Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions

E. Daems and A.B.J. Kuijlaars

Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200 B, 3001 Leuven, Belgium

evi.daems@wis.kuleuven.be
arno.kuijlaars@wis.kuleuven.be

Abstract

We present a generalization of multiple orthogonal polynomials of type I and type II, which we call multiple orthogonal polynomials of mixed type. Some basic properties are formulated, and a Riemann-Hilbert problem for the multiple orthogonal polynomials of mixed type is given. We derive a Christoffel-Darboux formula for these polynomials using the solution of the Riemann-Hilbert problem. The main motivation for studying these polynomials comes from a model of non-intersecting one-dimensional Brownian motions with a given number of starting points and endpoints. The correlation kernel for the positions of the Brownian paths at any intermediate time coincides with the Christoffel-Darboux kernel for the multiple orthogonal polynomials of mixed type with respect to Gaussian weights.

Keywords: Multiple orthogonal polynomials, Christoffel-Darboux formula, Riemann-Hilbert problem, non-intersecting Brownian motion

1 Introduction

In the early nineties, Fokas, Its and Kitaev [19] introduced a 2×2 matrix valued Riemann-Hilbert problem that characterizes orthogonal polynomials on the real line. This approach can be used to study various aspects of the theory of orthogonal polynomials [14]. Combined with the Deift-Zhou steepest descent method for Riemann-Hilbert problems it has been very succesful in deriving asymptotics for orthogonal polynomials and solving basic questions in the theory of random matrices [13, 15]. The Deift-Zhou steepest descent method was first introduced in [17] and further developed in for example [15] and [16].

In [31], the Riemann-Hilbert problem was extended to the case of multiple orthogonal polynomials of type I and type II. Multiple orthogonal polynomials are polynomials which satisfy orthogonality conditions with respect to a number of measures. The definition will be given in section 2. The Riemann-Hilbert problem is now of size $(p+1) \times (p+1)$, where p is the number of measures.

In this paper, we introduce a generalization of multiple orthogonal polynomials. These new polynomials satisfy orthogonality conditions with respect to two sets of weights $w_{1,1}, \ldots, w_{1,p}$ and $w_{2,1}, \ldots, w_{2,q}$. We call these polynomials multiple orthogonal polynomials of mixed type. The definition will be given in section 2, together with some basic properties concerning existence and uniqueness. We can again characterize these polynomials by a Riemann-Hilbert problem which is now of size $(p+q) \times (p+q)$. This will be given in section 3.

The usual monic orthogonal polynomials P_n on the real line with weight function w satisfy a three term recurrence relation and this gives rise to the basic Christoffel-Darboux formula (see for example [11])

$$\sum_{j=0}^{n-1} \frac{1}{h_j} P_j(x) P_j(y) = \frac{1}{h_{n-1}} \frac{P_n(x) P_{n-1}(y) - P_{n-1}(x) P_n(y)}{x-y}, \tag{1.1}$$

where

$$h_j = \int P_j(x) x^j w(x) dx. \tag{1.2}$$

This formula was generalized to multiple orthogonal polynomials in [4] in the case of two weights and in [12] in the general case of p weights. We will derive using the Riemann-Hilbert problem a Christoffel-Darboux formula for the multiple orthogonal polynomials of mixed type.

Our main motivation for studying this new kind of orthogonality comes from non-intersecting Brownian paths. Consider n independent one-dimensional Brownian motions that start in n different fixed points at time $t = 0$ and end in n different fixed points at time $t = 1$, conditioned on the fact that they do not intersect in the full time interval $(0, 1)$. At any intermediate time $t \in (0, 1)$ the positions of the Brownian paths are distributed according to a determinantal point process on the real line. This is a consequence of a classical theorem of Karlin and McGregor [22] and it applies not only to Brownian motion, but to any one-dimensional strong Markov process with continuous sample paths.

Of special interest is the confluent case in which many of the starting points and many of the endpoints coincide. The formulae simplify for Brownian motion because of the properties of the Gaussian transition probabilities. In the extreme case of one starting point and one endpoint, the positions of the Brownian paths have the same distribution, up to simple scaling, as the eigenvalues of a matrix from the Gaussian unitary ensemble (GUE) which is a basic ensemble from random matrix theory [23]. In this case the kernel for the determinantal point process is constructed out of Hermite polynomials and the Christoffel-Darboux formula (1.1) expresses this kernel in terms of Hermite polynomials of degrees n and $n - 1$ only.

In the case of one starting point and q endpoints the positions of the Brownian paths have the same distribution as the eigenvalues of a Gaussian unitary matrix with external source [3, 1, 29]. Then the kernel is constructed out of multiple Hermite polynomials of type I and II. The Christoffel-Darboux formula for multiple orthogonal polynomials of [12] expresses this kernel in terms of a sum of $q + 1$ terms, in which each term involves products of multiple Hermite polynomials of type I and type II.

The next step is to consider Brownian motions which begin in p starting points and end in q endpoints. This gives rise to the multiple orthogonal polynomials of mixed type with respect to Gaussian weights, and so we call these polynomials multiple Hermite polynomials of mixed type. We will discuss this in more detail in section 6. Unfortunately we do not know if there exists a corresponding random matrix model.

For a summary of the above discussion, see Table 1.

Non-intersecting Brownian motions	Associated polynomials	Random matrix ensemble
1 starting point and 1 endpoint	Hermite polynomials	Gaussian unitary ensemble
1 starting point and $q \geq 2$ endpoints	multiple Hermite polynomials	Gaussian unitary ensemble with external source
$p \geq 2$ starting points and $q \geq 2$ endpoints	multiple Hermite polynomials of mixed type	unknown

Table 1: Overview of the connection between non-intersecting Brownian motions, associated polynomials, and random matrix ensembles. The three cases are illustrated in Figures 1, 2, and 3. See section 6 for a more detailed discussion.

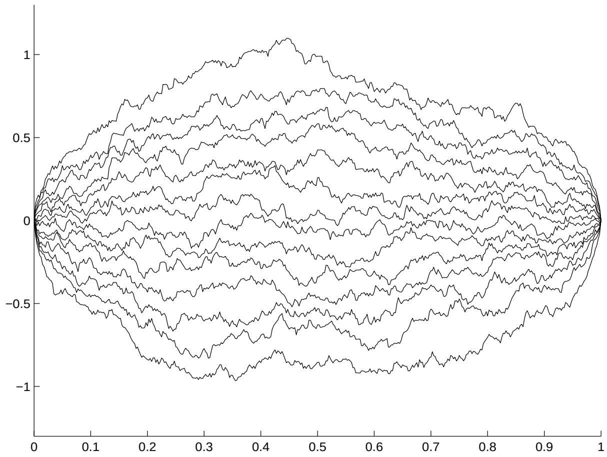

Figure 1: Non-intersecting Brownian motions which start and end at 1 point. At any intermediate time the positions of the paths have the same distribution as the eigenvalues of a GUE matrix. The correlation kernel is built out of Hermite polynomials.

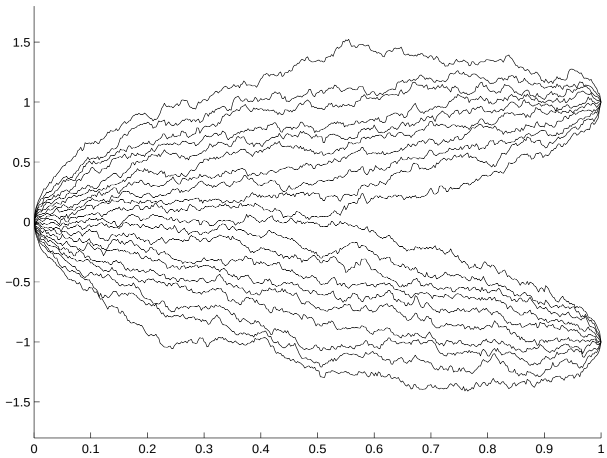

Figure 2: Non-intersecting Brownian motions which start at 1 point and end at 2 different points. At any intermediate time the positions of the paths have the same distribution as the eigenvalues of a Gaussian unitary random matrix with external source. The correlation kernel is built out of multiple Hermite polynomials.

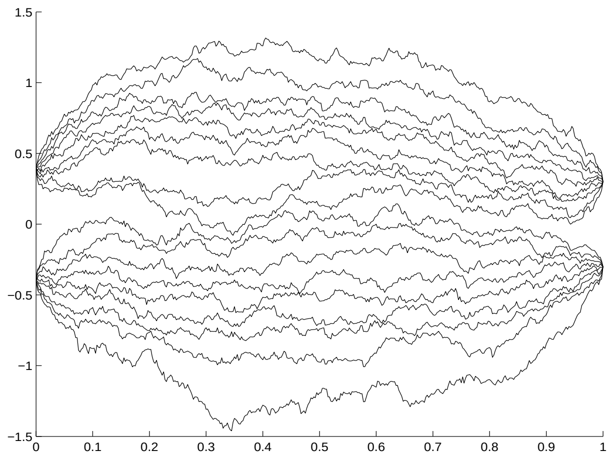

Figure 3: Non-intersecting Brownian motions which start at 2 points and end at 2 points. At any intermediate time the positions of the paths are distributed according to a determinantal point process with a kernel that is built out of multiple Hermite polynomials of mixed type.

2 Multiple orthogonal polynomials of mixed type

In this section we define the multiple orthogonal polynomials of mixed type and we give conditions for existence. These polynomials can be seen as a generalization of multiple orthogonal polynomials of type I and type II which we discuss first. Throughout this paper we will say that w is a weight on $\mathbb{R}$ if $w(x) \geq 0$ for $x \in \mathbb{R}$ and $\int x^k w(x)dx < +\infty$ for every $k \in \mathbb{N} \cup \{0\}$.

2.1 Multiple orthogonal polynomials of type I

Let $w_1, w_2, \ldots, w_p$ be p weights on the real line and let $\vec{n} = (n_1, \ldots, n_p)$ be a multi-index consisting of non-negative integers. If $A_1, \ldots, A_p$ are polynomials and

$$Q(x) = \sum_{j=1}^{p} A_j(x) w_j(x), \qquad \deg A_j \leq n_j - 1, \tag{2.1}$$

such that

$$\int Q(x) x^j dx = 0 \qquad \text{for } j = 0, \ldots, |\vec{n}| - 2, \tag{2.2}$$

then the A_j are called multiple orthogonal polynomials of type I and Q is the linear form built out of the multiple orthogonal polynomials of type I. Here we follow the usual multi-index notation

$$|\vec{n}| = \sum_{i=1}^{p} n_i.$$

The relations (2.2) give us $|\vec{n}| - 1$ homogeneous linear equations for the in total $|\vec{n}|$ coefficients of the polynomials A_j. So there is always a non-zero solution. If the solution is unique up to a multiplicative factor, then the multi-index $\vec{n}$ is called normal for type I. The multi-index $\vec{n}$ is called strongly normal for type I if we have

$$\int Q(x)x^{|\vec{n}|-1}dx \neq 0$$

for any non-zero Q satisfying (2.1)–(2.2). In that case we can normalize the multiple orthogonal polynomials of type I so that

$$\int Q(x)x^{|\vec{n}|-1}dx = 1. \tag{2.3}$$

We call (2.3) a type I normalization.

2.2 Multiple orthogonal polynomials of type II

Let $w_1, w_2, \ldots, w_q$ be q weights on the real line and let $\vec{m} = (m_1, \ldots, m_q)$ be a multi-index of length q. If P is a polynomial of degree $|\vec{m}|$ such that

$$\int P(x)x^j w_k(x)dx = 0 \qquad \text{for } j = 0, \ldots, m_k - 1 \text{ and } k = 1, \ldots, q, \tag{2.4}$$

then P is called a multiple orthogonal polynomial of type II. The $|\vec{m}|$ equations (2.4) are homogeneous linear equations for the $|\vec{m}| + 1$ coefficients of P. So there is always a non-zero solution. If the solution is unique up to a multiplicative factor then the multi-index $|\vec{m}|$ is called normal for type II. If every non-zero solution has a non-zero leading coefficient then the multi-index $|\vec{m}|$ is called strongly normal for type II. In that case we can normalize the multiple orthogonal polynomial of type II so that

$$P_n(x) = x^n + \cdots \tag{2.5}$$

and we call (2.5) a type II normalization.

For more details and examples of multiple orthogonal polynomials of type I and type II, and about their relation with Hermite-Padé approximation, we refer the interested reader to [2, 24, 30] and the references cited therein.

2.3 Multiple orthogonal polynomials of mixed type

To define the multiple orthogonal polynomials of mixed type, we need two sets of weights on $\mathbb{R}$: $w_{1,1}, w_{1,2}, \ldots, w_{1,p}$ and $w_{2,1}, w_{2,2}, \ldots, w_{2,q}$, which we collect in two row vectors

$$\vec{w}_1 = (w_{1,1}, \ldots, w_{1,p}), \qquad \vec{w}_2 = (w_{2,1}, \ldots, w_{2,q}),$$

and two multi-indices $\vec{n} = (n_1, \ldots, n_p)$ and $\vec{m} = (m_1, \ldots, m_q)$ of length p and q, respectively, such that

$$|\vec{n}| = |\vec{m}| + 1. \tag{2.6}$$

Definition 2.1 We call the polynomials $A_1, \ldots, A_p$ with

$$\deg A_j \leq n_j - 1 \qquad \text{for } j = 1, \ldots, p \tag{2.7}$$

multiple orthogonal polynomials of mixed type for the pair of multi-indices $(\vec{n}, \vec{m})$ and with respect to the vectors of weights $\vec{w}_1$ and $\vec{w}_2$ if the function

$$Q(x) = \sum_{j=1}^{p} A_j(x) w_{1,j}(x) \tag{2.8}$$

satisfies the following orthogonality conditions:

$$\int Q(x) x^j w_{2,k}(x) dx = 0 \qquad \text{for } j = 0, 1, \ldots, m_k - 1 \text{ and } k = 1, \ldots, q. \tag{2.9}$$

To emphasize the dependence on the multi-indices we also write

$$A_j = A_{j,\vec{n},\vec{m}}, \qquad Q = Q_{\vec{n},\vec{m}},$$

and to emphasize the role of the two vectors of weights we will occasionally write

$$A_j(x) = A_{j,\vec{n},\vec{m}}(x; \vec{w}_1, \vec{w}_2), \qquad Q(x) = Q_{\vec{n},\vec{m}}(x; \vec{w}_1, \vec{w}_2),$$

although mostly we drop the explicit mentioning of the weights. Note that the role of the two vectors of weights is not symmetric. The function Q is a linear form with respect to the weights from $\vec{w}_1$ as in multiple orthogonality of type I, and the linear form has a number of orthogonality conditions with respect to the weights from $\vec{w}_2$ as in multiple orthogonality of type II.

The conditions (2.9) lead to $|\vec{m}|$ homogeneous linear equations for the in total $|\vec{n}|$ free coefficients of the polynomials A_j. Because of the assumption (2.6) there is always a non-zero solution. If the polynomials A_j are unique up to a multiplicative constant, then we call $(\vec{n}, \vec{m})$ a **normal** pair of indices for the two sets of weights $\vec{w}_1$ and $\vec{w}_2$.

For a normal pair of indices we can choose a certain normalization in order to define a unique multiple orthogonal polynomial of mixed type. In this paper, we are going to use two types of normalization:

- **Type I normalization:** Fix $k = 1, \dots, q$ and normalize Q such that

$$\int Q(x) x^{m_k} w_{2,k}(x) dx = 1. \tag{2.10}$$

If we choose this normalization, we write

$$A_j = A_j^{(I,k)}, \qquad Q = Q^{(I,k)}, \tag{2.11}$$

or in full notation, if we want to emphasize the dependence on the multi-indices and the weights

$$A_j(x) = A_{j,\vec{n},\vec{m}}^{(I,k)}(x; \vec{w}_1, \vec{w}_2), \qquad Q(x) = Q_{\vec{n},\vec{m}}^{(I,k)}(x; \vec{w}_1, \vec{w}_2).$$

- **Type II normalization:** Fix $k = 1, \dots, p$ and normalize Q such that A_k is a monic polynomial of degree $n_k - 1$. If we choose this normalization, we write

$$A_j = A_j^{(II,k)}, \qquad Q = Q^{(II,k)}, \tag{2.12}$$

or in full

$$A_j(x) = A_{j,\vec{n},\vec{m}}^{(II,k)}(x; \vec{w}_1, \vec{w}_2), \qquad Q(x) = Q_{\vec{n},\vec{m}}^{(II,k)}(x; \vec{w}_1, \vec{w}_2).$$

We emphasize that the above normalizations may not always be possible.

2.4 Conditions for normality

We will state the conditions using the Hilbert space geometry of $L^2(\mathbb{R})$ and to do so we assume that

$$x^j w_{1,k} \in L^2(\mathbb{R}), \qquad \text{for } k = 1, \dots, p, \quad j = 0, 1, \dots, n_k - 1, \tag{2.13}$$

$$x^j w_{2,k} \in L^2(\mathbb{R}), \qquad \text{for } k = 1, \ldots, q, \quad j = 0, 1, \ldots, m_k - 1. \tag{2.14}$$

Associated with $\vec{n}$ and $\vec{w}_1$ we have the vector space

$$F_{\vec{n}} = \{\sum_{j=1}^{p} A_j w_{1,j} \mid A_j \text{ is a polynomial of degree } \leq n_j - 1 \text{ for } j = 1, \ldots, p\} \tag{2.15}$$

and associated with $\vec{m}$ and $\vec{w}_2$ we have

$$G_{\vec{m}} = \{\sum_{j=1}^{q} B_j w_{2,j} \mid B_j \text{ is a polynomial of degree } \leq m_j - 1 \text{ for } j = 1, \ldots, q\}. \tag{2.16}$$

Hence $F_{\vec{n}}$ is the linear span of the functions in (2.13) and $G_{\vec{m}}$ is the linear span of the functions in (2.14).

Let $\vec{e}_k$ be the standard basis vector

$$\vec{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0), \qquad \text{where 1 is in the } k\text{th position.} \tag{2.17}$$

We de not specify the length of the vector $\vec{e}_k$, but this should be clear from the context.

Lemma 2.2 *Suppose that the vector space $F_{\vec{n}}$ is $|\vec{n}|$-dimensional and let $|\vec{n}| = |\vec{m}| + 1$.*

(a) *Then Q is a linear form (2.8) of multiple orthogonal polynomials of mixed type for the pair $(\vec{n}, \vec{m})$ if and only if $Q \in F_{\vec{n}} \cap G_{\vec{m}}^{\perp}$.*

(b) *The pair $(\vec{n}, \vec{m})$ is a normal pair of indices if and only if $F_{\vec{n}} \cap G_{\vec{m}}^{\perp}$ is one-dimensional.*

(c) *The pair $(\vec{n}, \vec{m})$ allows a type I normalization with respect to the kth index of $\vec{m}$ if and only if $F_{\vec{n}} \cap G_{\vec{m}+\vec{e}_k}^{\perp} = \{0\}$.*

(d) *The pair $(\vec{n}, \vec{m})$ allows a type II normalization with respect to the kth index of $\vec{n}$ if and only if $F_{\vec{n}-\vec{e}_k} \cap G_{\vec{m}}^{\perp} = \{0\}$.*

Proof. (a) This is immediate from the definitions.

(b) If $F_{\vec{n}} \cap G_{\vec{m}}^{\perp}$ is one-dimensional, then by part (a) the linear form Q is unique up to a multiplicative constant. Then the polynomials A_j are also unique up to a multiplicative constant, since $F_{\vec{n}}$ is $|\vec{n}|$-dimensional. Thus

$(\vec{n}, \vec{m})$ is a normal pair of indices. The converse is obvious in view of part (a).

(c) Suppose $F_{\vec{n}} \cap G^{\perp}_{\vec{m}+\vec{e}_k} = \{0\}$. Let Q be a non-zero linear form of multiple orthogonal polynomials of mixed type. Then $Q \in G^{\perp}_{\vec{m}}$ by part (a), but $Q \notin G^{\perp}_{\vec{m}+\vec{e}_k}$. Since $G_{\vec{m}} \subset G_{\vec{m}+\vec{e}_k}$ with codimension one, it follows that Q is not orthogonal to any function in $G_{\vec{m}+\vec{e}_k} \setminus G_{\vec{m}}$ and since $x^{m_k} w_{2,k}$ belongs to this set, this implies

$$\int Q(x) x^{m_k} w_{2,k}(x) dx \neq 0.$$

Then we can normalize Q so that this integral is 1 and the pair $(\vec{n}, \vec{m})$ allows a type I normalization with respect to the kth index.

On the other hand if $F_{\vec{n}} \cap G^{\perp}_{\vec{m}+\vec{e}_k} \neq \{0\}$, then any non-zero Q in this space would be a non-zero linear form of multiple orthogonal polynomials of mixed type such that

$$\int Q(x) x^{m_k} w_{2,k}(x) dx = 0.$$

Then either $(\vec{n}, \vec{m})$ is not normal, or if it is normal, it does not allow a type I normalization with respect to the kth index.

(d) Suppose $F_{\vec{n}-\vec{e}_k} \cap G^{\perp}_{\vec{m}} = \{0\}$. Let Q be a non-zero linear form of multiple orthogonal polynomials of mixed type for the pair $(\vec{n}, \vec{m})$. Then $Q \in F_{\vec{n}}$ by part (a), but $Q \notin F_{\vec{n}-\vec{e}_k}$. This implies that A_k has exact degree $n_k - 1$. Then we can normalize Q so that the leading coefficient of A_k is 1 and thus the pair $(\vec{n}, \vec{m})$ allows a type II normalization with respect to the kth index.

On the other hand if $F_{\vec{n}-\vec{e}_k} \cap G^{\perp}_{\vec{m}} \neq \{0\}$, then any non-zero Q in this space would be a non-zero linear form of multiple orthogonal polynomials of mixed type for the pair $(\vec{n}, \vec{m})$ with the degree of A_k less than $n_k - 1$. Then either $(\vec{n}, \vec{m})$ is not normal, or if it is normal, it does not allow a type II normalization with respect to the kth index. □

We have the following easy corollary of Lemma 2.2.

Corollary 2.3 *Suppose that*

(1) $|\vec{n}| = |\vec{m}|$ *and* $F_{\vec{n}}$ *is a* $|\vec{n}|$*-dimensional subspace of* $L^2(\mathbb{R})$.

(2) $F_{\vec{n}} \cap G^{\perp}_{\vec{m}} = \{0\}$.

Then the following holds:

(a) *For every* $k = 1, \ldots, q$ *we have that* $(\vec{n}, \vec{m} - \vec{e}_k)$ *is a normal pair of multi-indices which allows a type I normalization with respect to the* k*th index. Hence* $Q^{(I,k)}_{\vec{n},\vec{m}-\vec{e}_k}$ *exists and is unique.*

(b) *For every* $k = 1, \ldots, p$ *we have that* $(\vec{n} + \vec{e}_k, \vec{m})$ *is a normal pair of multi-indices which allows a type II normalization with respect to the* k*th index. Hence* $Q^{(II,k)}_{\vec{n}+\vec{e}_k,\vec{m}}$ *exists and is unique.*

Proof. Part (a) follows immediately from part (c) of Lemma 2.2 applied to the multi-indices $\vec{n}$ and $\vec{m} - \vec{e}_k$, and part (b) follows from part (d) of Lemma 2.2 applied to the multi-indices $\vec{n} + \vec{e}_k$ and $\vec{m}$. □

3 The Riemann-Hilbert problem

Fokas, Its, and Kitaev [19] found a Riemann-Hilbert problem that characterizes the orthogonal polynomials. Van Assche, Geronimo, and Kuijlaars [31] extended this Riemann-Hilbert problem to multiple orthogonal polynomials of type I and type II. We are now going to give a further extension to multiple orthogonal polynomials of mixed type.

3.1 Riemann-Hilbert problem for multiple orthogonal polynomials of mixed type

Assume that $|\vec{n}| = |\vec{m}|$ and let $\vec{w}_1$ and $\vec{w}_2$ be as before. In addition to (2.13) and (2.14) we assume that the weights $w_{1,k}$ and $w_{2,k}$ are continuous and a.e. differentiable with

$$x^j w'_{i,k}(x) \in L^2(\mathbb{R}), \qquad \text{for } i = 1, 2, \quad k = 1, \ldots, p, \quad j = 0, 1, \ldots, n_k - 1.$$

This will ensure that the boundary values $Y_\pm(x)$ in (3.1) can be taken pointwise and uniformly for $x \in \mathbb{R}$, see also the discussion in [13, Section 3.2].

For convenience we also assume that $n_k > 0$ and $m_l > 0$ for $k = 1, \ldots, p$ and $l = 1, \ldots, q$. We can modify the arguments in case one or several of the n_k and m_l are zero, but we will not discuss that here.

Consider the following Riemann-Hilbert problem: determine a $(p+q) \times (p+q)$ matrix valued function $Y : \mathbb{C} \setminus \mathbb{R} \to \mathbb{C}^{(p+q)\times(p+q)}$ such that

(1) Y is analytic on $\mathbb{C} \setminus \mathbb{R}$,

(2) for $x \in \mathbb{R}$, we have

$$Y_+(x) = Y_-(x) \begin{bmatrix} I_p & W(x) \\ 0 & I_q \end{bmatrix} \tag{3.1}$$

where I_p and I_q denote the identity matrices of sizes p and q, respectively, and

$$W = \vec{w}_1^t \vec{w}_2 = \begin{bmatrix} w_{1,1}w_{2,1} & w_{1,1}w_{2,2} & \cdots & w_{1,1}w_{2,q} \\ w_{1,2}w_{2,1} & w_{1,2}w_{2,2} & \cdots & w_{1,2}w_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1,p}w_{2,1} & w_{1,p}w_{2,2} & \cdots & w_{1,p}w_{2,q} \end{bmatrix}, \tag{3.2}$$

(3) as $z \to \infty$, we have that

$$Y(z) = \left(I_{p+q} + O\left(\frac{1}{z}\right)\right) \begin{bmatrix} z^{n_1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & z^{n_2} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & z^{n_p} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & z^{-m_1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & z^{-m_q} \end{bmatrix} \tag{3.3}$$

where I_{p+q} denotes the identity matrix of size $p+q$.

As before we use $F_{\vec{n}}$ and $G_{\vec{m}}$ to denote the spaces (2.15) and (2.16) associated with the indices $\vec{n}$ and $\vec{m}$. The main result of this section is that the Riemann-Hilbert problem has a unique solution if the conditions of Corollary 2.3 are satisfied.

Theorem 3.1 *Let $|\vec{n}| = |\vec{m}|$. Suppose that $F_{\vec{n}}$ is a $|\vec{n}|$-dimensional subspace of $L^2(\mathbb{R})$ such that $F_{\vec{n}} \cap G_{\vec{m}}^{\perp} = \{0\}$. Then the above Riemann-Hilbert problem has a unique solution, given in terms of the multiple orthogonal polynomials of mixed type with respect to the vectors of weights $\vec{w}_1$ and $\vec{w}_2$. We have*

- *for $k, l = 1, \ldots, p$:*

$$Y_{k,l}(z) = A^{(II,k)}_{l,\vec{n}+\vec{e}_k,\vec{m}}(z), \tag{3.4}$$

- *for $k = 1, \ldots, q$ and $l = 1, \ldots, p$:*

$$Y_{p+k,l}(z) = -2\pi i A^{(I,k)}_{l,\vec{n},\vec{m}-\vec{e}_k}(z), \tag{3.5}$$

- *for $k = 1, \ldots, p$ and $l = 1, \ldots, q$:*

$$Y_{k,p+l}(z) = \frac{1}{2\pi i} \int \frac{Q^{(II,k)}_{\vec{n}+\vec{e}_k,\vec{m}}(x) w_{2,l}(x)}{x - z} dx, \tag{3.6}$$

- *for* $k, l = 1, \ldots, q$*:*

$$Y_{p+k,p+l}(z) = -\int \frac{Q^{(I,k)}_{\vec{n},\vec{m}-\vec{e}_k}(x) w_{2,l}(x)}{x-z} dx. \tag{3.7}$$

Proof. First note that the functions $Q^{(II,k)}_{\vec{n}+\vec{e}_k,\vec{m}}$ and $Q^{(I,l)}_{\vec{n},\vec{m}-\vec{e}_l}$ uniquely exist for $k = 1, \ldots, p$ and $l = 1, \ldots, q$ by Corollary 2.3. Partition Y as

$$Y(z) = \begin{bmatrix} K(z) & L(z) \\ M(z) & N(z) \end{bmatrix}, \tag{3.8}$$

where K is a $p \times p$ matrix, L is a $p \times q$ matrix, M is a $q \times p$ matrix, and N is a $q \times q$ matrix.

The jump condition (3.1) implies that

$$K^+_{k,l}(x) = K^-_{k,l}(x) \quad \text{for } x \in \mathbb{R}, \tag{3.9}$$

so that K is analytic on the full complex plane. From (3.9), the asymptotic condition (3.3), and an extension of Liouville's theorem to polynomials, it then follows that each diagonal element $K_{k,k}$, with $k = 1, \ldots, p$, is a monic polynomial of degree n_k, and that each off-diagonal element $K_{k,l}$ is a polynomial of degree at most $n_l - 1$.

For the matrix L the jump condition (3.1) implies that for $x \in \mathbb{R}$,

$$L^+_{k,l}(x) = L^-_{k,l}(x) + Q_k(x) w_{2,l}(x) \quad \text{for } k = 1, \ldots, p \text{ and } l = 1, \ldots, q, \tag{3.10}$$

where

$$Q_k(x) = \sum_{j=1}^{p} K_{k,j}(x) w_{1,j}.$$

The Sokhotsky-Plemelj formula then gives that

$$L_{k,l}(z) = \frac{1}{2\pi i} \int \frac{Q_k(x) w_{2,l}(x)}{x-z} dx \quad \text{if } z \in \mathbb{C} \setminus \mathbb{R}. \tag{3.11}$$

If we now use the expansion

$$\frac{1}{z-x} = \sum_{l=0}^{n-1} \frac{x^l}{z^{l+1}} + \frac{x^n}{z^n} \frac{1}{z-x} \quad \text{for } n \in \mathbb{N}, \tag{3.12}$$

we find that

$$L_{k,l}(z) = -\sum_{j=0}^{n-1} \frac{1}{2\pi i z^{j+1}} \int Q_k(x) x^j w_{2,l}(x) dx - \frac{1}{2\pi i z^n} \int \frac{Q_k(x) x^n w_{2,l}(x)}{z-x} dx \quad \text{for } k=1,\dots,p \text{ and } l=1,\dots,q. \tag{3.13}$$

The asymptotic condition (3.3) gives that

$$\lim_{z\to\infty} L_{k,l}(z) z^{m_l} = 0, \tag{3.14}$$

such that from (3.13) with $n = m_l - 1$ we get

$$\int Q_k(x) x^j w_{2,l}(x) dx = 0 \quad \text{for } j=0,\dots,m_l-1, k=1,\dots,p \text{ and } l=1,\dots,q. \tag{3.15}$$

Because $K_{k,l}$ is a polynomial of degree at most $n_l - 1$ if $l \neq k$ and $K_{k,k}$ is a monic polynomial of degree n_k, as mentioned in the beginning of the proof, and because of (3.15) we see that

$$K_{k,l} = A^{(II,k)}_{l,\vec{n}+\vec{e}_k,\vec{m}} \tag{3.16}$$

for $k, l = 1,\dots,p$. Because of (3.11) and (3.16) we see that

$$L_{k,l}(z) = \frac{1}{2\pi i} \int \frac{Q^{(II,k)}_{\vec{n}+\vec{e}_k,\vec{m}}(x) w_{2,l}(x)}{x-z} dx \quad \text{if } z \in \mathbb{C} \setminus \mathbb{R} \tag{3.17}$$

for $k=1,\dots,p$ and $l=1,\dots,q$. This proves the formulae (3.4) and (3.6).

The jump condition (3.1) shows that for $k=1,\dots,q$ and $l=1,\dots,p$,

$$M^+_{k,l}(x) = M^-_{k,l}(x) \quad \text{for } x \in \mathbb{R}, \tag{3.18}$$

and consequently M is also analytic on the full complex plane. In the same way as for the matrix K, the asymptotic condition (3.3) implies that each $M_{k,l}$ is a polynomial of degree $\leq n_l - 1$. The jump condition (3.1) also implies that

$$N^+_{k,l}(x) = N^-_{k,l}(x) + Q_k(x) w_{2,l}(x) \quad \text{for } k,l=1,\dots,q \text{ and } x \in \mathbb{R}, \tag{3.19}$$

with

$$Q_k(x) = \sum_{j=1}^{p} M_{k,j}(x) w_{1,j}(x).$$

Using the Sokhotsky-Plemelj formula we get that

$$N_{k,l}(z) = \frac{1}{2\pi i} \int \frac{Q_k(x) w_{2,l}(x)}{x-z} dx \qquad \text{for } k,l = 1,\ldots,q \text{ if } z \in \mathbb{C} \setminus \mathbb{R}. \tag{3.20}$$

The asymptotic condition (3.3) gives that

$$\begin{cases} \lim_{z\to\infty} N_{k,l}(z) z^{m_l} = 0 & \text{if } k \neq l, \\ \lim_{z\to\infty} N_{k,k}(z) z^{m_k} = 1, \end{cases}$$

Using these conditions and the expansion (3.12) we get that

$$\begin{cases} \int Q_k(x) x^j w_{2,l}(x) dx = 0 & \text{for } j = 0,1,\ldots,m_l - 1 \text{ if } k \neq l, \\ \int Q_k(x) x^j w_{2,k}(x) dx = 0 & \text{for } j = 0,1,\ldots,m_k - 2, \\ \int Q_k(x) x^{m_k - 1} w_{2,k}(x) dx = -2\pi i. \end{cases} \tag{3.21}$$

The degree of the polynomials $M_{k,l}$ and the orthogonality conditions (3.21) imply that $M_{k,l} = -2\pi i A^{(I,k)}_{l,\vec{n},\vec{m}-\vec{e}_k}$ for $k = 1,\ldots,q$ and $l = 1,\ldots,p$. Because of (3.20) we get that

$$N_{k,l}(z) = -\int \frac{Q^{(I,k)}_{\vec{n},\vec{m}-\vec{e}_k}(x) w_{2,l}(x)}{x-z} dx \tag{3.22}$$

for $k,l = 1,\ldots,q$ and $z \in \mathbb{C}\setminus\mathbb{R}$. This proves the formulae (3.5) and (3.7). This completes the proof of Theorem 3.1. □

Remark 3.2 Here we point out a possible connection with the theory of orthogonal matrix polynomials, see [18] and references cited therein. The $p \times p$ left upper block of Y, which we called K in the proof of Theorem 3.1, is a $p \times p$ matrix polynomial. It satisfies certain orthogonality conditions which can be interpreted as matrix orthogonality in the special case that $p = q$ and $n_j = m_k = d$ for all $j,k = 1,\ldots,p$, that is, all indices are equal. Indeed in that case we have that K is monic of degree d,

$$K(x) = x^d I_p + O(x^{d-1}),$$

and satisfies

$$\int K(x)W(x)L(x)dx = 0$$

for every matrix polynomial L of degree $\leq d-1$, where W is the matrix of weights (3.2).

Note however that W is not necessarily symmetric and has rank one (so is non-invertible), which is in contrast to what is usually assumed for orthogonal matrix polynomials [18].

3.2 Riemann-Hilbert problem for the inverse

By standard arguments it follows that $\det Y(z) \equiv 1$ for $z \in \mathbb{C} \setminus \mathbb{R}$ so that the inverse $Y^{-1}(z)$ exists and is analytic for $z \in \mathbb{C} \setminus \mathbb{R}$. Define

$$X(z) = Y^{-t}(z). \tag{3.23}$$

From the Riemann-Hilbert problem for Y it is then straightforward to check that X is the solution of the following $(p+q)\times(p+q)$ matrix valued Riemann-Hilbert problem:

(1) $X : \mathbb{C} \setminus \mathbb{R} \to \mathbb{C}^{(p+q)\times(p+q)}$ is analytic,

(2) for $x \in \mathbb{R}$, we have

$$X_+(x) = X_-(x) \begin{bmatrix} I_p & 0 \\ -W^t(x) & I_q \end{bmatrix} \tag{3.24}$$

where W is given by (3.2),

(3) as $z \to \infty$, we have that

$$X(z) = \left(I_{p+q} + O\left(\frac{1}{z}\right)\right) \begin{bmatrix} z^{-n_1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & z^{-n_2} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & z^{-n_p} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & z^{m_1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & z^{m_q} \end{bmatrix}. \tag{3.25}$$

The solution of the Riemann-Hilbert problem for X can again be written in terms of multiple orthogonal polynomials of mixed type, but with the roles of the vectors of weights $\vec{w}_1$ and $\vec{w}_2$ as well as the multi-indices $\vec{n}$ and $\vec{m}$ interchanged. Therefore we use the full notation $Q_{\vec{m},\vec{n}}(x;\vec{w}_2,\vec{w}_1)$.

Lemma 3.3 *Suppose that the conditions of Theorem* 3.1 *are satisfied. Then the above Riemann-Hilbert problem has a unique solution given by*

- *For* $k,l=1,\dots,p$*:*

$$X_{k,l}(z) = -\int \frac{Q^{(I,k)}_{\vec{m},\vec{n}-\vec{e}_k}(x;\vec{w}_2,\vec{w}_1)w_{1,l}(x)}{x-z}dx, \tag{3.26}$$

- *for* $k=1,\dots,q$ *and* $l=1,\dots,p$*:*

$$X_{p+k,l}(z) = -\frac{1}{2\pi i}\int \frac{Q^{(II,k)}_{\vec{m}+\vec{e}_k,\vec{n}}(x;\vec{w}_2,\vec{w}_1)w_{1,l}(x)}{x-z}dx, \tag{3.27}$$

- *for* $k=1,\dots,p$ *and* $l=1,\dots,q$*:*

$$X_{k,p+l}(z) = 2\pi i A^{(I,k)}_{l,\vec{m},\vec{n}-\vec{e}_k}(z;\vec{w}_2,\vec{w}_1), \tag{3.28}$$

- *for* $k,l=1,\dots,q$*:*

$$X_{p+k,p+l}(z) = A^{(II,k)}_{l,\vec{m}+\vec{e}_k,\vec{n}}(z;\vec{w}_2,\vec{w}_1). \tag{3.29}$$

Proof. The lemma can be proven in the same way as Theorem 3.1, but it is also possible to derive it directly from Theorem 3.1 as follows. Let U be the solution of the Riemann-Hilbert problem described in Theorem 3.1, but with the roles of the vectors of weights $\vec{w}_1$ and $\vec{w}_2$ as well as the multi-indices $\vec{n}$ and $\vec{m}$ interchanged. By comparing the jump conditions and the asymptotic conditions of the Riemann-Hilbert problem of X and U, we can easily see that

$$U(z) = \begin{bmatrix} 0 & -I_q \\ I_p & 0 \end{bmatrix} X(z) \begin{bmatrix} 0 & I_p \\ -I_q & 0 \end{bmatrix}. \tag{3.30}$$

Theorem 3.1 therefore implies that the solution of the Riemann-Hilbert problem defined above is unique and is given by the formulae (3.26)–(3.29). □

4 The kernel

Suppose that $\vec{w}_1 = (w_{1,1}, w_{1,2}, \ldots, w_{1,p})$ and $\vec{w}_2 = (w_{2,1}, w_{2,2}, \ldots, w_{2,q})$ are two vectors of weights on the real line, and define $F_{\vec{n}}$ and $G_{\vec{m}}$ as in (2.15) and (2.16). Suppose as in section 3 that $|\vec{n}| = |\vec{m}| = n$ and that $F_{\vec{n}}$ and $G_{\vec{m}}$ are both n-dimensional subspaces of $L^2(\mathbb{R})$. Two bases $\phi_1, \ldots, \phi_n$ of $F_{\vec{n}}$ and $\psi_1, \ldots, \psi_n$ of $G_{\vec{m}}$ are called biorthogonal if

$$\int \phi_j(x)\psi_k(x)dx = \delta_{j,k}.$$

The following lemmas are well-known, but we include their proofs for completeness.

Lemma 4.1 *There exist biorthogonal bases for $F_{\vec{n}}$ and $G_{\vec{m}}$ if and only if $F_{\vec{n}} \cap G_{\vec{m}}^{\perp} = \{0\}$.*

Proof. First suppose that $\phi_1, \ldots, \phi_n \in F_{\vec{n}}$ and $\psi_1, \ldots, \psi_n \in G_{\vec{m}}$ are biorthogonal bases. Then every $f \in F_{\vec{n}}$ can be written as

$$f = \sum_{j=1}^{n} c_j \phi_j \tag{4.1}$$

with

$$c_j = \int f(x)\psi_j(x)dx, \qquad j = 1, \ldots, n. \tag{4.2}$$

Since every $c_j = 0$ if $f \in G_{\vec{m}}^{\perp}$ it follows that $F_{\vec{n}} \cap G_{\vec{m}}^{\perp} = \{0\}$.

Conversely, suppose that $F_{\vec{n}} \cap G_{\vec{m}}^{\perp} = \{0\}$. Let $\psi_1, \ldots, \psi_n$ be any basis of $G_{\vec{m}}$. Consider the linear mapping

$$\mathcal{F} : F_{\vec{n}} \to \mathbb{R}^n : f \mapsto (\int f(x)\psi_k(x)dx)_{k=1,\ldots,n}. \tag{4.3}$$

Because $F_{\vec{n}} \cap G_{\vec{m}}^{\perp} = \{0\}$ it is clear that $\mathcal{F}$ is injective. Since the dimensions of $\mathbb{R}^n$ and $F_{\vec{n}}$ are equal, and the mapping is linear, $\mathcal{F}$ is bijective. Consequently there exist functions $\phi_j \in F_{\vec{n}}$, $j = 1, \ldots, n$ such that $\int \phi_j(x)\psi_k(x)dx = \delta_{j,k}$. Then $\phi_1, \ldots, \phi_n$ is a basis of $F_{\vec{n}}$ which is biorthogonal to $\psi_1, \ldots, \psi_n$. □

From now on, we assume that $|\vec{n}| = |\vec{m}|$ and $F_{\vec{n}} \cap G_{\vec{m}}^{\perp} = \{0\}$. According to Lemma 4.1 there exist biorthogonal bases $\phi_1, \ldots, \phi_n$ of $F_{\vec{n}}$ and $\psi_1, \ldots, \psi_n$ of $G_{\vec{m}}$. We define the kernel $K(x, y)$ as

$$K(x, y) = \sum_{j=1}^{n} \phi_j(x)\psi_j(y). \tag{4.4}$$

Lemma 4.2 *The kernel $K(x, y)$ is the kernel of the (non-orthogonal) projection operator onto $F_{\vec{n}}$ parallel to $G_{\vec{m}}^{\perp}$.*

Proof. Define the operator K on $L^2(\mathbb{R})$ as

$$(Kh)(x) = \int K(x, y)h(y)dy. \tag{4.5}$$

By the definition (4.4) of the function $K(x, y)$ it is clear that $Kh = 0$ if $h \in G_{\vec{m}}^{\perp}$. In the same way it is obvious that $Kh = h$ if $h = \phi_k$, for $k = 1, \ldots, n$. By linearity it then follows that $Kh = h$ for every $h \in F_{\vec{n}}$. Because $F_{\vec{n}} \oplus G_{\vec{m}}^{\perp} = L^2(\mathbb{R})$, Lemma 4.2 follows immediately. □

Lemma 4.2 also implies that the kernel K is independent of the chosen biorthogonal bases. Now we arrive at the main result of this paper:

Theorem 4.3 *The kernel $K(x, y)$, as defined in* (4.4)*, can be written in terms of the solution of the Riemann-Hilbert problem for Y of section 3.1 in the following way:*

$$K(x, y) = \frac{1}{2\pi i(x - y)} \begin{bmatrix} 0 & \cdots & 0 & w_{2,1}(y) & \cdots & w_{2,q}(y) \end{bmatrix} Y_+^{-1}(y)Y_+(x) \begin{bmatrix} w_{1,1}(x) \\ \vdots \\ w_{1,p}(x) \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4.6}$$

Proof. Define the operator L on $L^2(\mathbb{R})$ as

$$(Lh)(x) = \int L(x, y)h(y)dy, \tag{4.7}$$

where $L(x, y)$ denotes the right-hand side of (4.6). It is enough to prove the following two things:

(a) $Lh = 0$ if $h \in G_{\vec{m}}^{\perp}$,

(b) $Lh = h$ if $h \in F_{\vec{n}}$.

Indeed, if this is the case, then L is the projection operator onto $F_{\vec{n}}$ parallel to $G_{\vec{m}}^{\perp}$, and according to Lemma 4.2, we get that $K(x, y)$ is the kernel of L, and consequently $K(x, y) = L(x, y)$.

(a) First let $h \in G_{\vec{m}}^{\perp}$. For ease of notation we will use $\begin{bmatrix}\vec{w}_1(x) & \vec{0}\end{bmatrix}$ and $\begin{bmatrix}\vec{0} & \vec{w}_2(y)\end{bmatrix}$ instead of $\begin{bmatrix}w_{1,1}(x) & \cdots & w_{1,p}(x) & 0 & \cdots & 0\end{bmatrix}$ and $\begin{bmatrix}0 & \cdots & 0 & w_{2,1}(y) & \cdots & w_{2,q}(y)\end{bmatrix}$, where in each case the number of zeros is such that the length of the vectors is $p+q$. We then have that

$$\begin{aligned}(Lh)(x) =& \frac{1}{2\pi i}\int h(y)\begin{bmatrix}\vec{0} & \vec{w}_2(y)\end{bmatrix}\frac{Y_+^{-1}(y)-Y_+^{-1}(x)}{x-y}Y_+(x)\begin{bmatrix}\vec{w}_1(x) & \vec{0}\end{bmatrix}^t dy \\ &+\frac{1}{2\pi i}\int h(y)\begin{bmatrix}\vec{0} & \vec{w}_2(y)\end{bmatrix}\frac{Y_+^{-1}(x)}{x-y}Y_+(x)\begin{bmatrix}\vec{w}_1(x) & \vec{0}\end{bmatrix}^t dy.\end{aligned} \tag{4.8}$$

Because $Y_+^{-1}(x)Y_+(x) = I$ and $\begin{bmatrix}\vec{0} & \vec{w}_2(y)\end{bmatrix}\begin{bmatrix}\vec{w}_1(x) & \vec{0}\end{bmatrix}^t = 0$, the second term of (4.8) is equal to zero. The form of the solution of the Riemann-Hilbert problem for $X = Y^{-t}$ as given by (3.28) and (3.29) implies that the last q rows of

$$\frac{Y_+^{-1}(y)-Y_+^{-1}(x)}{x-y} \tag{4.9}$$

consist of polynomials in the variable y such that for $j = 1,\dots,q$, and $k = 1,\dots,p+q$,

$$\deg\left[\frac{Y_+^{-1}(y)-Y_+^{-1}(x)}{x-y}\right]_{p+j,k} \leq m_j - 1. \tag{4.10}$$

This implies that for each fixed $x \in \mathbb{R}$, each entry of the row vector

$$\begin{bmatrix}\vec{0} & \vec{w}_2(y)\end{bmatrix}\frac{Y_+^{-1}(y)-Y_+^{-1}(x)}{x-y}$$

belongs to $G_{\vec{m}}$. Because $h \in G_{\vec{m}}^{\perp}$, the first term of (4.8) is equal to zero as well. Thus $Lh = 0$ and this proves (a).

(b) Now let $h \in F_{\vec{n}}$. Then $h(x) = \sum_{j=1}^{p} A_j(x)w_{1,j}(x)$, where A_j is a polynomial of degree less than or equal to $n_j - 1$. We write $\vec{A} = (A_1,\dots,A_p)$ and $\begin{bmatrix}\vec{A}(x) & \vec{0}\end{bmatrix} = \begin{bmatrix}A_1(x) & \cdots & A_p(x) & 0 & \cdots & 0\end{bmatrix}$. We then have that

$h(x) = \left[\vec{A}(x) \quad \vec{0}\right] \left[\vec{w}_1(x) \quad 0\right]^t$ and so

$$(Lh)(x) = \frac{1}{2\pi i} \int \frac{\left[\vec{A}(y) - \vec{A}(x) \quad \vec{0}\right]}{x-y} \left[\vec{w}_1(y) \quad \vec{0}\right]^t \left[\vec{0} \quad \vec{w}_2(y)\right] Y_+^{-1}(y) Y_+(x) \left[\vec{w}_1(x) \quad \vec{0}\right]^t dy$$
$$+ \frac{1}{2\pi i} \int \left[\vec{A}(x) \quad \vec{0}\right] \left[\vec{w}_1(y) \quad \vec{0}\right]^t \left[\vec{0} \quad \vec{w}_2(y)\right] Y_+^{-1}(y) \frac{Y_+(x)}{x-y} \left[\vec{w}_1(x) \quad \vec{0}\right]^t dy. \tag{4.11}$$

We will deal first with the first term in the right-hand side of (4.11). We have the combination $\left[\vec{0} \quad \vec{w}_2(y)\right] Y_+^{-1}(y)$ which is a row vector whose kth entry is

$$\sum_{l=1}^{q} \left(Y_+^{-1}\right)_{p+l,k} (y) w_{2,l}(y) = \sum_{l=1}^{q} X_{k,p+l}(y) w_{2,l}(y)$$

since $Y^{-1} = X^t$, see (3.23). The functions $X_{k,p+l}$ are certain multiple orthogonal polynomials of mixed type given explicitly by formulae (3.28) and (3.29). Then it follows that

$$\left(\left[\vec{0} \quad \vec{w}_2(y)\right] Y_+^{-1}(y)\right)_k = 2\pi i Q^{(I,k)}_{\vec{m},\vec{n}-\vec{e}_k}(y; \vec{w}_2, \vec{w}_1) \qquad \text{for } k = 1, \dots, p \tag{4.12}$$

and

$$\left(\left[\vec{0} \quad \vec{w}_2(y)\right] Y_+^{-1}(y)\right)_{p+k} = Q^{(II,k)}_{\vec{m}+\vec{e}_k,\vec{n}}(y; \vec{w}_2, \vec{w}_1) \qquad \text{for } k = 1, \dots, q. \tag{4.13}$$

Since $\frac{A_j(y)-A_j(x)}{x-y}$ is a polynomial of degree $\leq n_j - 2$ in the variable y for $j = 1, \dots, p$, we have for each fixed $x \in \mathbb{R}$,

$$\frac{\left[\vec{A}(y) - \vec{A}(x) \quad \vec{0}\right]}{x-y} \left[\vec{w}_1(y) \quad \vec{0}\right]^t \in F_{\vec{n}-\sum_{j=1}^{p} \vec{e}_j}. \tag{4.14}$$

From the defining properties of the multiple orthogonal polynomials of mixed type, it follows that each of the functions (4.12) and (4.13) is orthogonal to $F_{\vec{n}-\sum_{j=1}^{p} \vec{e}_j}$. Then it follows that the first integral in the right-hand side of (4.11) is zero for every x.

Now we come to the second term in the right-hand side of (4.11). We are going to show that for every $x \in \mathbb{R}$,

$$\frac{1}{2\pi i} \int \left[\vec{w}_1(y) \quad \vec{0}\right]^t \left[\vec{0} \quad \vec{w}_2(y)\right] Y_+^{-1}(y) \frac{Y_+(x)}{x-y} dy = \begin{bmatrix} I_p & * \\ * & * \end{bmatrix}, \tag{4.15}$$

where $*$ represents an unspecified unimportant entry (which may actually be a divergent integral). Having (4.15) we easily see that the second term in the right-hand side of (4.11) reduces to $\left[\vec{A}(x) \quad \vec{0}\right] \left[\vec{w}_1(x) \quad \vec{0}\right]^t = h(x)$, independent of what the unspecified entries are (even if they are divergent integrals).

In order to establish (4.15) we note that the jump condition (3.1) written in the form

$$Y_+(y) = Y_-(y) \left(I + \left[\vec{w}_1(y) \quad \vec{0}\right]^t \left[\vec{0} \quad \vec{w}_2(y)\right] \right)$$

implies that

$$\begin{aligned} \left[\vec{w}_1(y) \quad \vec{0}\right]^t \left[\vec{0} \quad \vec{w}_2(y)\right] Y_+^{-1}(y) &= Y_-^{-1}(y) \left(Y_+(y) - Y_-(y)\right) Y_+^{-1}(y) \\ &= Y_-^{-1}(y) - Y_+^{-1}(y). \end{aligned} \tag{4.16}$$

Thus the left-hand side of (4.15) is

$$\frac{1}{2\pi i} \int \frac{Y_-^{-1}(y) - Y_+^{-1}(y)}{x-y} Y_+(x) dy. \tag{4.17}$$

Let $z \in \mathbb{C}$ with $\operatorname{Im} z > 0$. Then $Y^{-1}(y)/(z-y)$ is analytic in the lower half-plane and from the Riemann-Hilbert problem satisfied by $X = Y^{-t}$, it follows that for $k = 1, \ldots, p$, and $l = 1, \ldots, p+q$,

$$\frac{\left[Y^{-1}(y)\right]_{k,l}}{z-y} = O(y^{-n_k-1}) \qquad \text{as } y \to \infty. \tag{4.18}$$

This implies that (here we use $n_k \geq 1$)

$$\frac{1}{2\pi i} \int \frac{\left[Y_-^{-1}(y)\right]_{k,l}}{z-y} dy = 0, \qquad \text{for } k = 1, \ldots, p, \quad l = 1, \ldots, p+q. \tag{4.19}$$

Similarly, we have that $Y^{-1}(y)/(z-y)$ is analytic in the upper half-plane but with a pole at $y = z$. Then if we calculate the same integral as in (4.19) but with Y_- replaced by Y_+, and we use the decay property (4.18), the only contribution comes form the residue at $y = z$ and the result is

$$\frac{1}{2\pi i}\int \frac{\left[Y_+^{-1}(y)\right]_{k,l}}{z-y}dy = -\left(Y^{-1}(z)\right)_{k,l}, \qquad \text{for } k = 1,\ldots,p, \quad l = 1,\ldots,p+q. \tag{4.20}$$

From (4.19) and (4.20) it follows that

$$\left(\frac{1}{2\pi i}\int \frac{Y_-^{-1}(y) - Y_+^{-1}(y)}{z-y}dy\right) Y(z) = \begin{bmatrix} I_p & * \\ * & * \end{bmatrix} \qquad \text{for } \operatorname{Im} z > 0. \tag{4.21}$$

Letting $z \to x \in \mathbb{R}$, it follows that

$$\frac{1}{2\pi i}\int \frac{Y_-^{-1}(y) - Y_+^{-1}(y)}{x-y} Y_+(x)dy = \begin{bmatrix} I_p & * \\ * & * \end{bmatrix} \qquad \text{for } x \in \mathbb{R}, \tag{4.22}$$

which implies (4.15) by (4.16). As noted after (4.15) it then follows that the second term in the right-hand side of (4.11) is equal to $h(x)$. Since we already know that the first term is equal to 0, we have proven that $Lh = h$. This completes the proof of Theorem 4.3. □

5 The Christoffel-Darboux formula

Theorem 4.3 implies a Christoffel-Darboux formula for multiple orthogonal polynomials of mixed type. We assume as before that $|\vec{n}| = |\vec{m}|$ and that $F_{\vec{n}}$ and $G_{\vec{m}}$ are both $|\vec{n}|$-dimensional subspaces of $L^2(\mathbb{R})$ such that $F_{\vec{n}} \cap G_{\vec{m}}^{\perp} = \{0\}$.

Corollary 5.1 *Let K be the kernel defined in* (4.4). *We can write the kernel in terms of the multiple orthogonal polynomials of mixed type defined in* (2.11) *and* (2.12) *as follows:*

$$\begin{aligned}(x-y)K(x,y) = &\sum_{j=1}^{p} Q^{(II,j)}_{\vec{n}+\vec{e}_j,\vec{m}}(x;\vec{w}_1,\vec{w}_2) Q^{(I,j)}_{\vec{m},\vec{n}-\vec{e}_j}(y;\vec{w}_2,\vec{w}_1) \\ &- \sum_{k=1}^{q} Q^{(I,k)}_{\vec{n},\vec{m}-\vec{e}_k}(x;\vec{w}_1,\vec{w}_2) Q^{(II,k)}_{\vec{m}+\vec{e}_k,\vec{n}}(y;\vec{w}_2,\vec{w}_1).\end{aligned} \tag{5.1}$$

Proof. The kernel K and the multiple orthogonal polynomials of mixed type are well defined because of Corollary 2.3 and Lemma 4.1. The entries of $Y_+(x)$ in the first p columns are given by (3.4) and (3.5). The entries of $Y_+^{-1}(y) = X^t(y)$ in the last q rows are given by (3.28) and (3.29). Inserting these formulae into (4.6) we arrive at (5.1). $\square$

Remark 5.2 The usual monic orthogonal polynomials on the real line with weight function $w(x)$ satisfy the classical Christoffel-Darboux formula (1.1). By putting $p = q = 1$ in formula (5.1) and taking into account that the type I normalization for the multiple orthogonal polynomials of mixed type is different from the normalization used for monic orthogonal polynomials, we can see that (5.1) reduces to (1.1) in case $p = q = 1$.

In [4] the special case $p = 1$ and $q = 2$ is considered in connection with random matrices with external source. This leads to a kernel built out of multiple orthogonal polynomials with respect to 2 different weights for which a Christoffel-Darboux kernel was given. In [12] the Christoffel-Darboux formula was generalized to multiple orthogonal polynomials with respect to q different weights:

$$(x-y)K(x,y) = P_{\vec{n}}(x)Q_{\vec{n}}(y) - \sum_{k=1}^{q} \frac{h_{\vec{n}}^{(k)}}{h_{\vec{n}-\vec{e}_k}^{(k)}} P_{\vec{n}-\vec{e}_k}(x)Q_{\vec{n}+\vec{e}_k}(y). \qquad (5.2)$$

Here $P_{\vec{n}}$ is the multiple orthogonal polynomial of type II, $Q_{\vec{n}}$ is the linear form constructed out of the multiple orthogonal polynomials of type I, with the type I normalization as described in section 2, and the $h_{\vec{n}}^{(k)}$ and $h_{\vec{n}-\vec{e}_k}^{(k)}$ are certain constants. The formula (5.2) is the special case $p = 1$ of (5.1).

Remark 5.3 When we take the multi-indices $\vec{n}$ and $\vec{m}$ in a way such that

$$n_1 \le n_2 \le \cdots \le n_p \le n_1 + 1 \quad \text{and} \quad m_1 \le m_2 \le \cdots \le m_q \le m_1 + 1, \qquad (5.3)$$

then the multiple orthogonal polynomials of mixed type are vector polynomials orthogonal with respect to the weight

$$W(x) = \begin{bmatrix} w_{1,1}(x)w_{2,1}(x) & w_{1,1}(x)w_{2,2}(x) & \cdots & w_{1,1}(x)w_{2,q}(x) \\ w_{1,2}(x)w_{2,1}(x) & w_{1,2}(x)w_{2,2}(x) & \cdots & w_{1,2}(x)w_{2,q}(x) \\ \vdots & \vdots & \vdots & \vdots \\ w_{1,p}(x)w_{2,1}(x) & w_{1,p}(x)w_{2,2}(x) & \cdots & w_{1,p}(x)w_{2,q}(x) \end{bmatrix}. \qquad (5.4)$$

In [27], Sorokin and Van Iseghem obtained a Christoffel-Darboux formula for vector polynomials. Their formula has $\frac{p(p+1)}{2} + \frac{q(q+1)}{2}$ terms, while ours has only $p+q$ terms.

6 Non-intersecting Brownian motions

Our motivation for introducing the multiple orthogonal polynomials of mixed type came from the theory of non-intersecting Brownian motions. Consider n one-dimensional Brownian motions which start at n fixed points $a_1 < a_2 < \cdots < a_n$ at time $t = 0$ and end at n fixed points $b_1 < b_2 < \cdots < b_n$ at time $t = 1$. Let $p_{n,t}(x_1, \dots, x_n)$ denote the probability density that at time t, with $0 < t < 1$, the paths are at the positions $x_1, \dots, x_n$, conditioned on the event that the paths do not intersect in the full time interval $(0, 1)$. Then it follows from a result of Karlin and McGregor [22] that

$$p_{n,t}(x_1, \dots, x_n) = \frac{1}{Z_n} \det(P(t, a_j, x_k))_{j,k=1}^{n} \det(P(1-t, b_j, x_k))_{j,k=1}^{n}, \tag{6.1}$$

where

$$P(t, a, x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(x-a)^2} \tag{6.2}$$

is the transition probability for the one-dimensional Brownian motion and Z_n is a normalization constant. Note that (6.1) is an example of a biorthogonal ensemble [7].

Consider now the confluent case that some of the starting points and some of the endpoints coincide. Suppose that the n non-intersecting Brownian motions start at p different points a_j, $j = 1, \dots, p$, where a_j appears with multiplicity n_j, and end at q different points b_j, $j = 1, \dots, q$, where b_j appears with multiplicity m_j. Let $\vec{n} = (n_1, \dots, n_p)$ and $\vec{m} = (m_1, \dots, m_q)$ and

$$F_{\vec{n}} = \{\sum_{j=1}^{p} A_j(x) P(t, a_j, x) \mid A_j \text{ polynomial with } \deg(A_j) \leq n_j - 1\} \tag{6.3}$$

and

$$G_{\vec{m}} = \{\sum_{j=1}^{q} B_j(x) P(1-t, b_j, x) \mid B_j \text{ polynomial with } \deg(B_j) \leq m_j - 1\}. \tag{6.4}$$

So these are the spaces (2.15) and (2.16) associated with the vectors of weights $\vec{w}_1 = (w_{1,1}, \ldots, w_{1,p})$ and $\vec{w}_2 = (w_{2,1}, \ldots, w_{2,p})$ where

$$w_{1,j}(x) = P(t, a_j, x), \qquad \text{for } j = 1, \ldots, p,$$

and

$$w_{2,j}(x) = P(1-t, b_j, x), \qquad \text{for } j = 1, \ldots, q.$$

Lemma 6.1 *The spaces $F_{\vec{n}}$ and $G_{\vec{m}}$ are n-dimensional and*

$$F_{\vec{n}} \cap G_{\vec{m}}^{\perp} = \{0\}. \tag{6.5}$$

Proof. The statement about the dimensions are obvious. The proof of (6.5) is based on the following facts:

(1) Both $F_{\vec{n}}$ and $G_{\vec{m}}$ are Chebyshev spaces on $\mathbb{R}$, which means that any non-zero function in one of these spaces has at most $n-1$ zeros on $\mathbb{R}$, see e.g. [8]. To show this, we note that by an example given in [24, Chapter 4, §4], the functions

$$e^{\beta_1 x}, \ldots, x^{n_1-1} e^{\beta_1 x}, \ldots, e^{\beta_p x}, \ldots, x^{n_p-1} e^{\beta_p x} \tag{6.6}$$

form a Chebyshev system of order $n-1$ on $\mathbb{R}$ whenever $\beta_1, \ldots, \beta_p$ are distinct real numbers. Taking $\beta_j = a_j/t$, and multiplying the functions (6.6) by the common factor $e^{-\frac{1}{2t}x^2}$, we obtain a basis of $F_{\vec{n}}$, and so $F_{\vec{n}}$ is a Chebyshev space on $\mathbb{R}$. Similarly we have that $G_{\vec{m}}$ is a Chebyshev spaces on $\mathbb{R}$.

(2) For any set of distinct real points $x_1, \ldots, x_m$ with $m \leq n-1$, there exist functions $f \in F_{\vec{n}}$, $g \in G_{\vec{m}}$, such that f and g change sign exactly at each of these points. This is a general property of Chebyshev spaces, see [8, Chapter 3.1, exercise E.11].

Now let $f \in F_{\vec{n}}$ be non-zero. Then f has at most $n-1$ real zeros of by (1). Let $x_1, \ldots, x_m$ with $m \leq n-1$ be the zeros of odd multiplicity (so that f has a sign change at these points). By (2) there is a function $g \in G_{\vec{m}}$ which also changes sign exactly at these points. Then fg has no sign change on $\mathbb{R}$, and therefore $\int f(x)g(x)dx > 0$. Thus f does not belong to $G_{\vec{m}}^{\perp}$ and (6.5) follows. □

By Lemmas 6.1 and 4.1 there exist biorthogonal bases $\phi_1, \dots, \phi_n$ of $F_{\vec{n}}$ and $\psi_1, \dots, \psi_n$ of $G_{\vec{m}}$. Let K_n be the projection kernel

$$K_n(x, y) = \sum_{j=1}^{n} \phi_j(x)\psi_j(y). \tag{6.7}$$

Then in the confluent case the probability density (6.1) to find the Brownian paths at time t at the positions $x_1, \dots, x_n$ can be written as

$$p_{n,t}(x_1, \dots, x_n) = \frac{1}{n!} \det(\phi_j(x_k))_{j,k=1}^{n} \det(\psi_j(x_k))_{j,k=1}^{n} = \frac{1}{n!} \det(K_n(x_j, x_k))_{j,k=1}^{n}.$$

Moreover, all correlation functions have determinantal form with kernel K_n. That is, if

$$r_m(x_1, \dots, x_m) = \frac{n!}{(n-m)!} \underbrace{\int \cdots \int}_{n-m \text{ times}} p_{n,t}(x_1, \dots, x_m, x_{m+1}, \dots x_n) dx_{m+1} \cdots dx_n$$

denotes the m-point correlation function, then

$$r_m(x_1, \dots, x_m) = \det(K_n(x_j, x_k))_{j,k=1}^{m}$$

for every $m = 1, \dots, n$.

There is a substantial literature on determinantal point processes and non-intersecting random paths see e.g. the recent surveys [20, 21, 25, 28] and references cited therein.

Our Theorem 4.3 relates the kernel K_n to the Riemann-Hilbert problem for multiple orthogonal polynomials of mixed type. This opens up the possibility to analyze the kernel in the large n limit with the Deift/Zhou steepest descent method for Riemann-Hilbert problems.

To obtain interesting limit behavior, one first modifies the transition probability (6.2) to

$$P_n(t, a, x) = \frac{\sqrt{n}}{\sqrt{2\pi t}} e^{-\frac{n}{2t}(x-a)^2} \tag{6.8}$$

so that the overall variance of the Brownian paths is reduced with increasing n. With increasing n, the starting points $a_1, \dots, a_p$ and the endpoints $b_1, \dots, b_q$ remain fixed while the corresponding multiplicities $n_1, \dots, n_p$ and $m_1, \dots, m_q$ increase with n, such that the limits

$$\lim_{n\to\infty} \frac{n_j}{n} \quad \text{and} \quad \lim_{n\to\infty} \frac{m_j}{n}$$

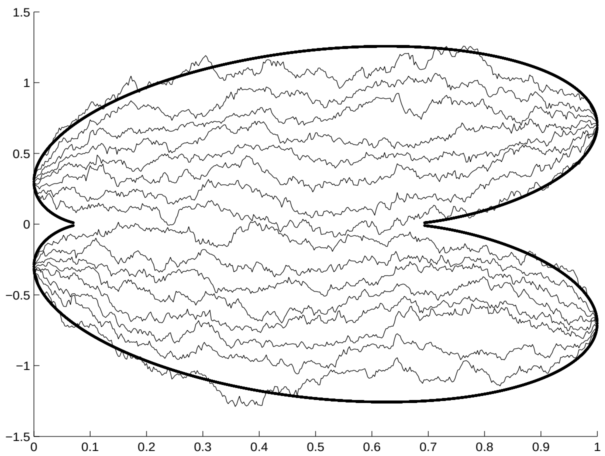

Figure 4: Non-intersecting Brownian motions which start and end at two different points.

exist and are positive.

Based on the experience with the case $p = 1$ and $q = 2$ that was developed in [3, 5, 6] we expect the following to hold true. In the limiting regime described above, with probability one, the Brownian paths fill out a bounded region as illustrated in Figure 4. Figure 4 shows a possible boundary curve in the case of two starting points and two endpoints. Here we see that two groups of paths start from two different starting points and they come together and merge at a certain critical time. Then they continue as one group until at a second critical time they split again into two groups that end at the two different endpoints. The boundary curve is smooth except for cusp singularities that arise when two groups of paths come together or split. We expect that this behavior is generic for general p and q.

At any time $t \in (0, 1)$ we further expect that the correlation kernel K_n has a scaling limit which is equal to the usual scaling limits from random matrix theory. That is, if we scale around a point (t, x) lying strictly inside the boundary curve then we expect the sine kernel in the limit and for a usual point (t, x) on the boundary (not a cusp point) we expect the Airy kernel.

At the cusp singularities we expect that the kernel K_n has the Pearcey kernel as a double scaling limit. This Pearcey kernel arose first in the works of Brézin and Hikami [9, 10] in the context of Gaussian random matrices with external source. In our notation this corresponds to $p = 1$ and $q = 2$.

A detailed treatment based on a double integral representation of the kernel was made by Tracy and Widom [29]. These authors also considered an extended Pearcey kernel and a Pearcey process which involves the limiting joint distributions at several scaled times near the critical times. The Pearcey process also appears in the recent papers [1, 26].

For an extension of the above results to more general values of p and q the Riemann-Hilbert problem that we gave in this paper might be useful. Indeed, if the Deift/Zhou steepest descent analysis can be made to work on this Riemann-Hilbert problem then the scaling limits of the kernel can be derived. We plan to report on this in a later publication.

Acknowledgements

The authors are supported by FWO research projects G.0176.02 and G.0455.04, by K.U.Leuven research grant OT/04/24, by INTAS Research Network NeCCA 03-51-6637, by NATO Collaborative Linkage Grant PST.CLG.979738, by grant BFM2001-3878-C02-02 of the Ministry of Science and Technology of Spain, by the European Science Foundation Program MISGAM and by the European Union through the FP6 Marie Curie RTN ENIGMA (Contract number MRTN-CT-2004-5652).

References

[1] M. Adler and P. van Moerbeke, PDE's for the Gaussian ensemble with external source and the Pearcey distribution, preprint math.PR/0509047.

[2] A.I. Aptekarev, Multiple orthogonal polynomials, J. Comput. Appl. Math. 99 (1998), 423–447.

[3] A.I. Aptekarev, P.M. Bleher, and A.B.J. Kuijlaars, Large n limit of Gaussian random matrices with external source, part II, Comm. Math. Phys. 259 (2005), 367–289.

[4] P.M. Bleher and A.B.J. Kuijlaars, Random matrices with external source and multiple orthogonal polynomials, Int. Math. Research Notices 2004, no 3 (2004), 109–129.

[5] P.M. Bleher and A.B.J. Kuijlaars, Large n limit of Gaussian random matrices with external source, part I, Comm. Math. Phys. 252 (2004), 43–76.

[6] P.M. Bleher and A.B.J. Kuijlaars, Large n limit of Gaussian random matrices with external source, part III: double scaling limit, preprint math.ph/0602064, to appear in Comm. Math. Phys.

[7] A. Borodin, Biorthogonal ensembles, Nucl. Phys. B 536 (1998), 704–732.

[8] P. Borwein and T. Erdélyi, Polynomials and polynomial inequalities, Springer-Verlag, Graduate Texts in Mathematics, Volume 161, New York (1995).

[9] E. Brézin and S. Hikami, Universal singularity at the closure of the gap in a random matrix theory, Phys. Rev. E 57 (1998), 4140–4149.

[10] E. Brézin and S. Hikami, Level spacing of random matrices in an external source, Phys. Rev. E 58 (1998), 7176–7185.

[11] T.S. Chihara, An Introduction to Orthogonal Polynomials, Gordon and Breach, New York (1978).

[12] E. Daems and A.B.J. Kuijlaars, A Christoffel-Darboux formula for multiple orthogonal polynomials, J. Approx. Theory 130 (2004), 188–200.

[13] P. Deift, Orthogonal Polynomials and Random Matrices: a Riemann-Hilbert approach. Courant Lecture Notes in Mathematics Vol. 3, Amer. Math. Soc., Providence R.I. 1999.

[14] P. Deift, Riemann-Hilbert methods in the theory of orthogonal polynomials, preprint math.CA/0603309.

[15] P. Deift, T. Kriecherbauer, K.T.-R. McLaughlin, S. Venakides and X. Zhou, Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory, Commun. Pure Appl. Math. 52 (1999), no. 11, 1335–1425.

[16] P. Deift, T. Kriecherbauer, K.T.-R. McLaughlin, S. Venakides, and X. Zhou, Strong qsymptotics of orthogonal polynomials with respect to exponential weights, Commun. Pure Appl. Math. 52 (1999), no. 12, 1491–1552.

[17] P. Deift and X. Zhou, A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation, Ann. Math. (2) 137 (1993) no. 2, 295–368.

[18] A.J. Durán and F.A. Grünbaum, A survey on orthogonal matrix polynomials satisfying second order differential equations, J. Comput. Appl. Math. 178 (2005), 169–190.

[19] A.S. Fokas, A.R. Its, and A.V. Kitaev, The isomonodromy approach to matrix models in 2D quantum gravity, Commun. Math. Phys. 147 (1992), 395–430.

[20] J.B. Hough, M. Krishnapur, Y. Peres, and B. Virag Determinantal processes and independence, Probab. Surv. 3 (2006), 206–229.

[21] K. Johansson, Random matrices and determinantal processes, preprint math-ph/0510038.

[22] S. Karlin and J. McGregor, Coincidence probabilities, Pacific J. Math., 9 (1959), 1141–1164.

[23] M.L. Mehta, Random Matrices, 2nd edition, Academic Press, Boston, 1991.

[24] E. M. Nikishin and V. N. Sorokin, Rational approximations and orthogonality, Amer. Math. Soc. Transl. (2), vol. 92, Amer. Math. Soc., Providence, Rhode Island, 1991.

[25] N. O'Connell, Random matrices, non-colliding processes and queues. In: Séminaire de Probabilités, XXXVI, Lecture Notes Math. 1801, Springer, Berlin, 2003, pp. 165–182.

[26] A. Okounkov and N. Reshetikhin, Random skew plane partitions and the Pearcey process, preprint math.CO/0503508.

[27] V.N. Sorokin and J. Van Iseghem, Algebraic aspects of matrix orthogonality for vector polynomials, J. Approx. Theory 90 (1997), 97–116.

[28] A. Soshnikov, Determinantal random point fields, Russian Mathematical Surveys 55 (2000), 923–975.

[29] C. Tracy and H. Widom, The Pearcey process, Comm. Math. Phys. 263 (2006), no. 2, 381–400.

[30] W. Van Assche and E. Coussement, Some classical multiple orthogonal polynomials, J. Comput. Appl. Math. 127 (2001), 317–347.

[31] W. Van Assche, J.S. Geronimo and A.B.J. Kuijlaars, Riemann-Hilbert problems for multiple orthogonal polynomials, Special Functions 2000: Current Perspectives and Future Directions (J. Bustoz et al., eds.), Kluwer, Dordrecht, 2001, pp. 23–59.