}_1(z)W_{\beta^*}(z) - \mathbf{R}_{\alpha\beta^*}(z)$.

Consequently,

$$\begin{aligned}&\mathbf{q}_1^T(w)\Pi\mathbf{p}_2(z) = \mathbf{q}_1^T(w)\Pi\mathbf{p}_1(z)W_{\beta^*}(z) - \mathbf{q}_1^T(w)\Pi\mathbf{R}_{\alpha\beta^*}(z) = \\ &\left(\frac{\mathbf{q}_1^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_1(z)-1}{w+z}\right)W_{\beta^*(z)} - \left(\mathbf{q}_1^T(w)\Pi\mathbf{p}_0(z)W_{\alpha\beta^*}(z)\right)_{-,z}.\end{aligned}$$

Using the existing identities and carrying out the projection in the second term we obtain:

$$(w+z)\mathbf{q}_1^T(w)\Pi\mathbf{p}_2(z) = \mathbf{q}_1^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_2(z) - W_{\beta^*}(z) - W_\beta(w).$$

(2, 2):

The computation is similar to the one for $(1,2)$ entry; we use both projections. The projection in the z variable gives:

$$\mathbf{q}_2^T(w)\Pi\mathbf{p}_2(z) = \frac{\mathbf{q}_2^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_2(z)}{w+z} + \frac{F(w) - (W_{\alpha^*}(w) + W_\alpha(z))W_{\beta^*}(z) + W_{\alpha\beta^*}(z)}{w+z}.$$

On the other hand, carrying out the projection in the w variable we obtain:

$$\mathbf{q}_2^T(w)\Pi\mathbf{p}_2(z) = \frac{\mathbf{q}_2^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_2(z)}{w+z} + \frac{G(z) - (W_\beta(w) + W_{\beta^*}(z))W_{\alpha^*}(w) + W_{\beta\alpha^*}(w)}{w+z}.$$

Upon comparing the two expressions and using Lemma 7.2 we obtain $F(w) = -W_{\alpha^*\beta}(w)$, hence

$$\begin{aligned}&(w+z)\mathbf{q}_2^T(w)\Pi\mathbf{p}_2(z) = \mathbf{q}_2^T(w)\mathbb{A}(-w)\widehat{\mathbf{p}}_2(z) - W_{\alpha^*\beta}(w) - (W_{\alpha^*}(w) + W_\alpha(z))W_{\beta^*}(z) + W_{\alpha\beta^*}(z) = \\ &\mathbf{q}_2^T(w)[\Pi, \mathbb{A}(-w)]\widehat{\mathbf{p}}_2(z) - (W_{\alpha^*}(w)W_{\beta^*}(z) + W_{\alpha^*\beta}(w) + W_{\beta^*\alpha}(z)),\end{aligned}$$

where in the last step we used again Lemma 7.2. □

We point out that if we set $w = -z$ in the CDI's contained in Theorem 7.3, the left hand side vanishes identically and the RHS contains terms of the form $\mathbf{q}_a(-z)\mathbb{A}(z)\widehat{\mathbf{p}}_b(z)$ minus $\mathbf{F}_{ab}(-z,z)$. The main observation is that the second term is **constant**, independent of both z and n, and hence one ends up with the **perfect pairing** (see [15]) between the auxiliary vectors. For the reader's convenience we recall the definition of $\mathbb{A}(z)$ to emphasize the implicit dependence on the index n hidden in the projection Π.

Theorem 7.4. *(Perfect Duality)*

Let

$$\mathbb{J} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Then

$$\mathbf{q}_a^T(-z)\mathbb{A}(z)\widehat{\mathbf{p}}_b(z) = \mathbb{J}_{ab},$$

where

$$\mathbb{A}(z) = [(z-\mathbf{X})\widehat{L}, \Pi].$$

Proof. The only nontrivial entry to check is $(2,2)$. In this case, after one substitutes $w = -z$ into $W_{\alpha^*}(w)W_{\beta^*}(z) + W_{\alpha^*\beta}(w) + W_{\beta^*\alpha}(z)$, one obtains the identity of Lemma 7.2. □

There also exists an analog of the extended Christoffel-Darboux identities of Theorem 7.3 for the "hatted" quantities.

We first define:

Definition 7.6. *For $a = 0, 1, 2$,*

$$\widehat{\mathbf{q}}_a^T := \mathbf{q}_a^T\widehat{L}. \tag{7-18}$$

The following identities follow directly from the respective definitions.

Lemma 7.3.

$$w\widehat{\mathbf{q}}_a^T(w) = \begin{cases} \mathbf{q}_a^T(w)\mathbf{Y}^T\widehat{L}, & a = 0,1 \\ \mathbf{q}_2^T(w)\mathbf{Y}^T\widehat{L} - \langle 1|\widehat{\mathbf{q}}_0^T\rangle, & a = 2. \end{cases}$$

$$(z-\mathbf{X})\widehat{L}\widehat{\mathbf{p}}_b(z) = \begin{cases} 0, & b = 0, \\ \frac{\langle \mathbf{p}_0|z+y\rangle}{\beta_0}, & b = 1, \\ -\langle \mathbf{p}_0|1\rangle + \frac{\langle \mathbf{p}_0|z+y\rangle W_{\beta^*}(z)}{\beta_0}, & b = 2. \end{cases}$$

Theorem 7.5 (Extended Christoffel-Darboux Identities)**.** *Let $a, b = 0, \dots 2$. Then*

$$(w+z)\widehat{\mathbf{q}}_a^T(w)\Pi\widehat{\mathbf{p}}_b(z) = \mathbf{q}_a^T(w)\mathbb{A}(z)\widehat{\mathbf{p}}_b(z) - \widehat{\mathbb{F}}(w,z)_{ab} \tag{7-19}$$

where

$$\widehat{\mathbb{F}}(w,z) = \mathbb{F}(w,z) - \frac{w+z}{\beta_0}\begin{bmatrix} 0 & 1 & W_{\beta^*}(z) \\ 0 & W_\beta(z) & W_\beta(w)W_{\beta^*}(z) \\ 1 & W_{\alpha^*\beta^*}(w) & W_{\alpha^*\beta^*}(w)W_{\beta^*}(z) \end{bmatrix}. \tag{7-20}$$

Proof. We give an outline of the proof. For $a = 0, 1$, in view of Lemma 7.3

$$(w+z)\widehat{\mathbf{q}}_a^T(w)\Pi\widehat{\mathbf{p}}_b(z) = \mathbf{q}_a^T(w)\mathbb{A}(z)\widehat{\mathbf{p}}_b(z) + \mathbf{q}_a^T(w)\Pi(z-\mathbf{X})\widehat{L}\widehat{\mathbf{p}}_b(z).$$

The second term equals, again by Lemma 7.3,

$$\mathbf{q}_a^T(w)\Pi \begin{cases} 0, & b=0, \\ \frac{\langle \mathbf{p}_0|z+y\rangle}{\beta_0}, & b=1, \\ -\langle \mathbf{p}_0|1\rangle + \frac{\langle \mathbf{p}_0|z+y\rangle W_{\beta^*}(z)}{\beta_0}, & b=2. \end{cases}$$

Now, one goes case by case, using biorthogonality of $\mathbf{q}_0^T$ and $\mathbf{p}_0$, and the definition of $\mathbf{q}_1^T(w)$. After a few elementary steps one arrives at the claimed result. The computation for $a = 2$ is only slightly more involved. From Lemma 7.3 we obtain:

$$(w+z)\widehat{\mathbf{q}}_2^T(w)\Pi\widehat{\mathbf{p}}_b(z) = \mathbf{q}_2^T(w)\mathbb{A}(z)\widehat{\mathbf{p}}_b(z) - \langle 1|\widehat{\mathbf{q}}_0\rangle\Pi\widehat{\mathbf{p}}_b(z) + \mathbf{q}_2^T(w)\Pi(z-\mathbf{X})\widehat{L}\widehat{\mathbf{p}}_b(z).$$

In view of biorthogonality of $\widehat{\mathbf{q}}_0^T$ and $\widehat{\mathbf{p}}$, after some intermediate computations, one obtains:

$$\langle 1|\widehat{\mathbf{q}}_0\rangle\Pi\widehat{\mathbf{p}}_b(z) = \begin{cases} 1, & b=0 \\ W_\alpha(z) + \frac{\langle 1|1\rangle}{\beta_0}, & b=1, \\ W_{\beta^*\alpha}(z) + \frac{\langle 1|1\rangle}{\beta_0}W_{\beta^*}(z), & b=2. \end{cases}$$

Likewise,

$$\mathbf{q}_2^T(w)\Pi(z-\mathbf{X})\widehat{L}\widehat{\mathbf{p}}_b(z) = \begin{cases} 0, & b=0 \\ \frac{w+z}{\beta_0}W_{\alpha^*\beta}(w) - W_{\alpha^*}(w) + \frac{\langle 1|1\rangle}{\beta_0}, & b=1, \\ \frac{w+z}{\beta_0}W_{\beta^*}(z)W_{\alpha^*}(w) - W_{\alpha^*\beta}(w) - W_{\beta^*}(z)W_{\alpha^*}(w) + \frac{\langle 1|1\rangle}{\beta_0}W_{\beta^*}(z), & b=2, \end{cases}$$

and the claim follows.

□

8 Riemann–Hilbert problems

In this section we set up two Riemann–Hilbert problems characterizing the Cauchy BOPs that enter the Christoffel–Darboux identities of the previous section. This is done in anticipation of possible applications to the study of universality for the corresponding two–matrix model. Moreover, since the Christoffel–Darboux kernels contain also the hatted polynomials, it is useful to formulate the Riemann–Hilbert problems for those polynomials as well.

We will also make the **assumption** (confined to this section) that the measures $\mathrm{d}\alpha, \mathrm{d}\beta$ are *absolutely continuous with respect to Lebesgue's measure* on the respective axes. Thus one can write

$$\frac{\mathrm{d}\alpha}{\mathrm{d}x} = \mathrm{e}^{-\frac{U(x)}{\hbar}}, \qquad \frac{\mathrm{d}\beta}{\mathrm{d}y} = \mathrm{e}^{-\frac{V(y)}{\hbar}}, \tag{8-1}$$

for the respective (positive!) densities on the respective supports: the signs in the exponents are conventional so as to have (in the case of an unbounded support) the *potentials* U, V bounded from below. The constant $\hbar$ is only for convenience when studying the asymptotics of biorthogonal polynomials for large degrees (small $\hbar$).

Since the Christoffel–Darboux identities involve the expressions $\mathbf{q}_a \mathbb{A} \widehat{\mathbf{p}}_b$, we are naturally led to characterize the sequences $\mathbf{q}$ and $\widehat{\mathbf{p}}$. However, the other sequences can be characterized in a similar manner by swapping the rôles of the relevant measures and symbols.

8.1 Riemann–Hilbert problem for the q–BOPs

We will be describing here only the RHP characterizing the polynomials $q_n(y)$, where the characterization of the polynomials $p_n(x)$ is obtained by simply interchanging α with β (see for example Theorem 7.2).

We consider the real axis $\mathbb{R}$ oriented as usual and define

$$\begin{aligned} \vec{\mathbf{q}}_0^{(n)}(w) &:= \begin{bmatrix} q_{n-2}(w) \\ q_{n-1}(w) \\ q_n(w) \end{bmatrix} , \\ \vec{\mathbf{q}}_1^{(n)}(w) &:= \int \vec{\mathbf{q}}^{(n)}(y) \frac{\mathrm{d}\beta(y)}{w-y} , \\ \vec{\mathbf{q}}_2^{(n)}(w) &:= \int \vec{\mathbf{q}}_1^{(n)}(x) \frac{\mathrm{d}\alpha^*(x)}{w-x} \end{aligned} \tag{8-2}$$

For simplicity of notation we will suppress the superscript $^{(n)}$ in most of the following discussions, only to restore it when necessary for clarity; the main point is that an arrow on top of the corresponding vector will denote a "window" of three consecutive entries of either the ordinary vector $\mathbf{q}$ (index $a = 0$), or the auxiliary vectors $\mathbf{q}_a$ (index $a = 1, 2$, see Def. 7.5) which, as we might recall at this point, combine the polynomials and the corresponding remainders in the Hermite-Padè approximation problem given by Theorem 7.1. Some simple observations are in order;

- the vector $\vec{\mathbf{q}}_1(w)$ is an analytic vector which has a jump–discontinuity on the support of $\mathrm{d}\beta$ contained in the positive real axis. As $w \to \infty$ (away from the support of $\mathrm{d}\beta$) it decays as $\frac{1}{w}$. Its jump-discontinuity is (using Plemelj formula)

$$\vec{\mathbf{q}}_1(w)_+ = \vec{\mathbf{q}}_1(w)_- - 2\pi i \frac{\mathrm{d}\beta}{\mathrm{d}w} \vec{\mathbf{q}}_0(w) \ , \quad w \in supp(\mathrm{d}\beta). \tag{8-3}$$

  Looking at the leading term at $w = \infty$ we see that

$$\vec{\mathbf{q}}_1(w) = \begin{bmatrix} \eta_{n-2} \\ \eta_{n-1} \\ \eta_n \end{bmatrix} \frac{1}{w} + \mathcal{O}(1/w^2) \ . \tag{8-4}$$

- The vector $\vec{\mathbf{q}}_2(w)$ is also analytic with a jump discontinuity on the **reflected support** of $\mathrm{d}\alpha$ (i.e. on $supp(\mathrm{d}\alpha^*)$). In view of Theorem 7.1, recalling that $\mathbf{q}_2$ are remainders of the Hermite-Padè approximation problem of type II, we easily see that

$$\vec{\mathbf{q}}_2(w) = \begin{bmatrix} \frac{c_{n-2}}{(-w)^{n-1}} \\ \frac{c_{n-1}}{(-w)^{n}} \\ \frac{c_{n}}{(-w)^{n+1}} \end{bmatrix} (1 + \mathcal{O}(1/w))$$

$$c_n := \langle x^n | q_n \rangle = \sqrt{\frac{D_{n+1}}{D_n}} > 0. \tag{8-5}$$

The jump-discontinuity of $\vec{\mathbf{q}}_2$ is

$$\vec{\mathbf{q}}_2(w)_+ = \vec{\mathbf{q}}_2(w)_- - 2\pi i \frac{\mathrm{d}\alpha^*}{\mathrm{d}w} \vec{\mathbf{q}}_1(w) \qquad w \in supp(d\alpha^*). \tag{8-6}$$

- The behavior of $\vec{\mathbf{q}}_0(w)$ at infinity is

$$\vec{\mathbf{q}}_0(w) = \begin{bmatrix} \frac{w^{n-2}}{c_{n-2}} \\ \frac{w^{n-1}}{c_{n-1}} \\ \frac{w^{n}}{c_{n}} \end{bmatrix} (1 + \mathcal{O}(1/w)), \tag{8-7}$$

with the same c_n's as in 8-5.

Define the matrix

$$\Gamma(w) := \overbrace{\begin{bmatrix} 1 & -c_n\eta_n & 0 \\ 0 & 1 & 0 \\ 0 & (-1)^{n-1}\frac{\eta_{n-2}}{c_{n-2}} & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & c_n \\ 0 & \frac{1}{\eta_{n-1}} & 0 \\ \frac{(-1)^n}{c_{n-2}} & 0 & 0 \end{bmatrix}}^{=:\mathcal{N}_q} [\vec{\mathbf{q}}_0^{(n)}(w), \vec{\mathbf{q}}_1^{(n)}(w), \vec{\mathbf{q}}_2^{(n)}(w)] \tag{8-8}$$

Proposition 8.1. *The matrix $\Gamma(w)$ is analytic on $\mathbb{C} \setminus (supp(\mathrm{d}\beta) \cup supp(\mathrm{d}\alpha^*)$. Moreover, it satisfies the jump conditions*

$$\begin{aligned} \Gamma(w)_+ &= \Gamma(w)_- \begin{bmatrix} 1 & -2\pi i \frac{\mathrm{d}\beta}{\mathrm{d}w} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad w \in supp(\mathrm{d}\beta) \subset \mathbb{R}_+ \\ \Gamma(w)_+ &= \Gamma(w)_- \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2\pi i \frac{\mathrm{d}\alpha^*}{\mathrm{d}w} \\ 0 & 0 & 1 \end{bmatrix}, \qquad w \in supp(\mathrm{d}\alpha^*) \subset \mathbb{R}_- \end{aligned} \tag{8-9}$$

and its asymptotic behavior at $w=\infty$ is

$$\Gamma(w) = (\mathbf{1} + \mathcal{O}(w^{-1}))\begin{bmatrix} w^n & 0 & \\ 0 & w^{-1} & 0 \\ 0 & 0 & w^{-n+1} \end{bmatrix} \tag{8-10}$$

Moreover, $\Gamma(w)$ can be written as:

$$\Gamma(w) = \begin{bmatrix} c_n\eta_n & 0 & 0 \\ 0 & \frac{1}{\eta_{n-1}} & 0 \\ 0 & 0 & \frac{(-1)^{n-1}\eta_{n-2}}{c_{n-2}} \end{bmatrix}\begin{bmatrix} \widehat{q}_{n-1} & \widehat{q}_{1,n-1} & \widehat{q}_{2,n-1} \\ q_{n-1} & q_{1,n-1} & q_{2,n-1} \\ \widehat{q}_{n-2} & \widehat{q}_{1,n-2} & \widehat{q}_{2,n-2} \end{bmatrix}. \tag{8-11}$$

Proof. All the properties listed are obtained from elementary matrix computations.

□

Remark 8.1. *An analogous problem with the rôles of α, β, etc., interchanged, characterizes the monic orthogonal polynomials $p_{n-1}(x)$ of degree $n-1$ in x.*

Corollary 8.1. *Given $n \in \mathbb{N}$, the absolutely continuous measures $\mathrm{d}\beta \subset \mathbb{R}_+$ and $\mathrm{d}\alpha^* \subset \mathbb{R}_-$, and assuming the existence of all the bimomoments I_{ij} there exists a unique matrix $\Gamma(w)$ solving the RHP specified by equations (8-9), (8-10). The solution characterizes uniquely the polynomials q_{n-1} as well as $\widehat{q}_{n-1}$. In particular, the normalization constants c_{n-1}, η_{n-1} (i.e. the "norm" of the monic orthogonal polynomials and the β average of the q_{n-1}) are read off the following expansions*

$$\Gamma_{2,1}(w) = \frac{1}{c_{n-1}\eta_{n-1}} w^{n-1} + \mathcal{O}(w^{n-2}) \ , \tag{8-12}$$

$$\Gamma_{2,3}(w) = (-1)^n \frac{c_{n-1}}{\eta_{n-1} w^n} + \mathcal{O}(w^{-n-1}) \tag{8-13}$$

or, equivalently,

$$\begin{aligned} \frac{1}{\eta_{n-1}^2} &= (-1)^n \lim_{w\to\infty} w\Gamma_{2,1}(w)\Gamma_{2,3}(w), \\ c_{n-1}^2 &= (-1)^n \lim_{w\to\infty} w^{2n-1}\frac{\Gamma_{2,3}(w)}{\Gamma_{2,1}(w)}. \end{aligned} \tag{8-14}$$

Proof. Given $\mathrm{d}\beta$ and $\mathrm{d}\alpha^*$ it suffices to construct the Nikishin systems $W_\beta, W_{\beta\alpha^*}$ and $W_{\alpha^*}, W_{\alpha^*\beta}$ followed by solving the Hermite-Padé approximation problems given by equations (7-5). The existence of the solution is ensured by the existence of all bimoments I_{ij} (see equation (4-5) for the definition). Then one constructs the polynomials $\widehat{q}_j$, finally the matrix $\Gamma(w)$ using equation (8-11). By construction $\Gamma(w)$ satisfies the Riemann-Hilbert factorization problem specified by equations (8-9) and (8-10). Since the determinant of $\Gamma(w)$ is constant in w (and equal to one), the solution of the Riemann–Hilbert problem is unique. The formulas for η_{n-1} and c_{n-1} follow by elementary matrix computations. □

8.1.1 A Riemann–Hilbert problem with constant jumps

Let us recall that

$$\frac{\mathrm{d}\alpha}{\mathrm{d}x} = \mathrm{e}^{-\frac{1}{\hbar}U(x)} \ , \quad x \in supp(\mathrm{d}\alpha) \qquad \frac{\mathrm{d}\beta}{\mathrm{d}y} = \mathrm{e}^{-\frac{1}{\hbar}V(y)} \ , \quad y \in supp(\mathrm{d}\beta). \tag{8-15}$$

In order to modify the RHP into one with constant jumps we make the (restrictive) assumption (only for this subsection) that the *potentials* can be extended to analytic functions off the real axis. An example is if $U(x), V(y)$ are real–analytic functions. The matrix

$$\mathbb{Y}(w) := \Gamma(w) \begin{bmatrix} \exp\left(-\frac{2V+U^\star}{3\hbar}\right) & 0 & 0 \\ 0 & \exp\left(\frac{V-U^\star}{3\hbar}\right) & 0 \\ 0 & 0 & \exp\left(\frac{2U^\star+V}{3\hbar}\right) \end{bmatrix} \tag{8-16}$$

solves a similar RH problem but with **constant** jump-discontinuity (and still with unit determinant); hence one can conclude immediately that it solves a linear ODE in the complex plane (or in the maximal domain of meromorphicity of U, V). The detailed singularity structure of this ODE depends on the analyticity properties of the potentials but the issue is of no relevance for the moment. For similar situation in the ordinary OP case see [27].

This new RHP exhibits the following jumps

$$\begin{aligned} \mathbb{Y}_+(w) &= \mathbb{Y}_-(w) \begin{bmatrix} 1 & -2\pi i & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \ , \quad w \in supp(\mathrm{d}\beta) \\ \mathbb{Y}_+(w) &= \mathbb{Y}_-(w) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2\pi i \\ 0 & 0 & 1 \end{bmatrix} \ , \quad w \in supp(\mathrm{d}\alpha^*) \end{aligned} \tag{8-17}$$

Remark 8.2. *Since the jumps are constants it follows straightforwardly that $\mathbb{Y}(w)$ solves a linear ODE with the same singularities as $V', U^{\star\prime}$; for example if U', V' are rational functions then so is the coefficient matrix of the ODE and the orders of poles do not exceed those of V', U'. It would be possible to express this ODE directly in terms of the coefficients of the recurrence relations using methods already exploited in [28, 29, 30, 15]; such expressions should be used in proving that the principal minors D_n of the matrix of bimoments I_{ij} are also isomonodromic tau–functions in the sense of Jimbo–Miwa–Ueno [31]. In a certain sense this is to be expected* a priori *because the vanishing of the isomonodromic tau function characterizes the non–solvability of the Riemann–Hilbert problem, i.e. the (non)–existence of the BOPs, exactly as the principal minors of I do.*

8.2 Riemann–Hilbert problem for the $\widehat{\mathrm{p}}$–BOPs

Referring to the defining properties of $\widehat{p}_n(x)$ as indicated in Prop. 6.2 we are going to define a second 3×3 local RHP that characterizes them.

Define

$$\vec{\widehat{\mathbf{p}}}_0(z) := \begin{bmatrix} \widehat{p}_{n-2}(z) \\ \widehat{p}_{n-1}(z) \\ \widehat{p}_n(z) \end{bmatrix} \tag{8-18}$$

and $\vec{\widehat{\mathbf{p}}}_{1,2}(z)$ as the same **windows** of the auxiliary vectors $\widehat{\mathbf{p}}_{1,2}$ introduced in Definition 7.5. We first study the large z asymptotic behavior of $\widehat{p}_{0,n}(z), \widehat{p}_{1,n}(z), \widehat{p}_{2,n}(z)$.

Lemma 8.1. *The asymptotic behavior at $z \to \infty, z \in \mathbb{C}_\pm$ is given by:*

$$\widehat{p}_{0,n}(z) = -\frac{\eta_n}{c_n} z^n (1 + \mathcal{O}(1/z)), \tag{8-19}$$

$$\widehat{p}_{1,n}(z) = -1 + \mathcal{O}(1/z), \tag{8-20}$$

$$\widehat{p}_{2,n}(z) = (-1)^n \frac{c_{n+1}\eta_{n+1}}{z^{n+2}} (1 + \mathcal{O}(1/z)). \tag{8-21}$$

Proof. We give a proof for $\widehat{p}_{1,n}(z) = \int \frac{\widehat{p}_{0,n}(x)}{z-x} \mathrm{d}\alpha(x) + \frac{1}{\beta_0} \langle \widehat{p}_{0,n} | 1 \rangle$. The first term is $\mathcal{O}(\frac{1}{z})$, while the second term can be computed using biorthogonality and the fact that $\widehat{p}_{0,n} = -(\eta_n p_{0,n} + \eta_{n-1} p_{0,n-1} + \cdots + \eta_0 p_{0,0})$. Thus the second term equals $-\frac{\eta_0}{\beta_0} \langle p_{0,0} | 1 \rangle = -1$, since $\eta_0 = q_0 \beta_0$, hence the claim for $\widehat{p}_{1,n}(z)$ follows. The remaining statements are proved in a similar manner. □

For reasons of normalization, and in full analogy with equation (8-8), we arrange the window of all $\widehat{\mathbf{p}}$s wave vectors into the matrix

$$\widehat{\Gamma}(z) = \overbrace{\begin{bmatrix} 0 & 0 & -\frac{c_n}{\eta_n} \\ 0 & -1 & 0 \\ \frac{(-1)^n}{c_{n-1}\eta_{n-1}} & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}}^{=:\mathcal{N}_{\widehat{p}}} \left[\vec{\widehat{\mathbf{p}}}(z), \vec{\widehat{\mathbf{p}}}_1(z), \vec{\widehat{\mathbf{p}}}_2(z) \right]. \tag{8-22}$$

Proposition 8.2. *The matrix $\widehat{\Gamma}(z)$ is analytic in $\mathbb{C} \setminus supp(\mathrm{d}\alpha) \cup supp(\mathrm{d}\beta^*)$. Moreover, it satisfies the jump conditions*

$$\begin{aligned} \widehat{\Gamma}(z)_+ &= \widehat{\Gamma}(z)_- \begin{bmatrix} 1 & -2\pi i \frac{\mathrm{d}\alpha}{\mathrm{d}z} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad z \in supp(\mathrm{d}\alpha) \subseteq \mathbb{R}_+ \\ \widehat{\Gamma}(z)_+ &= \widehat{\Gamma}(z)_- \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2\pi i \frac{\mathrm{d}\beta^*}{\mathrm{d}z} \\ 0 & 0 & 1 \end{bmatrix}, \quad z \in supp(\mathrm{d}\beta^*) \subseteq \mathbb{R}_-, \end{aligned} \tag{8-23}$$

and its asymptotic behavior at $z = \infty$ is

$$\widehat{\Gamma}(z) = \left(\mathbf{1} + \mathcal{O}\left(\frac{1}{z}\right) \right) \begin{bmatrix} z^n & 0 & \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{z^n} \end{bmatrix}. \tag{8-24}$$

$\widehat{\Gamma}(z)$ *can be written as:*

$$\Gamma(z) = \begin{bmatrix} c_n & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{(-1)^n}{c_{n-1}} \end{bmatrix} \begin{bmatrix} p_{0,n} & p_{1,n} & p_{2,n} \\ \widehat{p}_{0,n-1} & \widehat{p}_{1,n-1} & \widehat{p}_{2,n-1} \\ p_{0,n-1} & p_{1,n-1} & p_{2,n-1} \end{bmatrix}. \tag{8-25}$$

The existence and uniqueness of the solution of the Riemann-Hilbert problem (8-23), (8-24) is proved in a similar way to the proof of Corollary 8.1.

Corollary 8.2. *Given* $n \in \mathbb{N}$*, the absolutely continuous measures* $\mathrm{d}\alpha \subset \mathbb{R}_+$ *and* $\mathrm{d}\beta^* \subset \mathbb{R}_-$*, and assuming the existence of all the bimomoments* I_{ij} *there exists a unique matrix* $\Gamma(z)$ *solving the RHP specified by equations* (8-23)*,* (8-24)*. The solution characterizes uniquely the polynomials* $\widehat{p}_{n-1}$ *and* p_n*.*

9 Acknowledgments

M.B. would like to thank the Department of Mathematics of the University of Notre Dame for hospitality during which the project was initiated and J. Harnad for insight on the relationship of Cauchy biorthogonal polynomials with matrix models.

While working on this project, M. B. and M. G. enjoyed the hospitality of the Department of Mathematics, University of Saskatchewan and M. G. and J. S. enjoyed the hospitality of the Centre de recherches mathématiques, Université de Montréal.

J.S wishes to thank R. Beals for sharing with us his insight regarding the Parseval Identity for an ordinary inhomogeneous string of M.G. Krein which influenced the formulation of Theorem 3.4. J.S would also like to thank H. Lundmark for an ongoing collaboration on the cubic string problem which motivated many of the questions addressed in this paper.

References

[1] R. Camassa and D. D. Holm. An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.*, 71(11):1661–1664, 1993.

[2] R. Beals, D. H. Sattinger, and J. Szmigielski. Multi-peakons and a theorem of Stieltjes. *Inverse Problems*, 15(1):L1–L4, 1999.

[3] R. Beals, D. Sattinger, and J. Szmigielski. Multipeakons and the classical moment problem. *Advances in Mathematics*, 154:229–257, 2000.

[4] T. J. Stieltjes. *Œuvres complètes/Collected papers. Vol. I, II.* Springer-Verlag, Berlin, 1993.

[5] A. Degasperis and M. Procesi. Asymptotic integrability. In A. Degasperis and G. Gaeta, editors, *Symmetry and perturbation theory (Rome, 1998)*, pages 23–37. World Scientific Publishing, River Edge, NJ, 1999.

[6] A. Degasperis, D. D. Holm, and A. N. W. Hone. A new integrable equation with peakon solutions. *Theoretical and Mathematical Physics*, 133:1461–1472, 2002. Preprint nlin.SI/0205023.

[7] A. Degasperis, D. D. Holm, and A. N. W. Hone. Integrable and non-integrable equations with peakons. In M. J. Ablowitz, M. Boiti, F. Pempinelli, and B. Prinari, editors, *Nonlinear Physics: Theory and Experiment (Gallipoli, 2002)*, volume II, pages 37–43. World Scientific, 2003. Preprint nlin.SI/0209008.

[8] H. Lundmark and J. Szmigielski. Multi-peakon solutions of the Degasperis–Procesi equation. *Inverse Problems*, 19:1241–1245, December 2003.

[9] H. Lundmark and J. Szmigielski. Degasperis-Procesi peakons and the discrete cubic string. *IMRP Int. Math. Res. Pap.*, (2):53–116, 2005.

[10] M. L. Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[11] M. Bertola, B. Eynard, and J. Harnad. Duality, biorthogonal polynomials and multi-matrix models. *Comm. Math. Phys.*, 229(1):73–120, 2002.

[12] M. Bertola, B. Eynard, and J. Harnad. The duality of spectral curves that arises in two-matrix models. *Teoret. Mat. Fiz.*, 134(1):32–45, 2003.

[13] M. Bertola, B. Eynard, and J. Harnad. Differential systems for biorthogonal polynomials appearing in 2-matrix models and the associated Riemann-Hilbert problem. *Comm. Math. Phys.*, 243(2):193–240, 2003.

[14] N. M. Ercolani and K. T.-R. McLaughlin. Asymptotics and integrable structures for biorthogonal polynomials associated to a random two-matrix model. *Phys. D*, 152/153:232–268, 2001. Advances in nonlinear mathematics and science.

[15] M. Bertola. Biorthogonal polynomials for two-matrix models with semiclassical potentials. *J. Approx. Theory*, 144(2):162–212, 2007. CRM preprint CRM-3205 (2005).

[16] I. S. Kac and M. G. Krein. On the spectral functions of the string. *Amer. Math. Soc. Transl.*, 103(2):19–102, 1974.

[17] M. Bertola, M. Gekhtman, and J. Szmigielski. The Cauchy two–matrix model. *in preparation*, 2008.

[18] M. Bertola, M. Gekhtman, and J. Szmigielski. Strong asymptotics of Cauchy biorthogonal polynomials. *in preparation*, 2008.

[19] B. Eynard and M. L. Mehta. Matrices coupled in a chain. I. Eigenvalue correlations. *J. Phys. A*, 31(19):4449–4456, 1998.

[20] M. Adler and P. van Moerbeke. String-orthogonal polynomials, string equations and 2-Toda symmetries. *Comm.Pure Appl.Math*, 50(3):241–290, 1997.

[21] S. Karlin. *Total positivity. Vol. I.* Stanford University Press, Stanford, Calif, 1968.

[22] E. M. Nikishin and V. N. Sorokin. *Rational approximations and orthogonality*, volume 92 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991. Translated from the Russian by Ralph P. Boas.

[23] F. P. Gantmacher and M. G. Krein. *Oscillation matrices and kernels and small vibrations of mechanical systems*. AMS Chelsea Publishing, Providence, RI, revised edition, 2002. Translation based on the 1941 Russian original, edited and with a preface by Alex Eremenko.

[24] C. W. Cryer. Some properties of totally positive matrices. *Linear Algebra and Appl.*, 15(1):1–25, 1976.

[25] C. W. Cryer. The *LU*-factorization of totally positive matrices. *Linear Algebra and Appl.*, 7:83–92, 1973.

[26] W. Van Assche, J. S. Geronimo, and A. B. J. Kuijlaars. Riemann-Hilbert problems for multiple orthogonal polynomials. In *Special functions 2000: current perspective and future directions (Tempe, AZ)*, volume 30 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 23–59. Kluwer Acad. Publ., Dordrecht, 2001.

[27] M. Bertola, B. Eynard, and J. Harnad. Semiclassical orthogonal polynomials, matrix models and isomonodromic tau functions. *Comm. Math. Phys.*, 263(2):401–437, 2006.

[28] M. Bertola and B. Eynard. The PDEs of biorthogonal polynomials arising in the two-matrix model. *Math. Phys. Anal. Geom.*, 9(1):23–52, 2006.

[29] M. Bertola, B. Eynard, and J. Harnad. Partition functions for matrix models and isomonodromic tau functions. *J. Phys. A*, 36(12):3067–3083, 2003. Random matrix theory.

[30] M. Bertola and M. Gekhtman. Biorthogonal Laurent polynomials, Toeplitz determinants, minimal Toda orbits and isomonodromic tau functions. *Constructive Approximation*, 26(3):383–430, 2007. Preprint, nlin.SI/0503050.

[31] M. Jimbo, T. Miwa, and K. Ueno. Monodromy preserving deformation of linear ordinary differential equations with rational coefficients. i. general theory. *Physica D*, 2(2):306–352, 1981.