A. I. Aptekarev, A. Kuijlaars, Hermite–Padé approximations and multiple orthogonal polynomial ensembles, *Russian Mathematical Surveys*, 2011, Volume 66, Issue 6, 1133–1199

DOI: 10.1070/RM2011v066n06ABEH004771

Russian Math. Surveys **66**:6 1133–1199 © 2011 RAS(DoM) and LMS

Uspekhi Mat. Nauk **66**:6 123–190 DOI 10.1070/RM2011v066n06ABEH004771

Hermite–Padé approximations and multiple orthogonal polynomial ensembles

A. I. Aptekarev and A. B. J. Kuijlaars

Abstract. This paper is concerned with Hermite–Padé rational approximants of analytic functions and their connection with multiple orthogonal polynomial ensembles of random matrices. Results on the analytic theory of such approximants are discussed, namely, convergence and the distribution of the poles of the rational approximants, and a survey is given of results on the distribution of the eigenvalues of the corresponding random matrices and on various regimes of such distributions. An important notion used to describe and to prove these kinds of results is the equilibrium of vector potentials with interaction matrices. This notion was introduced by A. A. Gonchar and E. A. Rakhmanov in 1981.

Bibliography: 91 titles.

Keywords: Hermite–Padé approximants, multiple orthogonal polynomials, weak and strong asymptotics, extremal equilibrium problems for a system of measures, matrix Riemann–Hilbert problem, Christoffel–Darboux formula, matrix model with an external source, non-intersecting paths, two-matrix model.

Contents

The first author was supported by the Russian Foundation for Basic Research (grant nos. 11-01-00245 and 11-01-12045-офи-м), Programme no. 1 of the Department of Mathematical Sciences of the Russian Academy of Sciences, and the programme "Catedras de Exelencia" of Universidad Carlos III de Madrid and Bank Santander. The second author was supported by K. U. Leuven Research Grant OT/08/33, FWO-Flanders projects G.0427.09 and G.0641.11, the Belgian Interuniversity Attraction Pole P06/02, and the Spanish Ministry of Science and Innovation (grant no. MTM2008-06689-C02-01).

AMS 2010 *Mathematics Subject Classification.* Primary 41A21, 42C05, 60B20; Secondary 31A15, 60G17, 60G55.

Introduction

We proceed to define a general construction of rational functions with a common denominator which furnish an approximation of a vector of power series

$$\vec{f} = (f_1, \dots, f_p), \qquad f_j(z) = \sum_{k=0}^{\infty} \frac{f_{j,k}}{z^{k+1}}, \quad j = 1, \dots, p. \tag{0.1}$$

A vector

$$\pi_{\vec{n}} = \left(\frac{Q_{\vec{n},1}}{P_{\vec{n}}}, \dots, \frac{Q_{\vec{n},p}}{P_{\vec{n}}} \right), \qquad \vec{n} = (n_1, \dots, n_p) \in \mathbb{N}^p, \tag{0.2}$$

of rational functions with the common denominator $P_{\vec{n}}$ is called a *Hermite–Padé* approximant (of the second kind) with multi-index $\vec{n}$ to the vector $\vec{f}$ of power series if

$$P_{\vec{n}} \not\equiv 0, \qquad \deg P_{\vec{n}} \leqslant |\vec{n}\,| := n_1 + \cdots + n_p, \tag{0.3}$$

$$f_j(z) P_{\vec{n}}(z) - Q_{\vec{n},j}(z) =: R_{\vec{n},j}(z) = O\left(\frac{1}{z^{n_j+1}} \right) \quad \text{as } z \to \infty, \ j = 1, \dots, p. \tag{0.4}$$

This construction was proposed by Hermite [1] in connection with his celebrated proof of the transcendence of e. For $p = 1$, the approximants (0.2) are known as *Padé approximants*. The relations (0.4) provide $|\vec{n}\,|$ homogeneous linear equations to determine the $|\vec{n}\,|+1$ coefficients of the polynomial (0.3). When each polynomial

satisfying (0.4) has degree $\vec{n}$ (and so is uniquely defined up to a multiplicative constant), the multi-index is said to be *normal*, and the polynomial $P_{\vec{n}}$ is normalized as follows:

$$P_{\vec{n}}(z) = z^{|\vec{n}|} + \cdots . \tag{0.5}$$

For the special case (see [2]) of Markov functions

$$f_j(z) = \int \frac{d\mu_j(x)}{z - x}, \qquad j = 1, \dots, p,$$

with measures μ_j supported on the real line, the denominator $P_{\vec{n}}$ of the Hermite–Padé approximant (0.2) satisfies the orthogonality relations

$$\int P_{\vec{n}}(x) x^k \, d\mu_j(x) = 0, \qquad k = 0, 1, \dots, n_j - 1, \quad j = 1, \dots, p. \tag{0.6}$$

The equalities (0.6) are called multiple orthogonality relations, and the polynomial $P_{\vec{n}}$ is called a multiple orthogonal polynomial.

Apart from traditional applications to the theory of Diophantine approximations and the theory of approximations of analytic functions (see [3]–[5]), Hermite–Padé approximants and multiple orthogonal polynomials have proved useful in the spectral theory of higher-order non-symmetric difference operators (see [6]–[8]). Recently a connection was discovered between them and the theory of random matrices (see [9]–[14]).

A multiple orthogonal polynomial (MOP) ensemble is a probability density function on $\mathbb{R}^n$ of the form

$$\mathscr{P}(x_1, \dots, x_n) = \frac{1}{Z_n} \prod_{j>i} (x_j - x_i) \det\bigl[\varphi_i(x_j)\bigr]_{i,j=1,\dots,n}, \tag{0.7}$$

where the set $\{\varphi_1, \dots, \varphi_n\}$ of functions has the same linear span as the set

$$\{x^k w_j(x) \mid k = 0, \dots, n_j - 1, \ j = 1, \dots, p\}.$$

The latter MOP ensemble is generated by the p weight functions $w_1, \dots, w_p$ and the multi-index $\vec{n} = (n_1, \dots, n_p)$. The average characteristic polynomial

$$P_{\vec{n}}(z) = \mathsf{E}\biggl[\prod_{j=1} (z - x_j)\biggr]$$

with expectation taken with respect to (0.7) then satisfies the multiple orthogonality relations (0.6) with respect to the measures $d\mu_j(x) = w_j(x)\,dx$ (see [9], [15]). Important examples of MOP ensembles appear in random matrix theory, and in the theory of non-intersecting random paths.

In Chapter 1 of this survey we give a historical introduction to the analytic aspects of Hermite–Padé rational approximations and consider recent developments in this theory. One of these developments is considered in § 1. It is a generalization of the Markov theorem to the case when the approximated system of functions is defined by means of directed graphs connected with Markov (real-analytic)

functions. This statement comes from the fundamental paper [16] of Gonchar, Rakhmanov, and Sorokin, where a tree graph was considered for the generating system of functions. The minimization problem for the energy functional is investigated for a vector measure whose components are connected by a given interaction matrix and supported on a given system of intervals. The weak asymptotics of the approximants are obtained in terms of the solution of this problem. Here (following [17]) we allow the generating graph to contain undirected cycles, and thus consider the minimization problem for this case in the class of measures whose masses are not fixed but allowed to 'flow' between intervals.

Another problem discussed in Chapter 1 is the asymptotic behaviour of Hermite–Padé approximants of analytic functions with branch points in the complex plane. Namely, in § 2 we consider the asymptotic behaviour of Hermite–Padé approximants ($\vec{n} = (n, n)$) for two functions f_1 and f_2 with branch points in $A_1 = \{a_1, b_1\}$ and $A_2 = \{a_2, b_2\}$, respectively. Even this simplest 'complex' example exhibits a variety of different asymptotic behaviours, depending on the disposition of the pairs of branch points in the complex plane (see the detailed analysis in [18]). We choose for presentation here one generic geometric case of arrangement of the branch points. For this example the main term of the Hermite–Padé asymptotics is described by an algebraic function of third order and genus zero. We connect its Riemann surface with an equilibrium problem for a vector potential, which is analogous to the equilibrium problem considered in the previous section § 1.

The second part (Chapter 2) of this survey gives an overview of some of the recent work that has been done on MOP ensembles (0.7). § 3 presents the basic definitions. A MOP ensemble is a special case of a determinantal point process: a random point process in which all correlation functions (that is, marginal densities) are given as determinants of a correlation kernel. In the case of a MOP ensemble this correlation kernel can be expressed as a combination of multiple orthogonal polynomials of type I and type II. The Riemann–Hilbert problem for multiple orthogonal polynomials [19] is formulated, and by means of the Christoffel–Darboux formula [20] the correlation kernel is expressed in terms of the solution of the Riemann–Hilbert problem.

The remaining sections deal with three specific examples of MOP ensembles. The random matrix model with an external source leads to multiple orthogonal polynomials with weight functions of the form

$$w_j(x) = e^{-(V(x)-a_j x)}, \qquad j = 1, \dots, p,$$

where V is a fixed function and $a_1 < a_2 < \cdots < a_p$ are real numbers that come from the external source. The Gaussian case $V(x) = x^2/2$ leads to multiple Hermite polynomials. We discuss the large-n asymptotics of the Gaussian model with an external source ([11], [21], [22]) in the case $p = 2$, and in a situation with a more general V. In the latter case a crucial role is played by the vector equilibrium problem for two measures with both an external field and an upper constraint. Non-intersecting random paths and the two-matrix model are discussed in the final two sections. For non-intersecting squared Bessel paths and for the two-matrix model a vector equilibrium problem is given that is relevant for the large-n asymptotic behaviour.

We wish to emphasize that the notion of vector equilibrium problem, which is a powerful tool for investigating the problems discussed here, was first formulated by Gonchar and Rakhmanov in 1981 in their pioneering paper [23] (see also the more general statement in [24]). The authors would like to dedicate this survey to the 80th birthday of A. A. Gonchar.

Chapter 1

Hermite–Padé approximants

1. Hermite–Padé approximants of Markov-type functions

1.1. Historical remarks. Markov's paper [2] was devoted to power expansions of functions

$$f(z)=\sum_{k=0}^{\infty}\frac{c_k}{z^{k+1}}=\int_{\mathbb{R}}\frac{d\mu(x)}{z-x}\,,\qquad c_k=\int_{\mathbb{R}}x^k\,d\mu(x),\quad \mu>0, \tag{1.1}$$

which are Cauchy transforms of positive measures with compact support. In it Markov considered continued fraction representations

$$f(z)\doteq\cfrac{c_0}{z-b_0-\cfrac{a_1^2}{z-b_1-\cfrac{a_2^2}{z-b_2-\ddots}}} \tag{1.2}$$

and proved that the convergents $\pi_n(z)$ to $f(z)$ of (1.2) converge uniformly on compact subsets of the complex plane outside the interval E which supports the measure:

$$\operatorname{supp}\mu\subset E\Subset\mathbb{R}\quad\Rightarrow\quad\lim_{n\to\infty}\pi_n(z)=\int\frac{d\mu(x)}{z-x}\,,\quad z\in\overline{\mathbb{C}}\setminus E. \tag{1.3}$$

In the theory of rational approximations, functions of the form (1.1) are called *Markov functions* (also known as *resolvent functions* or *Weyl functions* in the theory of operators). The Markov functions form a class of analytic functions that are useful for investigating rational approximations, and Markov's theorem (1.3) is the starting point for these studies. Rational functions obtained by truncating the continued fractions (1.2) at finite levels (convergents) are a particular case of Padé approximants.

The denominator $P_{\vec n}$ of the Hermite–Padé approximant (0.2)–(0.4) to the system of Markov functions

$$f_j(z)=\hat\mu_j(z)=\int_{E_j}\frac{d\mu_j(x)}{z-x}\,,\qquad E_j\subset\mathbb{R},\quad j=1,\dots,p, \tag{1.4}$$

satisfies the orthogonality relations

$$\int_{E_j}P_{\vec n}(x)x^k\,d\mu_j(x)=0,\qquad k=0,1,\dots,n_j-1,\quad j=1,\dots,p. \tag{1.5}$$

Polynomials satisfying the orthogonality relations (1.5) are also known as multiple orthogonal polynomials.

In contrast to the case of conventional orthogonal polynomials ($p = 1$), the orthogonality relations (1.5) with $p > 1$ do not guarantee that the index n is normal, and so they do not guarantee the existence of a multiple orthogonal polynomial (0.5) of degree $|\vec{n}\,|$. Some general systems of Markov functions are known to have normal Hermite–Padé approximants. Among these, for example, are the following.

An *Angelesco system* [23] is defined by

$$\mathscr{A}: \ \{\hat{\mu}_j(z)\}_{j=1}^p, \quad \operatorname{supp}\mu_j \subset E_j: \ \overset{\circ}{E}_k \cap \overset{\circ}{E}_j = \varnothing, \quad k \neq j, \ k, j = 1, \dots, p. \tag{1.6}$$

It is easily ascertained (see (1.5)) that for an Angelesco system the polynomial $P_{\vec{n}}$ has n_j changes of sign in the interior $\overset{\circ}{E}_j$ of the interval E_j. This forces any arbitrary multi-index $\vec{n}$ to be normal.

A *Nikishin system* [26] is defined by means of a family of measures

$$\sigma := \{\sigma_j(x)\}_{j=1}^p, \qquad \operatorname{supp}\sigma_j \subset E_j, \qquad E_j \cap E_{j-1} = \varnothing,$$

supported in the intervals $\{E_j\}_{j=1}^p$; this family in turn generates the vector of measures $\mu = \{\mu_j(x)\}_{j=1}^p$ as follows:

$$\begin{aligned}
&d\mu_1(x) := d\sigma_1(x),\\
&d\mu_2(x) := d\langle\sigma_1, \sigma_2\rangle(x) := \left(\int_{E_2} \frac{d\sigma_2(t)}{x-t}\right) d\sigma_1(x),\\
&d\mu_j(x) := d\langle\sigma_1, \sigma_2, \dots, \sigma_j\rangle := d\big\langle\sigma_1, \langle\sigma_2, \dots, \sigma_j\rangle\big\rangle, \qquad j = 3, \dots, p.
\end{aligned}$$

It is also worth noting that all the components of μ are supported in one interval: $\operatorname{supp}\mu_j \subset E_1$. The system of Markov functions $\{\hat{\mu}_j(z)\}_{j=1}^p$ which corresponds to the vector μ is called a Nikishin system. Thus,

$$\mathscr{N}: \quad \{\hat{\mu}_j(z)\}_{j=1}^p, \quad \operatorname{supp}\mu_j \subset E_1, \quad j = 1, \dots, p. \tag{1.7}$$

The conventional (see [26], [27]) condition for the normality of a multi-index $\vec{n}$ for the Hermite–Padé approximants to a Nikishin system is as follows:

$$n_k \leqslant n_j + 1 \quad \text{for } k > j. \tag{1.8}$$

Recently this condition was relaxed. In [28] all the indices $\vec{n}$ were shown to be normal.

The weak asymptotics of Hermite–Padé approximants to Angelesco systems (that is, the nth-root asymptotics and the limit measures of their distribution of poles), and therefore the answer to the question of when they converge, were obtained by Gonchar and Rakhmanov in [23]. We note that the convergence or divergence of an Angelesco system (1.6) depends on the configuration of the intervals $\{E_j\}_{j=1}^p$. The strong asymptotics of Hermite–Padé approximants to Angelesco systems (that is, the asymptotics of the approximants themselves and the determination of the positions of the individual poles for large $|\vec{n}\,|$) were obtained in [29].

The convergence of Hermite–Padé approximants to Nikishin systems (an analogue of Markov's theorem (1.3)) for $p = 2$ was established by Nikishin himself in [26]. In contrast to Angelesco systems, Hermite–Padé approximants for Nikishin systems always converge; this was shown in [30] for arbitrary p. The weak (strong) asymptotics of Hermite–Padé approximants to Nikishin systems were studied in [31] and [16] (in [32], respectively).

Generalized Nikishin systems (so-called $\mathscr{GN}$-systems) $\{\hat{\mu}_j(z)\}_{j=1}^p$ of Markov functions were introduced in [16] through the concept of a tree graph. Without going into the details of the definition (the process of generating systems of Markov functions by means of graphs will be discussed in detail further on), we note that the $\mathscr{GN}$-systems involve both Angelesco and Nikishin systems, as well as some mixed systems. For example, for $p = 3$, here are two such systems: the system

$$\begin{aligned} d\mu_1(x) &:= d\sigma_1(x), \qquad \operatorname{supp}\sigma_1 \subset E_1, \\ d\mu_2(x) &:= d\langle\sigma_1,\sigma_2\rangle(x) = \left(\int_{E_2} \frac{d\sigma_2(t)}{x-t}\right) d\sigma_1(x), \\ d\mu_3(x) &:= d\langle\sigma_1,\sigma_3\rangle(x) = \left(\int_{E_3} \frac{d\sigma_3(t)}{x-t}\right) d\sigma_1(x) \end{aligned} \tag{1.9}$$

with disjoint intervals $\{E_j\}_{j=1}^3$, and the system

$$\begin{aligned} d\mu_1(x) &:= d\sigma_1(x), \qquad \operatorname{supp}\sigma_1 \subset E_1, \\ d\mu_2(x) &:= d\sigma_2(x), \qquad \operatorname{supp}\sigma_2 \subset E_2, \\ d\mu_3(x) &:= d\langle\sigma_1,\sigma_3\rangle(x) = \left(\int_{E_3} \frac{d\sigma_3(t)}{x-t}\right) d\sigma_1(x), \end{aligned} \tag{1.10}$$

with $E_1 \cap E_2 = \varnothing$ and $E_1 \cap E_3 = \varnothing$. For Hermite–Padé approximants to a $\mathscr{GN}$-system, a condition for a multi-index $\vec{n}$ to be normal (similar to (1.8)) was determined in [16]. In the same paper the problem of weak asymptotics is solved.

The Hermite–Padé approximants to specific systems of Markov functions connected with graphs having cycles were investigated in [3] in connection with applications to number theory. In this section, following [17], we describe in more detail the analytic properties and the asymptotic behaviour of the Hermite–Padé approximants to general systems of Markov functions connected with graphs having cycles.

1.2. Graphs and the corresponding systems of Markov functions. Let us consider a directed graph with vertex set $\mathscr{V} := \{A, B, C, \dots\}$, $\#\mathscr{V} = p + 1$, and with edges $\mathscr{E} := \{\alpha, \beta, \gamma, \dots\}$, $\#\mathscr{E} = m$. We suppose that

1) the graph is acyclic (that is, it contains no directed cycles);
2) there is a vertex O such that for each vertex $A \in \mathscr{V}$ different from O there is a directed path from O to A.

The vertex O is unique by condition 1).

We denote this graph by

$$\mathscr{G} := \mathscr{G}(\mathscr{V}, \mathscr{E}, O). \tag{1.11}$$

Let (A, B) be the set of edges connecting two adjacent vertices A and B.

The vertex set $\mathscr{V}$ can be equipped with a partial order relation as follows: let $A \preccurlyeq B$ if either $A = B$ or there exists a directed path from A to B. In the latter case we shall also write $A \prec B$.

For a given vertex $A \in \mathscr{G}$ let

$$A_+ := \{B \in \mathscr{V} \colon \exists \alpha \in (A, B) \subset \mathscr{E}\} \quad \text{and} \quad A_- := \{B \in \mathscr{V} \colon \exists \alpha \in (B, A) \subset \mathscr{E}\}$$

be the two sets of vertices nearest to A, and let

$$\mathscr{E}_{A+} := \{\alpha \in (A, B) \colon B \in A_+\} \quad \text{and} \quad \mathscr{E}_{A-} := \{\alpha \in (B, A) \colon B \in A_-\}$$

be the sets of edges going out of or coming into A. We introduce the following relations on $\mathscr{E}$:

$$\begin{aligned}
&\alpha \to \beta \quad \Leftrightarrow \quad \exists A \in \mathscr{V} \colon \quad \alpha \in \mathscr{E}_{A-},\ \ \beta \in \mathscr{E}_{A+};\\
&\alpha \uparrow\uparrow \beta \quad \Leftrightarrow \quad \exists A, B \in \mathscr{V} \colon \quad \alpha, \beta \in (A, B);\\
&\alpha \leftrightarrow \beta \quad \Leftrightarrow \quad \exists A \in \mathscr{V} \colon \quad \alpha, \beta \in \mathscr{E}_{A-} \text{ or } \alpha, \beta \in \mathscr{E}_{A+}, \text{ but } \alpha \uparrow\uparrow \beta \text{ fails to hold.}
\end{aligned}$$

Following [16], we consider the system of Markov functions generated by the graph $\mathscr{G}$. To each edge α of $\mathscr{G}$ we assign an interval $E_\alpha := [a_\alpha, b_\alpha]$ of the real axis $\mathbb{R}$ and a positive Borel measure σ_α with support in E_α; that is,

$$\forall\, \alpha \in \mathscr{E} \quad \longrightarrow \quad E_\alpha := [a_\alpha, b_\alpha] \subset \mathbb{R}, \qquad \sigma_\alpha \colon\ \sigma'_\alpha > 0 \text{ a.e. on } E_\alpha. \tag{1.12}$$

We also assume that if the edges α and β have a common vertex, then the corresponding intervals E_α and E_β do not overlap, that is,

$$\alpha \to \beta \vee \beta \to \alpha \vee \alpha \uparrow\uparrow \beta \vee \alpha \leftrightarrow \beta \quad \Rightarrow \quad E_\alpha \cap E_\beta = \varnothing. \tag{1.13}$$

Corresponding to each vertex $A \in \mathring{\mathscr{V}} := \mathscr{V} \setminus \{O\}$ there is a non-empty set $\mathscr{T}_A$ of paths $t_A = (\omega, \dots, \beta, \alpha)$ from the root vertex O to A, that is,

$$\begin{gathered}
\forall\, A \in \mathring{\mathscr{V}} \quad \longrightarrow \quad \mathscr{T}_A := \{t_A\},\\
t_A := (\omega, \dots, \beta, \alpha) \colon \quad \omega \to \dots \to \beta \to \alpha, \quad \omega \in \mathscr{E}_{O+}, \quad \alpha \in \mathscr{E}_{A-}.
\end{gathered}$$

For each chain of paths t_A of this type there is a corresponding measure μ_{t_A} defined by Nikishin's rule as follows:

$$\mu_{t_A}(x) = \langle \sigma_\omega, \dots, \sigma_\beta, \sigma_\alpha \rangle(x);$$

here

$$d\langle \sigma_1, \sigma_2 \rangle(x) := \left(\int \frac{d\sigma_2(t)}{x - t} \right) d\sigma_1(x), \quad \dots, \quad d\langle \sigma_1, \sigma_2, \dots, \sigma_j \rangle := d\big\langle \sigma_1, \langle \sigma_2, \dots, \sigma_j \rangle \big\rangle.$$

To a vertex A we also assign the function

$$\hat{\mu}_A(x) := \sum_{t_A \in \mathscr{T}_A} \int \frac{d\mu_{t_A}(t)}{x - t}\,. \tag{1.14}$$

Definition 1.1. The family $\{\hat{\mu}_A(x),\ A \in \overset{\circ}{\mathscr{V}}\}$ of functions is called a *generalized Nikishin system* ($\mathscr{G}\mathscr{N}$-system) corresponding to the graph $\mathscr{G}$.

Remark 1.1. The concept of a generalized Nikishin system corresponding to a graph was introduced in [16]. In this paper we discuss a wider class of graphs than in [16], where only tree graphs characterized by the condition

$$\mathscr{G}\colon \quad \forall\, A \in \overset{\circ}{\mathscr{V}} \quad \Rightarrow \quad \#\mathscr{E}_{A-} = 1 \tag{1.15}$$

were considered. Consequently, the class of $\mathscr{G}\mathscr{N}$-systems under discussion is larger than the class of generalized Nikishin systems corresponding to tree graphs in [16].

We give examples of various $\mathscr{G}\mathscr{N}$-systems. As has already been noted, Angelesco and Nikishin systems are $\mathscr{G}\mathscr{N}$-systems. An Angelesco system is generated by the tree graph in Fig. 1*a*,

$$\mathscr{G}\colon \quad \mathscr{E} = \mathscr{E}_{O+} \to \mathscr{A},$$

and a Nikishin system is generated by the tree graph in Fig 1*b*,

$$\mathscr{G}\colon \quad \forall\, A \in \overset{\circ}{\mathscr{V}} \quad \Rightarrow \quad \#A_- = 1, \quad \#A_+ \leqslant 1 \quad \to \quad \mathscr{N}.$$

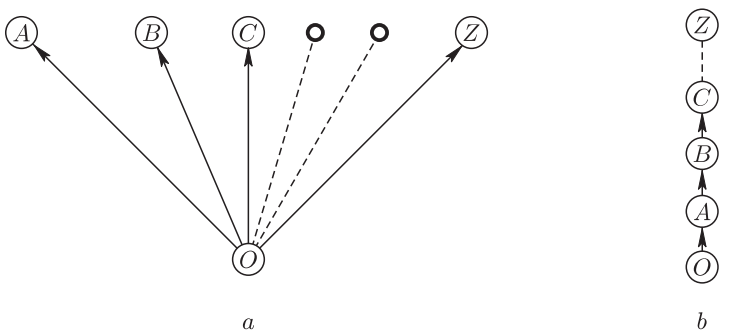

Figure 1. Graphs generating Angelesco (Fig. *a*) and Nikishin systems (Fig. *b*).

Examples of tree graphs are shown in Fig. 2. The graphs generating the systems of Markov functions (1.9) and (1.10) are shown in Figs. 2*a* and 2*b*, respectively.

Examples of graphs with undirected cycles are shown in Fig. 3 The graph in Fig. 3*a* generates a single Markov function supported in several (m) intervals. The graph in Fig. 3*b* generates two Markov functions. One of these, $\hat{\mu}_A$, has its support in the union of three intervals $E := E_\alpha \cup E_\beta \cup E_\gamma$, while the other, $\hat{\mu}_B$, has its support in the interval E_β. Also,

$$d\mu_B(x) := d\sigma_\beta(x), \qquad \operatorname{supp}\sigma_\beta \subset E_\beta,$$

$$d\mu_A(x) := \begin{cases} d\sigma_\alpha(x) & \text{on } E_\alpha, \\ \left(\displaystyle\int_{E_\delta} \frac{d\sigma_\delta(t)}{x-t}\right) d\sigma_\beta(x) & \text{on } E_\beta, \\ d\sigma_\gamma(x) & \text{on } E_\gamma. \end{cases}$$

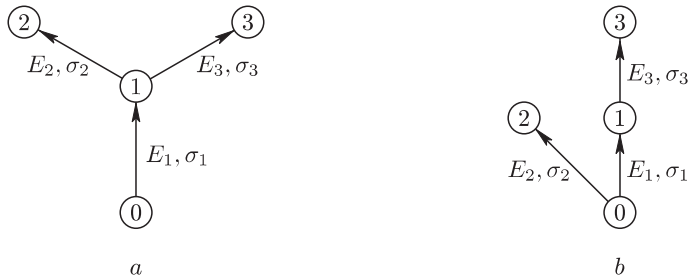

Figure 2. The tree graphs generating the systems of Markov functions (1.9) and (1.10).

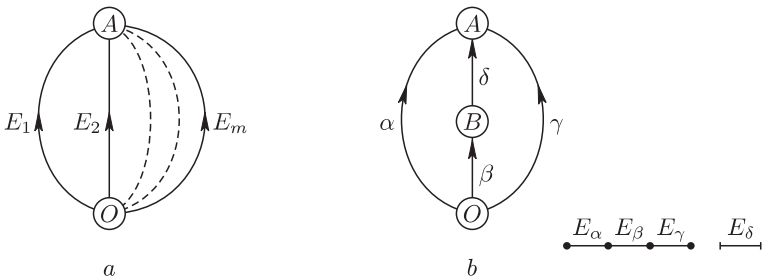

Figure 3. Graphs with undirected cycles.

Thus, graphs with undirected cycles are useful for representing Markov functions which have some part of their support in common (more precisely, one support contains the other), and the ratio of the weights on the common region of support is again a Markov function on some different interval.

We shall consider Hermite–Padé approximants to the $\mathscr{GN}$-system (1.14)

$$\vec{f} := \{\hat{\mu}_A(x),\ A \in \mathring{\mathscr{V}}\},$$

which corresponds to the graph $\mathscr{G}(\mathscr{V}, \mathscr{E}, O)$ (see (1.11)). We fix a multi-index

$$\vec{n} := \{n_A,\ A \in \mathring{\mathscr{V}}\}: \quad n_A \leqslant n_B + 1 \quad \text{if } B \prec A. \tag{1.16}$$

Then there exists a polynomial $P_{\vec{n}} \not\equiv 0$ of degree $\deg P_{\vec{n}} \leqslant |\vec{n}\,| := \sum_{A \in \mathring{\mathscr{V}}} n_A$ such that

$$R_{\vec{n},A} := P_{\vec{n}}\hat{\mu}_A - Q_{\vec{n},A} = O(z^{-n_A-1}), \qquad z \to \infty, \quad A \in \mathring{\mathscr{V}}, \tag{1.17}$$

where the $Q_{\vec{n},A}$ are some polynomials. This definition of Hermite–Padé approximants

$$\left\{ \frac{Q_{\vec{n},A}}{P_{\vec{n}}},\ A \in \mathring{\mathscr{V}} \right\} \tag{1.18}$$

leads to the following orthogonality relations:

$$\sum_{t_A \in \mathscr{T}_A} \int P_{\vec{n}}(x) x^k \, d\mu_{t_A}(x) = 0, \qquad k = 0, \dots, n_A - 1, \quad A \in \mathring{\mathscr{V}}. \tag{1.19}$$

To investigate these approximants, in addition to the functions $R_{\vec{n},A}$ of the second kind it proves useful to consider functions $\Psi_{\vec{n},A}$ which are defined by induction with respect to the partial order on the graph:

$$\Psi_{\vec{n},O} = P_{\vec{n}}, \qquad \Psi_{\vec{n},B}(x) = \sum_{A\in B_-} \sum_{\alpha\in(A,B)} \int \frac{\Psi_{\vec{n},A}(t)\, d\sigma_\alpha(t)}{t-x}, \qquad B \in \overset{\circ}{\mathscr{V}}. \quad (1.20)$$

In [16] the conventional condition (1.8) for the normality of a multi-index $\vec{n}$ for a Nikishin system was extended as follows to a generalized Nikishin system generated by a tree graph (1.15):

$$n_A \leqslant n_B + 1, \quad \text{if } B \prec A. \quad (1.21)$$

It was further shown that the polynomial $P_{\vec{n}}$ has $|\vec{n}|$ simple zeros on the union $\bigcup_{\alpha\in\mathscr{E}_{O+}} E_\alpha$ of intervals. Consequently, the indices (1.21) are normal, and hence the Hermite–Padé approximants are uniquely defined.

For a $\mathscr{G}\mathscr{N}$-system generated by an arbitrary graph $\mathscr{G}$ (see (1.11)), the questions of whether the indices are normal and whether the approximants are unique require further investigation. It can be shown, however, than under the condition (1.21) any such $P_{\vec{n}}$ has at least $|\vec{n}| - g$ simple zeros on the union $\bigcup_{\alpha\in\mathscr{E}_{O+}} E_\alpha$, where g is the cyclomatic number (the number of independent undirected cycles) of $\mathscr{G}$, that is,

$$g = \#\mathscr{E} - \#\mathscr{V} + 1. \quad (1.22)$$

Some other conclusions regarding the normality and uniqueness of Hermite–Padé approximants to arbitrary $\mathscr{G}\mathscr{N}$-systems will be stated below as corollaries to the asymptotic results.

Let

$$v = \{v_A,\ A \in \overset{\circ}{\mathscr{V}}\}, \quad \text{where } v_A > 0,\ A \in \overset{\circ}{\mathscr{V}},\ \sum_{A\in\overset{\circ}{\mathscr{V}}} v_A = 1, \text{ and } v_B \leqslant v_A \text{ if } A \prec B,$$

be a given probability distribution on $\mathscr{V}$. We consider a sequence $\mathbf{N}$ of multi-indices $\vec{n} = \{n_A,\ A \in \overset{\circ}{\mathscr{V}}\}$ such that the condition (1.21) holds and

$$\frac{n_A}{|\vec{n}\,|} \to v_A, \qquad A \in \overset{\circ}{\mathscr{V}}. \quad (1.23)$$

In this section we analyze the asymptotic behaviour of $P_{\vec{n}}$ for $\vec{n} \in \mathbf{N}$.

1.3. Weak asymptotics. By the weak asymptotics we mean the limit distributions of the zeros of the polynomial $P_{\vec{n}}$ and of the zeros of the functions $\Psi_{\vec{n},A}$ in (1.20). A way to attack such problems was proposed by Gonchar and Rakhmanov in [23]. They used an Angelesco system to show that the components of the limit measure for the distribution of the zeros of $P_{\vec{n}}$ (with supports in the intervals $\{E_j\}_{j=1}^p$ of the system (1.6)) must be components of the extremal measure in the problem of minimizing the energy functional of a vector measure with some interaction matrix of the components of the measure (see below for the details). This approach was extended in the paper [16] of Gonchar, Rakhmanov, and Sorokin to

generalized Nikishin systems generated by the tree graphs (1.15). This made it possible to broaden the class of interaction matrices under consideration. Following [17], we adapt this approach to arbitrary $\mathscr{GN}$-systems (1.14). The new feature here (apart from the fact that, as we said, the class of interaction matrices is broader) is that the extremal problem of minimizing the energy functional is treated in the class of vector measures whose components have masses that are not fixed (as they were before), but are subject to some constraints. A general problem of this kind will be formulated in § 1.3.1, where we also state Theorem 1.1 on the existence, uniqueness, and other features of the extremal vector measure. Then in § 1.3.2 we give the interaction matrices and linear relations in the general extremal problem specific to the masses of the components of the vector measures corresponding to an arbitrary graph of the form (1.11). Here we also state Theorem 1.2 on the limit distribution of the zeros of $P_{\vec{n}}$ and its multiple Cauchy transforms, and give a corollary on the weak asymptotics of Hermite–Padé approximants.

1.3.1. *Equilibrium of the potentials of vector measures with interaction matrices and linearly related masses.* We begin by setting up a general energy minimization problem for a vector measure subject to some linear restrictions on the masses of its components. As the initial data for the problem we have: a family

$$\vec{E} = (E_1, \dots, E_m) \Subset \mathbb{C}^m$$

of regular compact sets in the complex plane, a real symmetric non-negative-definite matrix

$$\mathscr{A} = (a_{kj})_{k,j=1}^m \in \mathbb{R}^{m\times m}, \qquad \mathscr{A} \geqslant 0,$$

a real $r \times m$ matrix of rank r

$$\mathscr{C} = (c_{kj})_{k,j=1}^{r,m} \in \mathbb{R}^{r\times m}, \qquad \operatorname{rank} \mathscr{C} = r,$$

and a non-zero vector

$$b = (b_1, \dots, b_r) \in \mathbb{R}^r, \qquad b \neq 0.$$

We assume in addition that the initial data satisfy the conditions

$$1)\ a_{jj} > 0, \qquad 2)\ a_{kj} = 0 \ \text{ for } \ k \neq j \ \text{ and } \ E_k \cap E_j \neq \varnothing, \qquad k, j = 1, \dots, m, \tag{1.24}$$

and that the polytope

$$\left\{ x \in \mathbb{R}^m \colon \sum_{j=1}^m c_{kj} x_j = b_k,\ k = 1, \dots, r;\ x_j \geqslant 0,\ j = 1, \dots, m \right\} \tag{1.25}$$

is bounded and non-empty.

We require some notation from potential theory. For a compact set $K \subset \mathbb{C}$ let $\mathbf{M}(K)$ be the set of all signed measures of finite variation and $\mathbf{M}^+(K)$ the set of all finite positive Borel measures ν with support $S(\nu)$ in K. The function

$$V^\nu(z) = \int_K \log \frac{1}{|z-t|}\, d\nu(t), \qquad z \in \mathbb{C}, \tag{1.26}$$

is called the logarithmic potential of a measure ν, and the integral

$$I(\nu_1,\nu_2)=\iint_{K\times K}\log\frac{1}{|x-t|}\,d\nu_1(x)\,d\nu_2(t) \tag{1.27}$$

is called the mutual energy of two measures ν_1 and ν_2. The total variation (the mass) of a measure ν will be denoted by $|\nu|$.

For a finite collection of compact sets E, we define the set $\mathbf{M}^+(\vec{E}\,)$ to be the direct product of the sets $\mathbf{M}^+(E_j)$ over all $j=1,\dots,m$. Thus, each element μ of the set $\mathbf{M}^+(\vec{E}\,)$ is a tuple of finite measures μ_j with $S(\mu_j)\subset E_j$.

For a measure $\mu\in\mathbf{M}^+(\vec{E}\,)$ with interaction matrix $\mathscr{A}$ the energy functional $J(\mu)$ and the vector potential $W^\mu=(W_1^\mu,\dots,W_m^\mu)$ are defined by

$$J(\mu)=\sum_{k,j=1}^m a_{kj}I(\mu_k,\mu_j)\quad\text{and}\quad W_k^\mu(x)=\sum_{j=1}^m a_{kj}V^{\mu_j}(x). \tag{1.28}$$

Finally, we introduce the class of measures with the masses of the components satisfying the linear relations in

$$\mathbf{M}^+_{\mathscr{C},b}(\vec{E}\,):=\left\{\mu\in\mathbf{M}^+(\vec{E}\,)\colon\ \sum_{j=1}^m c_{kj}|\mu_j|=b_k,\ k=1,\dots,r\right\} \tag{1.29}$$

(depending on the initial parameters $\mathscr{C}$ and b), and we consider the energy minimization problem

$$\begin{cases}J(\mu)\to\min,\\ \mu\in\mathbf{M}^+_{\mathscr{C},b}(\vec{E}\,).\end{cases} \tag{1.30}$$

We have the following result (see [17]).

Theorem 1.1. 1) *There exists a unique measure λ, called the extremal measure, which solves the problem* (1.30) (*that is, it minimizes the functional* (1.28) *subject to the linear relations in* (1.29) *on the masses of the components*).

2a) *There exists a tuple of constants $(l_1,\dots,l_r)$ such that the extremal measure λ minimizes the Lagrangian*:

$$\begin{cases}\mathscr{L}(\mu):=J(\mu)+\displaystyle\sum_{k=1}^r l_k\sum_{j=1}^m c_{kj}|\mu_j|\to\min,\\ \mu\in\mathbf{M}^+(\vec{E}\,).\end{cases} \tag{1.31}$$

2b) *If a measure $\lambda\in\mathbf{M}^+_{\mathscr{C},b}(\vec{E}\,)$ solves the problem* (1.31) *for some set of constants $l_1,\dots,l_r$, then λ is the extremal measure.*

3) *The extremal measure λ is the unique measure in the class* (1.29) *which satisfies the following equilibrium conditions with restrictions on the equilibrium constants*:

$$\begin{cases}W_k^\lambda(x):=\displaystyle\sum_{j=1}^m a_{kj}V^{\lambda_j}(x)\begin{cases}=\varkappa_k, & x\in S(\lambda_k),\\ \geqslant\varkappa_k, & x\in E_k,\end{cases}\quad k=1,\dots,m,\\ (\varkappa_1,\dots,\varkappa_m)\in\operatorname{Im}\mathscr{C}^T.\end{cases} \tag{1.32}$$

Remark 1.2. It is worth pointing out that the linear relations between the masses of the components of the vector measures in (1.29) are transformed in Theorem 1.1 into the linear relations between the equilibrium constants $(\varkappa_1, \dots, \varkappa_m)$ in (1.32). In other words, if the measures (with the initial data $\mathscr{C}$ and b) are allowed to 'flow' between the compact sets $E_1, \dots, E_m$, then this imposes additional restrictions on the equilibrium constants.

This remark can be illustrated by means of a trivial example. Suppose that the initial data correspond to an equilibrium measure concentrated on two disjoint intervals, that is,

$$\mathscr{A} := \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \mathscr{C} := \|1, 1\|, \qquad b := 1, \qquad E_1 \cap E_2 = \varnothing.$$

Then it is clear that

$$W_k^{\lambda}(x) := V^{\lambda_1}(x) + V^{\lambda_2}(x) = \varkappa_k, \qquad k = 1, 2, \quad x \in E_1 \cup E_2, \quad \Rightarrow \quad \varkappa_1 = \varkappa_2.$$

We note that the extremal problems (1.30) with linear restrictions on the masses (first introduced in [18]) play a key role in the study of the asymptotic behaviour of Hermite–Padé approximants for general classes of analytic functions with branch points (see [33]).

1.3.2. *The equilibrium problem for graphs and the limit distribution of the poles of the approximants.* In the general extremal problem (1.30) we specify the interaction matrices (1.28) and the linear relations in (1.29) for the total masses of the components of the vector measures; these components in turn correspond to the limit measures for the distributions of the poles of the Hermite–Padé approximants and of the zeros of the functions (1.20) for the system of Markov functions (1.14) as generated by an arbitrary graph (1.11). From the graph $\mathscr{G}$ we construct a symmetric matrix $\mathscr{A} = (a_{\alpha\beta})$ as follows:

$$a_{\alpha\beta} = \begin{cases} 2 & \text{if } \alpha = \beta \text{ or } \alpha \uparrow\uparrow \beta, \\ 1 & \text{if } \alpha \leftrightarrow \beta, \\ -1 & \text{if } \alpha \to \beta \text{ or } \beta \to \alpha, \\ 0 & \text{if the edges } \alpha \text{ and } \beta \text{ have no common vertices.} \end{cases} \tag{1.33}$$

The restrictions on the masses of the measures are as follows:

$$\left\{ \mu \in \mathbf{M}^+(\vec{E}) \colon \sum_{\alpha \in \mathscr{E}_{A-}} |\mu_\alpha| - \sum_{\beta \in \mathscr{E}_{A+}} |\mu_\beta| = v_A, \ A \in \overset{\circ}{\mathscr{V}} \right\}, \tag{1.34}$$

where the elements of $\{v_A > 0, \ A \in \overset{\circ}{\mathscr{V}}\}$ with $\sum v_A = 1$ are given in (1.23), and a sum over an empty set of indices is zero by definition.

The matrix $\mathscr{A}$ is non-negative definite. This follows either from the fact that $\mathscr{A}$ is the Gram matrix $a_{\alpha\beta} = (e_{A_2} - e_{A_1}, e_{B_2} - e_{B_1})$, where $\alpha \in (A_1, A_2)$, $\beta \in (B_1, B_2)$, and $\{e_A, \ A \in \mathscr{V}\}$ is the standard basis in $\mathbb{R}^{p+1}$, or from the equality

$$\sum_{\alpha, \beta \in \mathscr{E}} a_{\alpha\beta} x_\alpha x_\beta = \sum_{A \in \mathscr{V}} \left(\sum_{\alpha \in \mathscr{E}_{A-}} x_\alpha - \sum_{\beta \in \mathscr{E}_{A+}} x_\beta \right)^2.$$

The remaining conditions on $\mathscr{A}$ in the previous subsection are easily verified, as are the relations in (1.29).

Thus, there is a unique measure $\lambda = \{\lambda_\alpha,\ \alpha \in \mathscr{E}\}$ in the class (1.29) satisfying the equilibrium relations (1.32), that is,

$$W_\alpha^\lambda(x) := \sum_{\beta\in\mathscr{E}} a_{\alpha\beta} V^{\lambda_\beta}(x) \begin{cases} = \widetilde{\varkappa}_B - \widetilde{\varkappa}_A, & x \in S(\lambda_\alpha), \\ \geqslant \widetilde{\varkappa}_B - \widetilde{\varkappa}_A, & x \in E_\alpha, \end{cases} \tag{1.35}$$

where $\alpha \in (A, B) \subset \mathscr{E}$, and $\{\widetilde{\varkappa}_A,\ A \in \mathscr{V}\}$ is some distribution of constants over the vertices of the graph. As a result, the equilibrium constants $\varkappa_\alpha := \widetilde{\varkappa}_B - \widetilde{\varkappa}_A$ are subject to g linear relations, and we can take $\widetilde{\varkappa}_O = 0$.

The limit distributions of the zeros of the polynomials $P_{\vec{n}}$ and of the zeros of the functions $\Psi_{\vec{n},A}$ in (1.20) are represented in terms of the extremal measure λ. Suppose that $\alpha \in (A, B)$ and that $q_{\vec{n},\alpha}$ is a polynomial whose zeros, counted with multiplicities, are those of $\Psi_{\vec{n},A}$ on the interval E_α:

$$q_{\vec{n},\alpha}(z) := \prod_{x:\ \Psi_{\vec{n},A}(x)=0,\ x\in E_\alpha} (z - x) \quad \text{(by convention, } \prod_{x\in\varnothing} (z - x) := 1). \tag{1.36}$$

Let $\mu(q)$ be the equidistributed discrete measure of mass $\deg q$ on the zeros of the polynomial q:

$$\mu(q) = \sum_{x:\ q(x)=0} \delta_x.$$

We have the following result (see [17]).

Theorem 1.2. *For any $\alpha \in \mathscr{E}$ the limit relations*

$$\frac{1}{|\vec{n}\,|}\,\mu(q_{\vec{n},\alpha}) \to \lambda_\alpha$$

hold for $\vec{n} \in \mathbf{N}$ (see (1.23)). In particular,

$$\frac{1}{|\vec{n}\,|}\,\mu(P_{\vec{n}}) \to \sum_{\alpha\in O_+} \lambda_\alpha.$$

If we can show that the functions $\Psi_{\vec{n},A}\Big(\prod_{\alpha\in A_+} q_{\vec{n},\alpha}\Big)^{-1}$ have no zeros outside $\bigcup_{\alpha\in\mathscr{E}_{A+}} E_\alpha$, then we can write asymptotic formulae for $\Psi_{\vec{n},A}$. This condition holds, for example, if the graph $\mathscr{G}$ is a tree. Theorem 1.2 has the following corollary (we suppose $P_{\vec{n}}$ to be normalized so that its leading coefficient is 1).

Corollary 1.1. *Suppose that $\Psi_{\vec{n},A}(x) \neq 0$ for $x \in \mathbb{C} \setminus \bigcup_{\alpha\in\mathscr{E}_{A-}\cup\mathscr{E}_{A+}} E_\alpha$, where $A \in \mathscr{V}$, and $\widetilde{\mathbf{N}} \subset \mathbf{N}$ is some subsequence. Then*

$$\lim_{\vec{n}\in\widetilde{\mathbf{N}}} \frac{1}{|\vec{n}\,|} \log|\Psi_{\vec{n},A}(x)| = \sum_{\alpha\in\mathscr{E}_{A-}} V^{\lambda_\alpha}(x) - \sum_{\alpha\in\mathscr{E}_{A+}} V^{\lambda_\alpha}(x) - \widetilde{\varkappa}_A,$$
$$x \in \mathbb{C} \setminus \bigcup_{\alpha\in\mathscr{E}_{A-}\cup\mathscr{E}_{A+}} E_\alpha.$$

In particular,

$$\lim_{\vec{n}\in\widetilde{\mathbf{N}}} \frac{1}{|\vec{n}\,|} \log|P_{\vec{n}}(x)| = -\sum_{\alpha\in\mathscr{E}_{O+}} V^{\lambda_\alpha}(x), \qquad x\in\mathbb{C}\setminus\bigcup_{\alpha\in\mathscr{E}_{O+}} E_\alpha.$$

1.4. Proof of Theorem 1.1. 1) We shall need some properties of the energy functional (for proofs we refer the reader to [34]).

First, the principle of descent holds for the energy functional $I(\cdot)$: if $\mu_n \xrightarrow{*} \mu$, then $\varliminf_{n\to\infty} I(\mu_n) \geqslant I(\mu)$.

The functional $I(\,\cdot\,,\cdot\,)$ is a bilinear form on the linear space $\mathbf{M}(K)$ of all signed measures (charges) $\delta = \delta_+ - \delta_-$ supported on a compact set K and satisfying the condition $I(\delta_+ + \delta_-) < \infty$. The second condition that we require is that this form be positive definite. More precisely, under the (technical) assumption that $K \subset \{z\colon |z| < 1\}$ we have $I(\delta) \geqslant 0$ for any signed measure $\delta \in \mathbf{M}(K)$, and moreover, if $I(\delta) = 0$, then $\delta = 0$. In other words, $I(\,\cdot\,,\cdot\,)$ defines an inner product on $\mathbf{M}(K)$ by $(\delta_1, \delta_2) := I(\delta_1, \delta_2)$.

We prove the existence of the extremal measure. The principle of descent also holds for the functional $J(\mu) = \sum_{k,j} a_{kj} I(\mu^k, \mu^j)$ since it is valid for $a_{jj} I(\mu^j)$ and the non-diagonal terms are continuous, because $E_k \cap E_j = \varnothing$ for $a_{kj} \neq 0$ (see (1.24)).

Let $\mu_n \in \mathbf{M}^+_{\mathscr{C},b}(\vec{E}\,) = \bigl\{\mu \in \mathbf{M}^+(\vec{E}\,)\colon \sum_{j=1}^m c_{kj}|\mu_j| = b_k,\ k = 1,\dots,r\bigr\}$ be a minimizing sequence, that is,

$$J(\mu_n) \to J_0 := \inf\{J(\mu)\colon \mu \in \mathbf{M}^+_{\mathscr{C},b}(\vec{E}\,)\}.$$

Since the polytope (1.25) is bounded, the masses of the measures in $\mathbf{M}^+_{\mathscr{C},b}(\vec{E})$ are also bounded, and hence the set $\mathbf{M}^+_{\mathscr{C},b}(\vec{E})$ is compact in the weak topology. From the sequence μ_n we choose a convergent subsequence $\mu_{n_k} \to \lambda \in \mathbf{M}^+_{\mathscr{C},b}(\vec{E}\,)$, $n \in \Lambda$. On the one hand, $J(\lambda) \geqslant J_0$. On the other hand, $J_0 \geqslant J(\lambda)$ by semicontinuity. Thus, λ is an extremal measure.

To verify the uniqueness, observe that $J(\mu)$ is a convex functional, since $\mathscr{A}$ is non-negative definite. If λ and λ' were two extremal measures, then $(\lambda + \lambda')/2$ would also be an extremal measure: $J\bigl((\lambda + \lambda')/2\bigr) \leqslant \bigl(J(\lambda) + J(\lambda')\bigr)/2 = J_0$. But since $J(\lambda - \lambda') + J(\lambda + \lambda') = 2J(\lambda) + 2J(\lambda')$, we have $J(\lambda - \lambda') = 0$, and thus $(\lambda - \lambda')\mathscr{A} = 0$. Consider the charge $\nu = \lambda - \lambda'$: $\nu = (\nu_1, \dots, \nu_m)$, $S(\nu_k) \subset E_k$, $k = 1, \dots, m$. Since $\nu\mathscr{A} = 0$, the relation

$$-a_{kk}\nu_k = \sum_{j=1,\, j\neq k}^{m} a_{kj}\nu_j \tag{1.37}$$

holds for all k. In the last sum we have either $a_{kj} = 0$ or $E_k \cap E_j = \varnothing$ (see (1.24)). Hence, the supports of the charges on the right- and left-hand sides of (1.37) are disjoint, and $a_{kk}\nu_k = 0$. Since $a_{kk} > 0$, we get that $\nu_k = 0$. As a result, $\lambda = \lambda'$, and so the extremal measure is unique. This proves the assertion 1) of Theorem 1.1.

2) In essence, the assertions 2a) and 2b) are just variants of the Kuhn–Tucker theorem [35]. We proceed to prove that this is indeed the case.

Let λ be the extremal measure. Suppose that $J(\lambda) = J_0$, and consider the following set in $\mathbb{R}^{r+1}$:

$$\Xi := \Bigg\{(t_0, \dots, t_r) =: \mathbf{t} \;\Big|\; \exists \mu \in \mathbf{M}^+(E)\colon J(\mu) \leqslant t_0 + J_0, \\ \sum_{j=1}^{m} c_{kj}|\mu_j| = b_k + t_k,\ k = 1, \dots, r\Bigg\}.$$

It is readily verified that this set is non-empty, convex, and has no common points with the ray

$$\Upsilon := \{(\omega_0, 0, \dots, 0)\colon \omega_0 < 0\}.$$

By the finite-dimensional separation theorem, there is a non-zero vector $(l_0, l_1, \dots, l_r) \in \mathbb{R}^{r+1}$ such that

$$\inf_{\Xi} \sum_{j=0}^{r} l_j t_j \geqslant \sup_{\omega_0 < 0} l_0 \omega_0 \geqslant 0,$$

and thus

$$\sum_{j=0}^{r} l_j t_j \geqslant 0 \quad \forall\, \mathbf{t} \in \Xi. \tag{1.38}$$

We assert that $l_0 > 0$. Substituting $(1, 0, \dots, 0) \in \Xi$ in (1.38) gives us that $l_0 \geqslant 0$. Suppose that $l_0 = 0$. Then for any $x \in \mathbb{R}^m_+$

$$\sum_{k=1}^{r} l_k \sum_{j} c_{kj} x_j \geqslant \sum_{k=1}^{r} l_k \sum_{j} c_{kj} |\lambda_j|.$$

It follows that $\sum_{k=1}^{r} l_k \sum_j c_{kj}|\lambda_j| = 0$ and $\sum_{k=1}^{r} l_k c_{kj} \geqslant 0$ for each j. But this is possible only when $l_1 = \dots = l_r = 0$, contradicting the assumption that $l \neq 0$. Hence $l_0 > 0$, and thus we can take $l_0 = 1$.

Suppose now that $\mu \in \mathbf{M}^+(\vec{E})$. Then

$$J(\mu) - J_0 + \sum_{k} l_k \Bigg(\sum_{j} c_{kj}|\mu_j| - b_k \Bigg) \geqslant 0,$$

that is, $\mathscr{L}(\mu) \geqslant \mathscr{L}(\lambda)$.

Conversely, if a measure $\lambda \in \mathbf{M}^+_{\mathscr{C},b}(\vec{E}\,)$ satisfies

$$\mathscr{L}(\mu) \geqslant \mathscr{L}(\lambda) \quad \forall\, \mu \in \mathbf{M}^+(\vec{E}\,),$$

then for every $\mu \in \mathbf{M}^+_{\mathscr{C},b}(\vec{E}\,)$ we have $J(\mu) \geqslant J(\lambda)$. This proves the assertion 2).

3) We assert that the equilibrium conditions (1.32) are equivalent to the Lagrangian minimization problem (1.31).

Suppose that for some measure $\lambda \in \mathbf{M}^+_{\mathscr{C},b}(\vec{E}\,)$

$$\mathscr{L}(\lambda) \leqslant \mathscr{L}(\mu) \quad \forall\, \mu \in \mathbf{M}^+(\vec{E}\,), \tag{1.39}$$

and consider a charge $\nu^{(k)} = (0, \dots, \nu, \dots, 0) \in \mathbf{M}^+(\vec{E})$ with only the kth component non-zero. Then

$$\begin{aligned}\mathscr{L}\big(\lambda + \varepsilon\nu^{(k)}\big) - \mathscr{L}(\lambda) &= 2\varepsilon \sum_j a_{kj} I(\lambda_j, \nu) + \varepsilon \sum_j l_j c_{jk} |\nu| + O(\varepsilon^2) \\ &= \varepsilon \int \left(2W_k^\lambda + \sum_j l_j c_{jk} \right) d\nu + O(\varepsilon^2).\end{aligned}$$

Let

$$\varkappa_k = -\frac{1}{2} \sum_j l_j c_{jk}. \tag{1.40}$$

We shall show that $W_k^\lambda \geqslant \varkappa_k$ on E_k. Since $W_k^\lambda - a_{kk} V^{\lambda_k}$ is continuous on E_k, which is a regular compact set, it suffices to show that $W_k^\lambda \geqslant \varkappa_k$ on E_k outside a subset of capacity zero. Assume on the contrary that there is an $E \subset E_k$ with $\operatorname{cap} E > 0$ on which $W_k^\lambda < \varkappa_k$, and consider a (scalar) positive measure $\nu \in \mathbf{M}^+(E_k)$. Then $\mathscr{L}\big(\lambda + \varepsilon\nu^{(k)}\big) - \mathscr{L}(\lambda) < 0$ for sufficiently small $\varepsilon > 0$, which contradicts (1.39). Assume now that $W_k^\lambda(x_0) > \varkappa_k$ at a point $x_0 \in S(\lambda_k)$. Then since W_k^λ is lower semicontinuous, there is a neighbourhood $U(x_0)$ in which $W_k^\lambda > \varkappa_k$. Also, $\lambda\big(U(x_0)\big) > 0$ since $x_0 \in S(\lambda_k)$. We now pick a negative measure ν with support in $U(x_0)$ and choose an $\varepsilon > 0$ such that $\lambda_k + \varepsilon\nu$ is a positive measure. This again contradicts (1.39).

Conversely, assume that the equilibrium relations (1.32) hold. Then the constants $l_1, \dots, l_r$ satisfying (1.40) are known. Suppose that $\mathscr{L}(\mu) < \mathscr{L}(\lambda)$ for some $\mu \in \mathbf{M}^+(\vec{E})$. Then by convexity $\mathscr{L}\big((1-\varepsilon)\lambda + \varepsilon\mu\big) < \mathscr{L}(\lambda)$, and also $(1-\varepsilon)\lambda + \varepsilon\mu \in \mathbf{M}^+(\vec{E})$. Hence, the derivative of $\mathscr{L}$ in the direction $\mu - \lambda$ at λ is negative, that is,

$$2 \sum_{j,k} a_{jk} I(\lambda_j, \mu_k - \lambda_k) + \sum_j l_j \sum_k c_{jk} (|\mu_k| - |\lambda_k|) < 0.$$

The last inequality holds if and only if

$$2 \sum_{j,k} a_{jk} I(\lambda_j, \mu_k) + \sum_j l_j \sum_k c_{jk} (|\mu_k|) < \mathscr{L}(\lambda). \tag{1.41}$$

Integrating the kth equilibrium condition with respect to λ_k and summing over k, we get that $\mathscr{L}(\lambda) = 0$. Integrating the kth equilibrium condition with respect to μ_k, we get that

$$2 \sum_j a_{jk} I(\lambda_j, \mu_k) + \sum_j l_j c_{jk} (|\mu_k|) \geqslant 0,$$

and hence, summing over k, we arrive at a contradiction to (1.41). The proof of Theorem 1.1 is complete.

1.5. Example. The triangle graph. The Hermite–Padé problem for a system of functions on arbitrary graphs gives rise to new effects which are not found in tree graphs. For example, the energy functional may attain its minimum at the boundary of the admissible set of measures, that is, some components of the extremal measure may be zero.

Following [17] (see also [36]), we consider the graph on the vertices A, B, O (O is the least (root) vertex) joined by three edges α, β, and γ (see Fig. 4). To the

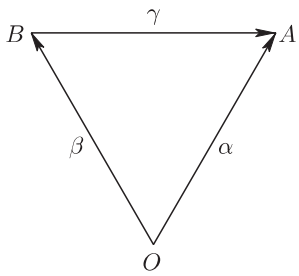

Figure 4. The triangle graph.

edges of the graph we assign non-overlapping intervals E_α, E_β, and E_γ of the real axis. On these intervals we are given positive Borel measures σ_α, σ_β, and σ_γ such that the derivative of the absolutely continuous component σ'_κ for $\kappa \in \mathscr{E} = \{\alpha, \beta, \gamma\}$ satisfies $\sigma'_\kappa > 0$ almost everywhere on E_κ with respect to Lebesgue measure.

Corresponding to the vertices A and B of the graph there is a generalized system of Nikishin functions:

$$f_A(z) = \hat{\sigma}_\alpha(z) + \int_{E_\beta} \frac{\hat{\sigma}_\gamma(x)\, d\sigma_\beta(x)}{z - x}\,, \tag{1.42}$$

$$f_B(z) = \hat{\sigma}_\beta(z), \tag{1.43}$$

where $\hat{\sigma}_\kappa(x) = \displaystyle\int_{E_\kappa} \frac{d\sigma_\kappa(t)}{x - t}\,.$

For a fixed diagonal multi-index (n, n), consider the Hermite–Padé approximants to this system. These are rational functions $(Q_{n,A}/P_n,\ Q_{n,B}/P_n)$ with a common denominator P_n of degree $\deg P_n \leqslant 2n$ such that, as $z \to \infty$,

$$R_{n,A} := f_A P_n - Q_{n,A} = O(z^{-n-1}), \tag{1.44}$$

$$R_{n,B} := f_B P_n - Q_{n,B} = O(z^{-n-1}). \tag{1.45}$$

These conditions are equivalent to the following orthogonality relations for $j = 0, \dots, n-1$:

$$\int P_n(x) x^j\, d\sigma_\alpha(x) + \int P_n(x) x^j \hat{\sigma}_\gamma(x)\, d\sigma_\beta(x) = 0, \tag{1.46}$$

$$\int P_n(x) x^j\, d\sigma_\beta(x) = 0. \tag{1.47}$$

For the function of the second kind $R_{n,B}$ this gives the relations

$$\int R_{n,B}(x) x^j\, d\sigma_\gamma(x) = -\int P_n(x) x^j \hat{\sigma}_\gamma(x)\, d\sigma_\beta(x), \qquad j = 0, \dots, n. \tag{1.48}$$

In fact, in view of the orthogonality relations (1.47) and Fubini's theorem,

$$\begin{aligned}\int_{E_\gamma} R_{n,B}(x)x^j\,d\sigma_\gamma(x) &= \int_{E_\gamma}\int_{E_\beta} \frac{P_n(t)\,d\sigma_\beta(t)}{x-t}\,x^j\,d\sigma_\gamma(x)\\ &= \int_{E_\gamma}\int_{E_\beta} P_n(t)\frac{x^j-t^j}{x-t}\,d\sigma_\beta(t)\,d\sigma_\gamma(x) + \int_{E_\gamma}\int_{E_\beta}\frac{P_n(t)t^j\,d\sigma_\beta(t)}{x-t}\,d\sigma_\gamma(x)\\ &= \int_{E_\gamma}\frac{d\sigma_\gamma(x)}{x-t}\,P_n(t)t^j\,d\sigma_\beta(t) = -\int_{E_\beta} P_n(t)t^j\hat\sigma_\gamma(t)\,d\sigma_\beta(t). \end{aligned} \tag{1.49}$$

Hence,

$$\int P_n(x)x^j\,d\sigma_\alpha(x) - \int R_{n,B}(x)x^j\,d\sigma_\gamma(x) = 0, \qquad j=0,\dots,n-1. \tag{1.50}$$

Let $p_{n,\gamma}$ be a normalized polynomial (with leading coefficient 1) formed from the zeros of $R_{n,B}$ on E_γ (if there are no zeros, then we set $p_{n,\gamma}=1$), and let $m_\gamma = \deg p_{n,\gamma}$. Then using Cauchy's integral formula, we get from (1.45) that

$$\int_{E_\beta} P_n(x)x^j\frac{d\sigma_\beta(x)}{p_{n,\gamma}(x)} = 0, \qquad j=0,\dots,m_\gamma+n-1, \tag{1.51}$$

and thus

$$\frac{R_{n,B}(z)}{p_{n,\gamma}(z)} = \int\frac{P_n(x)\,d\sigma_\beta(x)}{(z-x)p_{n,\gamma}(x)}\,. \tag{1.52}$$

Let $p_{n,\alpha}$ and $p_{n,\beta}$ be the polynomials (of degrees m_α and m_β) formed from the zeros of P_n on the intervals E_α and E_β, respectively. Then the conditions (1.50) and (1.51) imply that $m_\alpha+m_\gamma \geqslant n-1$ and $m_\beta \geqslant m_\gamma+n$. Hence $m_\alpha+m_\beta\geqslant 2n-1$. We write the orthogonality relations (1.50) as follows:

$$\int_{E_\alpha} p_{n,\alpha}(x)p_{n,\gamma}(x)x^j\frac{P_n(x)\,d\sigma_\alpha(x)}{p_{n,\alpha}(x)p_{n,\gamma}(x)} - \int_{E_\gamma} p_{n,\alpha}(x)p_{n,\gamma}(x)x^j\frac{R_{n,B}(x)\,d\sigma_\gamma(x)}{p_{n,\gamma}(x)p_{n,\alpha}(x)} = 0. \tag{1.53}$$

Let $\mu_{n,\alpha}$, $\mu_{n,\beta}$, and $\mu_{n,\gamma}$ be the zero-counting measures associated with the respective polynomials $p_{n,\alpha}$, $p_{n,\beta}$, and $p_{n,\gamma}$. We consider a subsequence $\Lambda\subset\mathbb{N}$ such that these measures have some weak limits: $\mu_{n,\alpha}/n\to\lambda_\alpha$, $\mu_{n,\beta}/n\to\lambda_\beta$, $\mu_{n,\gamma}/n\to\lambda_\gamma$. By the Gonchar–Rakhmanov theorem on the weak asymptotics of polynomials which are orthogonal with respect to a variable weight (see [37], [16]), and using the orthogonality conditions (1.53) and (1.51) and the representations (1.52), we obtain the following equilibrium relations for the potentials of the limit measures:

$$2V^{\lambda_\alpha}(x) + V^{\lambda_\beta}(x) + V^{\lambda_\gamma}(x)\begin{cases} = \varkappa_\alpha, & x\in S(\lambda_\alpha),\\ \geqslant \varkappa_\alpha, & x\in E_\alpha,\end{cases} \tag{1.54}$$

$$2V^{\lambda_\beta}(x) + V^{\lambda_\alpha}(x) - V^{\lambda_\gamma}(x)\begin{cases} = \varkappa_\beta, & x\in S(\lambda_\beta),\\ \geqslant \varkappa_\beta, & x\in E_\beta,\end{cases} \tag{1.55}$$

$$2V^{\lambda_\gamma}(x) + V^{\lambda_\alpha}(x) - V^{\lambda_\beta}(x)\begin{cases} = \varkappa_\gamma := \varkappa_\alpha-\varkappa_\beta, & x\in S(\lambda_\gamma),\\ \geqslant \varkappa_\gamma, & x\in E_\gamma.\end{cases} \tag{1.56}$$

Indeed, (1.55) follows from (1.51), (1.54) follows from (1.53) on the interval E_α, and (1.56) follows from (1.53) on the interval E_γ. Here we have taken into account the asymptotics

$$\frac{1}{n}\log\left|\frac{R_{n,B}(z)}{p_{n,\gamma}(z)}\right| \quad \rightrightarrows \quad V^{\lambda_\beta}(z) - \varkappa_\beta,$$

which hold uniformly outside of E_β. This asymptotic expression follows from (1.52). Since there is a unique vector measure $(\lambda_\alpha, \lambda_\beta, \lambda_\gamma)$ that satisfies these equilibrium relations and the conditions $|\lambda_\beta| - |\lambda_\gamma| = 1$ and $|\lambda_\alpha| + |\lambda_\gamma| = 1$, it follows that

$$\frac{1}{n}\,(\mu_{n,\alpha}, \mu_{n,\beta}, \mu_{n,\gamma}) \to (\lambda_\alpha, \lambda_\beta, \lambda_\gamma) \quad \text{as} \ \ n \to \infty.$$

Formally, two kinds of degenerate solutions of the equilibrium problem are possible: $|\lambda_\alpha| = 0$ and $|\lambda_\gamma| = 0$. In the first case, $(\lambda_\beta, \lambda_\gamma)$ is an equilibrium measure with Nikishin interaction matrix, and the inequality $V^\beta + V^\gamma \geqslant \varkappa_\beta + \varkappa_\gamma$ holds on E_α. In the second case, $(\lambda_\alpha, \lambda_\beta)$ is an equilibrium measure with Angelesco interaction matrix and $V^\beta(x) - V^\alpha(x) \leqslant \varkappa_\beta - \varkappa_\alpha$ holds for $x \in E_\gamma$.

For a Nikishin system the equilibrium potentials can be expressed in terms of the branches of the algebraic function $\Phi(z)$, where

$$\Phi_O \in \mathscr{H}(\mathbb{C} \setminus E_\beta), \qquad \Phi_O(z) = z^{-2}\left(1 + O\left(\frac{1}{z}\right)\right) \quad \text{as } z \to \infty, \tag{1.57}$$

$$\Phi_B \in \mathscr{H}\big(\mathbb{C} \setminus (E_\gamma \cup E_\beta)\big), \quad \Phi_A(z) = c_B z\left(1 + O\left(\frac{1}{z}\right)\right) \quad \text{as} \ \ z \to \infty, \tag{1.58}$$

$$\Phi_A \in \mathscr{H}(\mathbb{C} \setminus E_\gamma), \qquad \Phi_A(z) = c_A z\left(1 + O\left(\frac{1}{z}\right)\right) \quad \text{as} \ \ z \to \infty, \tag{1.59}$$

as follows:

$$V^\beta = \log|\Phi_O|, \tag{1.60}$$

$$V^\beta - V^\gamma = -\log|\Phi_B| + \varkappa_\beta, \tag{1.61}$$

$$V^\gamma = -\log|\Phi_A| + \varkappa_\beta + \varkappa_\gamma. \tag{1.62}$$

Also, $V^\beta + V^\gamma \geqslant \varkappa_\beta + \varkappa_\gamma$ if and only if $|\Phi_0| \geqslant |\Phi_2|$. However, it was shown in [32] that $|\Phi_2| > |\Phi_1| > |\Phi_0|$ for all $z \in \mathbb{C} \setminus (E_\gamma \cup E_\beta)$. Hence such a boundary solution is not realized.

We now consider the case when the measure λ_γ is identically zero. It is known that the equilibrium potentials for the Angelesco problem can be expressed in terms of the branches of the three-valued algebraic function $\Phi(z)$ as follows:

$$V^\alpha + V^\beta = \log|\Phi_O|, \tag{1.63}$$

$$V^\alpha = -\log|\Phi_A| + \varkappa_\alpha, \tag{1.64}$$

$$V^\beta = -\log|\Phi_B| + \varkappa_\beta, \tag{1.65}$$

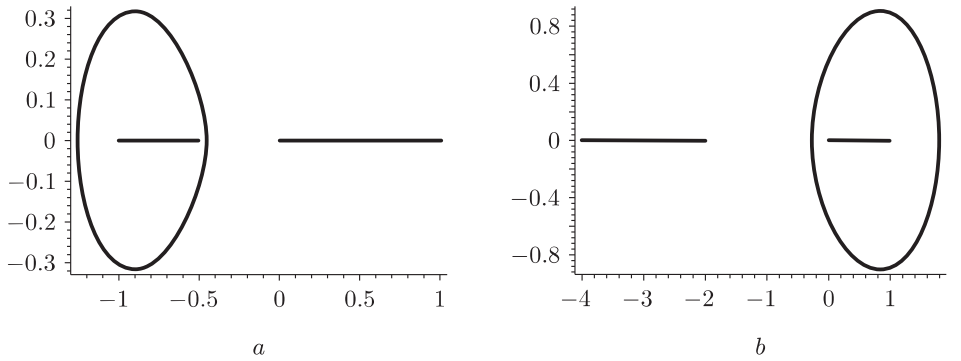

Figure 5. Critical trajectories for the Angelesco system: $E_\alpha = [0,1]$; $E_\beta = [-1,-1/2]$ (Fig. *a*), $E_\beta = [-4,-2]$ (Fig. *b*).

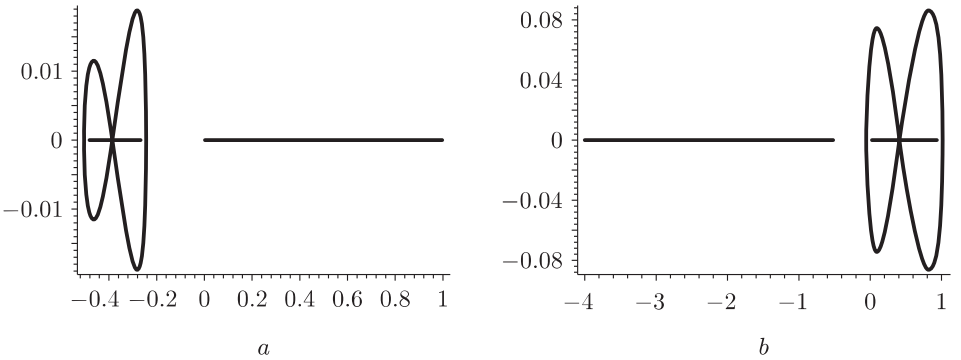

Figure 6. Critical trajectories for the Angelesco system: $E_\alpha = [0,1]$; $E_\beta = [-1/2,-1/4]$ (Fig. *a*), $E_\beta = [-4,-1/2]$ (Fig. *b*).

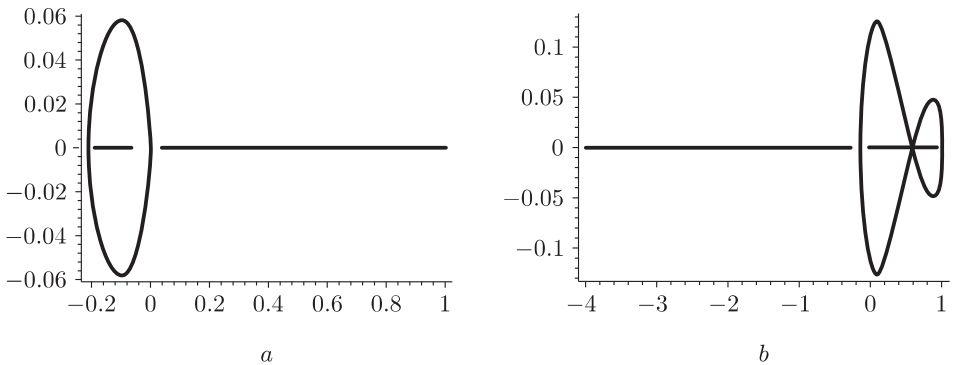

Figure 7. Critical trajectories for the Angelesco system in the case of collision: $E^*_\alpha = [0,1]$, $E_\beta = [-0.2,-0.0716]$ (Fig. *a*); $E_\alpha = [0,1]$, $E^*_\beta = [-4,-1/7]$ (Fig. *b*).

where Φ_O, Φ_A, and Φ_B are the branches of $\Phi(z)$ such that

$$\Phi_O \in \mathscr{H}\big(\mathbb{C} \setminus (E_\alpha^* \cup E_\beta^*)\big), \quad \Phi_O(z) = z^{-2}\left(1 + O\left(\frac{1}{z}\right)\right) \quad \text{as } z \to \infty, \tag{1.66}$$

$$\Phi_A \in \mathscr{H}(\mathbb{C} \setminus E_\alpha^*), \qquad \Phi_A(z) = c_A z\left(1 + O\left(\frac{1}{z}\right)\right) \quad \text{as } z \to \infty, \tag{1.67}$$

$$\Phi_B \in \mathscr{H}(\mathbb{C} \setminus E_\beta^*), \qquad \Phi_B(z) = c_B z\left(1 + O\left(\frac{1}{z}\right)\right) \quad \text{as } z \to \infty. \tag{1.68}$$

Hence, the condition $V^\beta - V^\alpha \leqslant \varkappa_\beta - \varkappa_\alpha$ is equivalent to the inequality $|\Phi_A| \leqslant |\Phi_B|$. The curves on which different pairs of branches of the function Φ are equal in absolute value are shown in Figs. 5–7. We now indicate regions in which $|\Phi_A| < |\Phi_B|$. In Figs. 5, 6*b*, and 7 the measure $|\lambda_\gamma|$ is 0 when the interval E_γ lies in the region containing the point $1 + 0$. In Fig. 6*a* the measure $|\lambda_\gamma|$ is 0 when the interval E_γ lies in the complement of the region containing the point $-0.5 - 0$.

2. Hermite–Padé approximants for functions with complex branch points

2.1. Historical remarks. The analytic theory of Hermite–Padé approximants for the *complex case* was initiated by Nuttall. In the two pioneering papers [38] and [39] from 1981 he obtained some asymptotic formulae for Hermite–Padé approximants to functions with separated complex branch points [38] (a complex analogue of an Angelesco system) and to functions meromorphic on the same Riemann surface [39] (that is, functions with the same set of branch points, like a Nikishin system for the real case). In [38] the results were verified by some heuristic considerations and numerical experiments, but [39] contains rigorous proofs of theorems. In his fundamental paper [40] from 1984, Nuttall tried to formulate a general conjecture about the asymptotic behaviour as $n \to \infty$ of the *diagonal* $\vec{n} = (n, n, \dots, n)$ Hermite–Padé polynomials. At the basis of his conjecture lies a $(p+1)$-sheeted Riemann surface which depends on the set of p functions which are being approximated. He showed how to determine this Riemann surface for some special classes of functions, but the general case was left as an open problem. Nevertheless, assuming the existence of the appropriate Riemann surface $\mathfrak{R}$, he conjectured that the strong asymptotics can be described by solutions of a certain boundary-value problem on $\mathfrak{R}$. For the main term of the asymptotics one would have

$$|P_{\vec{n}}(z)|^{1/n} \to |\Phi^{-1}(z)| = \exp\big(-\operatorname{Re} G(z)\big), \qquad n \to \infty, \tag{2.1}$$

where G is an Abelian integral of the third kind with logarithmic poles at $\infty^{(\ell)}$ $(\ell = 0, 1, \dots, p)$ with residues

$$G(z) = \begin{cases} -p \log z + \mathscr{O}(1), & z \to \infty^{(0)}, \\ \log z + \mathscr{O}(1), & z \to \infty^{(j)}, \quad j = 1, \dots, p, \end{cases} \tag{2.2}$$

and elsewhere G is analytic in the local variable. If the genus of $\mathfrak{R}$ is greater than zero, then an additional condition is imposed on G: all its periods are purely

imaginary. Such a function is unique up to an additive constant, and $\Re G(z)$ is a single-valued function on $\mathfrak{R}$ (see, for example, [41]). Another important single-valued function on $\mathfrak{R}$ is the derivative $h(z)$ of the Abelian integral $G(z)$. We single out the following branches of h at infinity:

$$h(z) := G'(z)\colon \begin{cases} h_0(z) = -\dfrac{2}{z} + \cdots, \\ h_j(z) = \dfrac{1}{z} + \cdots, \qquad j = 1, \dots, p, \end{cases} \qquad z \to \infty. \tag{2.3}$$

The function h is a rational function on its Riemann surface $\mathfrak{R}$.

We note that the condition $\operatorname{genus}(\mathfrak{R}) = 0$ implies single-valuedness of the function Φ in (2.1), therefore Φ is a rational function on $\mathfrak{R}$ that is uniquely determined (up to a multiplicative normalization) on $\mathfrak{R}$ by its divisor.

In the paper [18] there is a study of the asymptotic behaviour of the *diagonal* Hermite–Padé approximants ($\vec{n} = (n, n)$) for two functions f_1 and f_2 with branch points at the points in $A_1 = \{a_1, b_1\}$ and $A_2 = \{a_2, b_2\}$, respectively (such that $A_1 \setminus A_2 \neq \varnothing$ and $A_2 \setminus A_1 \neq \varnothing$). Here we present some of the results from [18] in order to show peculiarities and new features of the analytic properties of the Hermite–Padé approximants for functions with complex branch points.

We say that

$$f_j \in \mathscr{A}(\overline{\mathbb{C}} \setminus A_j), \qquad A_j = \{a_j, b_j\}, \tag{2.4}$$

if the Laurent expansion (0.1)) is convergent in a neighbourhood of infinity and has an analytic continuation along any path in $\mathbb{C} \setminus A_j$ and the function f_j has branch points of algebraic or logarithmic character in A_j. A typical example is the function

$$f_j(z) = \log \frac{z - a_j}{z - b_j}\,.$$

For the situation under consideration ($p = 2$ and $\vec{n} = (n, n)$) we use the notation

$$\pi_{\vec{n}} = \pi_n, \quad P_{\vec{n}} = P_n, \quad Q^{(j)}_{\vec{n}} = Q^{(j)}_n, \quad R^{(j)}_{\vec{n}} = R^{(j)}_n, \qquad j = 1, 2,$$

and we assume that $P_{\vec{n}}$ is monic.

2.2. Geometry of the problem. Asymptotic properties of Hermite–Padé approximants such as the limit distribution of their poles and the structure of the domains of convergence are strongly dependent on the positions of the branch points of the functions being approximated. In [18] various classes of arrangements of the points in

$$A := \{a_1, b_1; a_2, b_2\} \tag{2.5}$$

were considered in detail. In this survey we take one of these classes with a rather general and rich structure. The Hermite–Padé asymptotics for this class are described with the help of algebraic functions of third order with quadratic branches only at the points a_1, b_1, a_2, b_2, that is, by the Riemann–Hurwitz formula, we mean algebraic curves of third order and genus 0. To describe this class, we need to introduce some auxiliary notions.

A key role in the classification of possible geometric cases is played by the algebraic function h (see (2.3)), which satisfies the equation

$$h^3(z) - 3\,\frac{P_2(z)}{\Pi_4(z)}\,h(z) + 2\,\frac{P_1(z)}{\Pi_4(z)} = 0, \tag{2.6}$$

where $\Pi_4(z) = (z-a_1)(z-b_1)(z-a_2)(z-b_2)$ and the unknown parameters of the monic polynomials P_1 and P_2 are not yet known. The discriminant $\mathscr{D}$ of the equation (2.6) is

$$\mathscr{D} = \frac{\widetilde{\mathscr{D}}}{\Pi_4^3(z)}\,, \qquad \widetilde{\mathscr{D}} = P_2^3 - \Pi_4 P_1^2.$$

The condition that h does not have a branch point at infinity gives us a linear relation between the unknown coefficients of the polynomials P_1 and P_2. The other two algebraic relations (which are non-linear) must ensure — in our case — that h has genus zero. Hence, the discriminant $\mathscr{D}$ has zeros of even multiplicity, so that (2.5) are the only branch points of h. This gives us two algebraic equations for the unknown parameters of P_1 and P_2. In the paper [42] a rational uniformization of the algebraic curve $h(z)$ of genus 0 was obtained which made it possible to find P_1 and P_2.

2.2.1. *The function Φ and the contour Γ.* Since the algebraic function h has genus zero, the exponential of the Abelian integral

$$\mathbf{\Phi} = \exp\biggl(\int h(z)\,dz\biggr) \tag{2.7}$$

is also an algebraic function with the same Riemann surface as h.

It is clear that for a suitable normalization the function $\mathbf{\Phi}$ satisfies the equation

$$\mathbf{\Phi}^3(z) + q_1(z)\mathbf{\Phi}^2(z) + q_2(z)\mathbf{\Phi}(z) + q_0 = 0, \tag{2.8}$$

where the q_j are polynomials of degree $\leqslant j$ $(j = 0,1,2)$. In the paper [42] an explicit procedure for finding the coefficients $\{q_j\}$ is presented.

The function $\mathbf{\Phi}$ is defined in (2.7) up to a multiplicative constant. In what follows we often choose a normalization such that the product of all three branches of $\mathbf{\Phi}$ equals 1:

$$\Phi_0\Phi_1\Phi_2 = 1. \tag{2.9}$$

Therefore (see (2.3) and (2.7)), the three branches of the normalized function Φ behave at infinity as follows:

$$\begin{cases} \Phi_j(z) = \dfrac{z}{C_j} + \cdots, & j = 1,2, \\ \Phi_0(z) = \dfrac{1}{C_0 z^2} + \cdots, \end{cases} \qquad z \to \infty, \tag{2.10}$$

where

$$C_1C_2C_0 = 1.$$

From the equation for $\mathbf{\Phi}$ we can obtain a parametrization of the curve

$$\Gamma = \bigl\{z\colon |\Phi_j(z)| = |\Phi_k(z)| \text{ for some } 0 \leqslant j < k \leqslant 2\bigr\} \tag{2.11}$$

in terms of the function

$$J(\nu,z)=\nu^3+A(z)\nu^2+B(z)\nu+C(z), \tag{2.12}$$

where

$$\begin{aligned} A(z)&=\frac{3q_0-q_1(z)q_2(z)}{q_0},\\ B(z)&=\frac{q_0q_1^3(z)+q_2^3(z)-5q_0q_1(z)q_2(z)+3q_0^2}{q_0^2},\\ C(z)&=\frac{2q_0q_1^3(z)-q_1^2(z)q_2^2(z)+2q_2^3(z)-4q_0q_1(z)q_2(z)+q_0^2}{q_0^2}, \end{aligned} \tag{2.13}$$

and q_0,q_1,q_2 are the coefficients of the equation (2.8) for $\mathbf{\Phi}$. Thus, we have the following result.

Proposition 2.1 (see [18]). *The set* Γ *given in* (2.11) *for the function* $\mathbf{\Phi}$ *in* (2.7) *can be described as*

$$\Gamma=\big\{z\colon J(\nu,z)=\nu^3+A(z)\nu^2+B(z)\nu+C(z)=0 \text{ for some } \nu\in[-2,2]\big\}, \tag{2.14}$$

where A, B, *and* C *are given in* (2.13).

2.2.2. *Structure of* Γ. Here we shall use the global structure of Γ to define the geometric cases I and II. Since the polynomial $J(\nu,z)$ in (2.12) has degree 6 in the variable z, the contour Γ consists of six trajectories $z_j(\nu)$ parametrized by $\nu\in[-2,2]$. When $\nu=2$ we have

$$J(2,z)=-\frac{\Pi_4(z)\big[C_1P_1(z)+C_2\big]^2}{q_0^2}$$

with some constants C_1 and C_2. These trajectories start from the points a_1, b_1, a_2, b_2 and two trajectories start from the point

$$\alpha=P_1(0)+\frac{C_2}{C_1}.$$

Here we assume that α is different from a_j,b_j, $j=1,2$. See [18] regarding the case when α coincides with one of the points a_j,b_j, $j=1,2$. We denote these trajectories by

$$\gamma_{a_1},\ \gamma_{b_1},\ \gamma_{a_2},\ \gamma_{b_2},\ \gamma_{\alpha_1},\ \gamma_{\alpha_2}, \tag{2.15}$$

and then we extend them continuously with respect to ν to values less than 2 (see Fig. 8). When $\nu=-2$ we have

$$J(-2,z)=\frac{\big[q_0-q_1(z)q_2(z)\big]^2}{q_0^2}=\operatorname{const}(z-\beta_1)^2(z-\beta_2)^2(z-\beta_3)^2. \tag{2.16}$$

These six trajectories therefore meet pairwise at the points β_1,β_2,β_3, which are the zeros of the function $q_0-q_1q_2$ (see Fig. 9).

Now we are ready to describe those configurations of the branch points in A which will be a subject of our analysis in this survey.

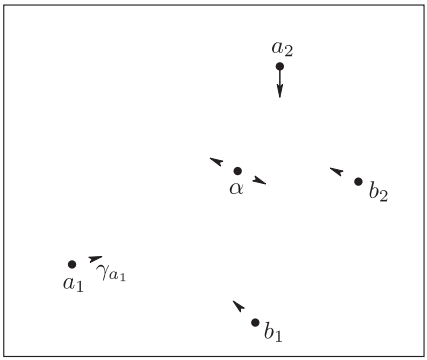

Figure 8. Start ($\nu = 2$) of the trajectories of Γ.

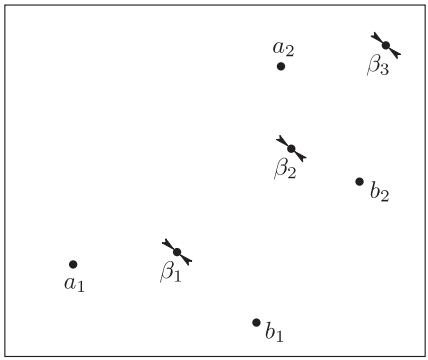

Figure 9. Finish ($\nu = -2$) of the trajectories of Γ.

Definition 2.1. We say that the set of points $\{a_1, b_1, a_2, b_2\}$ belongs to the geometric cases I or II,

$$A = \{a_1, b_1, a_2, b_2\} \in \mathrm{I} \cup \mathrm{II}, \tag{2.17}$$

if the set Γ in (2.11) has the following structure:

1) the algebraic function $z(\nu)$ defined by the equation $J(\nu, z) = 0$ (see (2.12)) has no branch points on $(-2, 2)$, so that by analytic continuation the trajectories (2.15) are defined globally on $[-2, 2]$,

$$\gamma_{a_1}(\nu),\ \gamma_{b_1}(\nu),\ \gamma_{a_2}(\nu),\ \gamma_{b_2}(\nu),\ \gamma_{\alpha_1}(\nu),\ \gamma_{\alpha_2}(\nu), \qquad \nu \in [-2, 2];$$

2) when $\nu = -2$ we have

$$\gamma_{a_1}(-2) = \gamma_{b_1}(-2), \qquad \gamma_{a_2}(-2) = \gamma_{b_2}(-2).$$

For $A \in \mathrm{I} \cup \mathrm{II}$ we can now define two arcs in $\mathbb{C}$,

$$\gamma_j = \gamma_{a_j, b_j} = \gamma_{a_j} \cup \gamma_{b_j}, \qquad j = 1, 2, \tag{2.18}$$

each connecting a_j and b_j, $j = 1, 2$, and a closed analytic curve

$$\gamma_\alpha := \gamma_{\alpha_1} \cup \gamma_{\alpha_2}.$$

Definition 2.2. Given $A \in \mathrm{I} \cup \mathrm{II}$, we say that (see Figs. 10 and 11)

1) $A \in \mathrm{I}$ if $\gamma_1 \cap \gamma_2 = \varnothing$;
2) $A \in \mathrm{II}$ if γ_1 and γ_2 have two points of intersection,

$$\gamma_1 \cap \gamma_2 = \{c_1, c_2\}.$$

In the case II we assume that the branch points a_j, b_j and the intersection points c_j are arranged as shown in Fig. 11. That is, if we follow the curve γ_j starting at a_j $(j = 1, 2)$ then we first meet c_1 and then c_2.

In what follows here we shall concentrate on the geometric case II, regarding the case I as a degeneration of it (see the details in [18]).

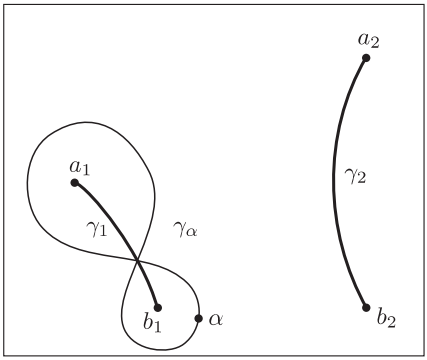

Figure 10. Global trajectories, case I.

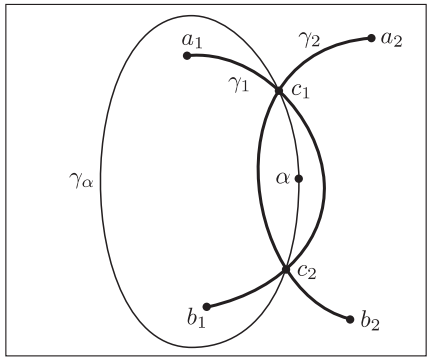

Figure 11. Global trajectories, case II.

2.2.3. *Riemann surface for the case* II. *Definition of the global branches for the algebraic functions* h *and* Φ. In the case II we have $\gamma_1 \cap \gamma_2 = \{c_1, c_2\}$. We assume that the points a_j, b_j, and c_j are arranged as shown in Fig. 11. We use γ_{a_j,c_1} to denote the part of the arc γ_j between a_j and c_1, and similarly for γ_{b_j,c_2}. Then in the case II we let

$$\widetilde{\Delta}_j := \gamma_{a_j,c_1} \cup \gamma_{b_j,c_2}, \quad j = 1, 2, \qquad \begin{cases} E_1 := \gamma_2 \setminus \widetilde{\Delta}_2, \\ E_2 := \gamma_1 \setminus \widetilde{\Delta}_1. \end{cases} \tag{2.19}$$

The arcs E_1 and E_2 form the boundary of a lens-shaped domain G:

$$\partial G := E_1 \cup E_2. \tag{2.20}$$

Note that the real-analytic curve γ_α has to pass through the points c_1, c_2 and that it divides the domain G into two parts (otherwise it would contradict the maximum principle for harmonic functions). Let

$$\Delta_{1,2} := \gamma_\alpha \cap G. \tag{2.21}$$

Finally, we define

$$\Delta_j := \widetilde{\Delta}_j \cup \Delta_{1,2}, \qquad \Delta_0 := \Delta_1 \cup \Delta_2 = \widetilde{\Delta}_1 \cup \Delta_{1,2} \cup \widetilde{\Delta}_2. \tag{2.22}$$

Definition 2.3. For a set of points $A \in$ II the corresponding Riemann surface

$$\mathfrak{R}(A) := \overline{\mathfrak{R}_0 \cup \mathfrak{R}_1 \cup \mathfrak{R}_2}$$

is formed by gluing the sheets

$$\begin{cases} \mathfrak{R}_1 := \overline{\mathbb{C}} \setminus (\Delta_1 \cup E_1), \\ \mathfrak{R}_2 := \overline{\mathbb{C}} \setminus (\widetilde{\Delta}_2 \cup E_1) = \overline{\mathbb{C}} \setminus \gamma_2 \end{cases} \tag{2.23}$$

of the cut complex plane to the sheet

$$\mathfrak{R}_0 := \overline{\mathbb{C}} \setminus \Delta_0 = \overline{\mathbb{C}} \setminus (\Delta_1 \cup \widetilde{\Delta}_2),$$

and then gluing the sheets $\mathfrak{R}_1$ and $\mathfrak{R}_2$ together along E_1 (see Fig. 12).

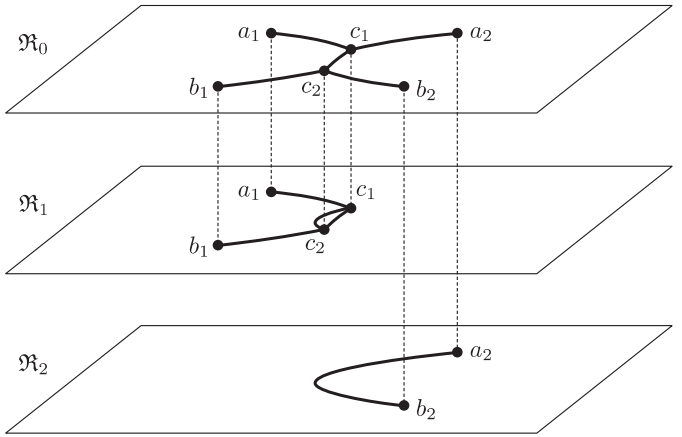

Figure 12. Riemann surface for case II.

Remark 2.1. 1. The so-defined structure of sheets of a Riemann surface possesses a certain non-symmetry with respect to the pairs $\{a_1, b_1\}$ and $\{a_2, b_2\}$. We can also consider a dual structure of sheets for $\mathfrak{R}$:

$$\mathfrak{R}_0 := \overline{\mathbb{C}} \setminus \Delta_0, \qquad \mathfrak{R}_2 := \overline{\mathbb{C}} \setminus (\Delta_2 \cup E_2), \qquad \mathfrak{R}_1 := \overline{\mathbb{C}} \setminus \gamma_1.$$

2. Although all three sheets are glued together at the points c_1 and c_2, it can easily be checked that these are not branch points of $\mathfrak{R}$.

3. Note that the sheet $\mathfrak{R}_1$ is a disconnected set (see (2.23)). It consists of two components: the domain G_1 bounded by E_1 and $\Delta_{1,2}$,

$$\partial G_1 := E_1 \cup \Delta_{1,2},$$

and the domain $\overline{\mathbb{C}} \setminus (\overline{G_1} \cup \widetilde{\Delta}_1)$.

The structure of the sheets (2.23) gives the global branches of the functions h and Φ:

$$\begin{aligned} &h_0, \Phi_0 \in H(\overline{\mathbb{C}} \setminus \Delta_0), \\ &h_1, \Phi_1 \in H\big(\mathbb{C} \setminus (\widetilde{\Delta}_1 \cup \overline{G_1})\big) \cup H(G_1), \\ &h_2, \Phi_2 \in H\big(\mathbb{C} \setminus (\widetilde{\Delta}_2 \cup E_1)\big). \end{aligned} \tag{2.24}$$

More precisely, in the domains

$$\overline{\mathbb{C}} \setminus \Delta_0, \quad \mathbb{C} \setminus (\widetilde{\Delta}_1 \cup \overline{G_1}), \quad \mathbb{C} \setminus (\widetilde{\Delta}_2 \cup E_1)$$

the branches

$$(h_0, \Phi_0), \quad (h_1, \Phi_1), \quad (h_2, \Phi_2)$$

are respectively the result of analytic continuation of (2.3) and (2.10) from the point at infinity, and the branch (h_1, Φ_1) in G_1 is the result of analytic continuation of

(h_0, Φ_0) across $\Delta_{1,2}$ or of (h_2, Φ_2) across E_1. The three branches Φ_j, $j = 0, 1, 2$, also determine a number of regions in the complex plane:

$$\Omega_{j,k,\ell} = \big\{z \in \mathbb{C}\colon |\Phi_j(z)| < |\Phi_k(z)| < |\Phi_\ell(z)|\big\}, \qquad j, k, \ell = 0, 1, 2. \tag{2.25}$$

We also define the sets

$$\Omega_{j,k} = \big\{z \in \mathbb{C}\colon |\Phi_j(z)| < |\Phi_k(z)|\big\}, \qquad j, k = 0, 1, 2. \tag{2.26}$$

Using the continuity of the global branches of Φ along γ_1, γ_2, and γ_α and using the maximum principle, we obtain a partition of $\mathbb{C}$ by Γ into the domains $\Omega_{j,k,\ell}$, as shown in Fig. 13.

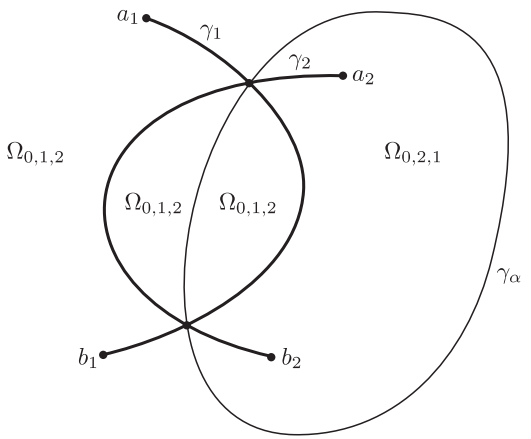

Figure 13. Partition of $\mathbb{C}$ by Γ into regions $\Omega_{j,k,\ell}$. The case II.

Proposition 2.2. *If $A \in$ II then the contour Γ defined in (2.11), (2.12) has the following structure*:

$$\begin{aligned}
\Gamma &= \Gamma_{0,1} \cup \Gamma_{0,2} \cup \Gamma_{1,2},\\
\Gamma_{0,1} &:= \{z\colon |\Phi_0(z)| = |\Phi_1(z)|\} = \Delta_1,\\
\Gamma_{0,2} &:= \{z\colon |\Phi_0(z)| = |\Phi_2(z)|\} = \widetilde{\Delta}_2,\\
\Gamma_{1,2} &:= \{z\colon |\Phi_1(z)| = |\Phi_2(z)|\} = E_1 \cup E_2 \cup (\gamma_\alpha \setminus \Delta_{1,2}),
\end{aligned}$$

and for the domains $\Omega_{j,k,\ell}$ (see Fig. 13),

$$\begin{cases}
\partial\Omega_{0,1,2} = \gamma_\alpha \cup \widetilde{\Delta}_1 \cup E_1 \cup E_2,\\
\partial\Omega_{0,2,1} = (\gamma_\alpha \setminus \Delta_{1,2}) \cup \widetilde{\Delta}_2 \cup E_2.
\end{cases} \tag{2.27}$$

See the proof of Proposition 2.2 and the proof of the following Theorem 2.1 in [18].

Theorem 2.1. *Let $A \in$ II.*

1) *The jump of h_0 on Δ_0 produces a positive measure λ of total mass 2,*

$$\frac{1}{2\pi i}\left(h_{0+}(\xi) - h_{0-}(\xi)\right) d\xi =: d\lambda(\xi), \qquad \xi \in \Delta_0,$$

and λ has two component measures λ_1 and $\widetilde{\lambda}_2$ supported on Δ_1 and $\widetilde{\Delta}_2$,

$$\lambda = \begin{cases} \lambda_1 & \text{on } \Delta_1, \\ \widetilde{\lambda}_2 & \text{on } \widetilde{\Delta}_2, \end{cases}$$

and with densities

$$\lambda_1'(\xi) = \frac{m_1(\xi)}{\sqrt{(\xi - a_1)(\xi - b_1)}}, \qquad \widetilde{\lambda}_2'(\xi) = \frac{m_2(\xi)}{\sqrt{(\xi - a_2)(\xi - b_2)}},$$
$$m_1 \in H(\widetilde{\Delta}_1) \cap H(\Delta_{1,2}), \qquad m_2 \in H(\widetilde{\Delta}_2).$$

2) *The jump of h_1 on E_1 produces a positive measure μ_1,*

$$\frac{1}{2\pi i}\left(h_{1+}(\xi) - h_{1-}(\xi)\right) d\xi =: d\mu_1(\xi), \qquad \xi \in E_1,$$

and $\mu_1' \in H(E_1)$.

3) *The total masses of these measures are connected by the relations*

$$|\lambda_1| + |\widetilde{\lambda}_2| = 2, \quad |\lambda_1| - |\mu_1| = 1.$$

Thus, the Riemann surface for the case II produces a system of three positive measures

$$\{\lambda_1, \widetilde{\lambda}_2, \mu_1\}, \qquad \begin{cases} \operatorname{supp}\lambda_1 = \Delta_1, \\ \operatorname{supp}\widetilde{\lambda}_2 = \widetilde{\Delta}_2, \\ \operatorname{supp}\mu_1 = E_1, \end{cases} \qquad \begin{cases} |\lambda_1| + |\widetilde{\lambda}_2| = 2, \\ |\lambda_1| - |\mu_1| = 1. \end{cases}$$

If we consider the dual Riemann surface (see Remark 2.1, item 1), then we arrive at a dual system of three positive measures

$$\{\lambda_2, \widetilde{\lambda}_1, \mu_2\}, \qquad \begin{cases} \operatorname{supp}\lambda_2 = \Delta_2, \\ \operatorname{supp}\widetilde{\lambda}_1 = \widetilde{\Delta}_1, \\ \operatorname{supp}\mu_2 = E_2, \end{cases} \qquad \begin{cases} |\lambda_2| + |\widetilde{\lambda}_1| = 2, \\ |\lambda_2| - |\mu_2| = 1, \end{cases} \tag{2.28}$$

and we have

$$\lambda_1 + \widetilde{\lambda}_2 = \lambda_2 + \widetilde{\lambda}_1 = \lambda.$$

2.3. Weak asymptotics, convergence, vector potential equilibrium problems. In this subsection we formulate corollaries from the strong asymptotics of the Hermite–Padé polynomials, obtained in [18] and relating to the weak asymptotics of the poles of the Hermite–Padé approximants and their convergence.

2.3.1. *Weak convergence.* We start with the weak limit of the counting measures ν_{P_n} that assign equal mass $1/(2n)$ to the poles of the Hermite–Padé approximants. Again we concentrate on the geometric case II. The results relating to the case I can be regarded as a degeneration of the stated results. The details and other geometric configurations can be found in [18].

Theorem 2.2. *Suppose that A corresponds to the case II. Then the poles of the Hermite–Padé approximants (0.2)–(0.4) for the functions in (2.4) have a weak limit*

$$\nu_{P_n} \xrightarrow{*} \frac{\lambda}{2}, \qquad n \to \infty, \tag{2.29}$$

where the limit measure λ is defined in Theorem 2.1.

We now state a result about the finite zeros of the functions of the second kind

$$R_n^{(j)} := f_j P_n - Q_n^{(j)}, \qquad j = 1, 2,$$

in (0.4). These zeros represent extra interpolation points. We use the notation $\nu_{R_n^{(j)}}$ for the counting measures with equal mass $1/n$ at the finite zeros of $R_n^{(j)}$.

Theorem 2.3. *Consider $R_n^{(j)}$ for the functions (2.4). For $A \in$ II*

$$\nu_{R_n^{(j)}} \xrightarrow{*} \mu_j, \qquad j = 1, 2, \tag{2.30}$$

where the limit measures μ_1 and μ_2 are defined in Theorem 2.1 and in (2.28).

The next theorem describes the nth-root asymptotics of the error term

$$f_j - \pi_n^{(j)} = \frac{R_n^{(j)}}{P_n}, \qquad \pi_n^{(j)} = \frac{Q_n^{(j)}}{P_n}, \qquad j = 1, 2,$$

and the convergence of the Hermite–Padé approximants (0.2)–(0.4).

Theorem 2.4. *Consider the Hermite–Padé approximants for the functions (2.4). If $A \in$ II, then (uniformly on compact subsets of the indicated sets)*

$$|f_1 - \pi_n^{(1)}|^{1/n} \to \left|\frac{\Phi_0}{\Phi_1}\right| \text{ on } \overline{\mathbb{C}} \setminus (\Delta_1 \cup E_1), \quad |f_2 - \pi_n^{(2)}|^{1/n} \to \begin{cases} \left|\dfrac{\Phi_0}{\Phi_2}\right| & \text{on } (\overline{\mathbb{C}} \setminus \Delta_2) \setminus G, \\ \left|\dfrac{\Phi_0}{\Phi_1}\right| & \text{on } G \setminus \Delta_{1,2}, \end{cases}$$

and therefore

$$\pi_n^{(j)} \to f_j \text{ on } \overline{\mathbb{C}} \setminus \Delta_j, \qquad j = 1, 2.$$

We recall that the definition of the branches of the algebraic function $\mathbf{\Phi}$ and the domains of divergence $\Omega_{j,0}$, $j = 1, 2$, of the Hermite–Padé approximants are given in (2.26), (2.7), and (2.24).

2.3.2. *Vector equilibrium problem.* In concluding this section we state a universal **vector equilibrium problem** for the logarithmic potentials of the measures $\lambda = \lambda_1 + \lambda_2$, μ_1, μ_2. These measures were introduced in Theorem 2.1 for the description of the weak limits of the poles and of the extra interpolation points.

Theorem 2.5. *Let f_1 and f_2 be the functions in (2.4) and let $A \in \mathrm{I} \cup \mathrm{II}$. Then the following statements hold.*

I. a) *There exist piecewise analytic arcs Δ_1 and Δ_2 off which f_1 and f_2 are holomorphic,*

$$f_j \in H(\overline{\mathbb{C}} \setminus \Delta_j), \qquad j = 1, 2,$$

and a piecewise analytic contour E bounding a domain containing their intersection

$$\Delta_{1,2} := \Delta_1 \cap \Delta_2 \qquad (\Delta_{1,2} = \varnothing \Rightarrow E = \varnothing).$$

b) *There exists a triple of measures $(\lambda_1, \widetilde{\lambda}_2, \mu_1)$ with supports (here $S(\mu) := \operatorname{supp} \mu$)*

$$S(\lambda_1) \subset \Delta_1, \qquad S(\widetilde{\lambda}_2) \subset \widetilde{\Delta}_2 := \Delta_2 \setminus \Delta_{1,2}, \qquad S(\mu_1) \subset E$$

and with total masses satisfying

$$\begin{cases} |\lambda_1| + |\widetilde{\lambda}_2| = 2, \\ |\lambda_1| - |\mu_1| = 1. \end{cases}$$

c) *The triple of measures $(\lambda_1, \widetilde{\lambda}_2, \mu_1)$ satisfies the following equilibrium relations with some constants κ_1 and $\widetilde{\kappa}_2$:*

$$\begin{aligned} U_1 &:= 2V^{\lambda_1} + V^{\widetilde{\lambda}_2} - V^{\mu_1} \begin{cases} = \kappa_1 & \text{on } S(\lambda_1), \\ \geqslant \kappa_1 & \text{on } \Delta_1, \end{cases} \\ U_2 &:= V^{\lambda_1} + 2V^{\widetilde{\lambda}_2} + V^{\mu_1} \begin{cases} = \widetilde{\kappa}_2 & \text{on } S(\widetilde{\lambda}_2), \\ \geqslant \widetilde{\kappa}_2 & \text{on } \widetilde{\Delta}_2, \end{cases} \\ U_3 &:= -V^{\lambda_1} + V^{\widetilde{\lambda}_2} + 2V^{\mu_1} \begin{cases} = \widetilde{\kappa}_2 - \kappa_1 & \text{on } S(\mu_1), \\ \geqslant \widetilde{\kappa}_2 - \kappa_1 & \text{on } E. \end{cases} \end{aligned} \tag{2.31}$$

d) *The supports of the measures $(\lambda_1, \widetilde{\lambda}_2, \mu_1)$ satisfy the following symmetry relations*:

$$\begin{cases} \dfrac{\partial U_1}{\partial n_+} = \dfrac{\partial U_1}{\partial n_-} & \text{on } S(\lambda_1), \\ \dfrac{\partial U_2}{\partial n_+} = \dfrac{\partial U_2}{\partial n_-} & \text{on } S(\widetilde{\lambda}_2), \\ \dfrac{\partial U_3}{\partial n_+} = \dfrac{\partial U_3}{\partial n_-} & \text{on } S(\mu_1), \end{cases} \tag{2.32}$$

where $\partial/\partial n_\pm$ denote the normal derivatives on the corresponding contours.

II. *There is also a dual problem for the triple $(\lambda_2, \widetilde{\lambda}_1, \mu_2)$ which is obtained from the problem* **I**.a)–**I**.d) *by interchanging the indices* 1 *and* 2.

III. *The equilibrium measures* $(\lambda_1, \widetilde{\lambda}_2, \mu_1)$ *and* $(\lambda_2, \widetilde{\lambda}_1, \mu_2)$ *are related as follows*:

$$\lambda := \lambda_1 + \widetilde{\lambda}_2 = \lambda_2 + \widetilde{\lambda}_1, \qquad S(\lambda) \subset \Delta_0 := \Delta_1 \cup \Delta_2,$$
$$\mu := \mu_1 + \mu_2, \qquad S(\mu_1) \cup S(\mu_2) = E,$$
$$V^{\mu}|_E = V^{\lambda_{1,2}}|_E, \qquad \lambda_{1,2} := \lambda|_{\Delta_{1,2}}.$$

IV. *The measure* $\lambda/2$ *is the weak limit* (2.29) *of the poles of the Hermite–Padé approximants of the functions* f_1 *and* f_2, *and the measures* μ_1 *and* μ_2 *are the weak limits* (2.30) *of the extra interpolation points.*

The proof of Theorem 2.5 (including generalizations for other geometric configurations) is given in [18].

In the case I we have $\Delta_{1,2} = \varnothing$ and therefore $E = \varnothing$,

$$\Delta_j = \widetilde{\Delta}_j, \qquad \lambda_j = \widetilde{\lambda}_j, \quad \mu_j = 0, \quad S(\lambda_j) = \Delta_j, \qquad j = 1, 2,$$

and in the systems of equilibrium and symmetry relations (2.31)–(2.32) only the first two relations are needed, which are the equilibrium relations for an Angelesco system (1.6) in the complex plane. In the case II the contours for the equilibrium problems are defined in (2.19)–(2.22). In this case (see Theorem 2.1) there are no degeneracies of the components in the equilibrium problem (2.31) and the dual problem.

The equilibrium relations (1.54) coincide with the equilibrium relations (1.54)–(1.56) from the previous section. This shows their universal character.

Chapter 2

Multiple orthogonal polynomial ensembles

3. Definitions, determinantal formulae

3.1. Orthogonal polynomial ensembles. The polynomials characterized by the multiple orthogonality conditions (1.5) appear in a natural way in certain models of random matrices and non-intersecting paths. This was first observed in [9] for the Hermitian random matrix model with an external source.

The connection between orthogonal polynomials and random matrix theory is much older. It was developed by Mehta and Gaudin first for Gaussian ensembles [43] and then extended to more general invariant ensembles (see [44]–[51]). Here the setup is the following. On the space of $n \times n$ Hermitian matrices M one defines a probability measure of the form

$$\frac{1}{\widetilde{Z}_n} e^{-\operatorname{Tr} V(M)}\, dM, \tag{3.1}$$

where $V\colon \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is a function such that

$$\lim_{x \to \pm\infty} \frac{V(x)}{\log(1 + |x|)} = +\infty,$$

for example, a polynomial of even degree with positive leading coefficient,

$$dM = \prod_{i=1}^{n} dM_{ii} \prod_{i=1}^{n-1} \prod_{j=2}^{n} d\operatorname{Re} M_{ij}\ d\operatorname{Im} M_{ij}$$

is the Lebesgue measure on the set of algebraically independent entries of M, and

$$\widetilde{Z}_n = \int e^{-\operatorname{Tr} V(M)}\, dM$$

is a normalization constant making (3.1) a probability measure. The joint probability density for the eigenvalues $x_1, \dots, x_n$ of the matrix M taken randomly according to (3.1) is equal to

$$\mathscr{P}(x_1, \dots, x_n) = \frac{1}{Z_n} \prod_{i<j} (x_j - x_i)^2 \prod_{j=1}^{n} e^{-V(x_j)}, \tag{3.2}$$

where Z_n is another normalization constant (see [44], [45]).

Let $(p_k)_{k=0}^{\infty}$ be the sequence of orthogonal polynomials with weight $e^{-V(x)}$ on the real line and let

$$K_n(x, y) = \sqrt{e^{-V(x)} e^{-V(y)}} \sum_{k=0}^{n-1} p_k(x) p_k(y) \tag{3.3}$$

be the reproducing kernel (also known as the orthogonal polynomial kernel, or the Christoffel–Darboux kernel). Then (3.2) can be rewritten as a determinant

$$\mathscr{P}(x_1, \dots, x_n) = \frac{1}{n!} \det\bigl[K_n(x_i, x_j)\bigr]_{i,j=1}^{n} \tag{3.4}$$

and for each $k = 1, \dots, n$ the k-point correlation function

$$R(x_1, \dots, x_k) = \frac{n!}{(n-k)!} \int_{\mathbb{R}^{n-k}} \mathscr{P}(x_1, \dots, x_n)\, dx_{k+1} \cdots dx_n \tag{3.5}$$

(which is, up to a constant factor, also equal to the marginal density) is the $k \times k$ determinant

$$R(x_1, \dots, x_k) = \det\bigl[K_n(x_i, x_j)\bigr]_{i,j=1,\dots,k}, \tag{3.6}$$

with the same kernel (3.3). The formula (3.6) is characteristic of a determinantal point process with correlation kernel K_n.

Another connection of the unitary ensemble (3.1) with orthogonal polynomials is that the average characteristic polynomial

$$\begin{aligned} P_n(x) &= \mathsf{E}[\det(xI - M)] \\ &= \frac{1}{Z_n} \int_{\mathbb{R}^n} \prod_{j=1} (x - x_j) \mathscr{P}(x_1, \dots, x_n)\, dx_1 \cdots dx_n \end{aligned} \tag{3.7}$$

is equal to the monic orthogonal polynomial of degree n:

$$\int_{-\infty}^{\infty} P_n(x) x^k e^{-nV(x)}\,dx = 0, \qquad k = 0, 1, \dots, n-1,$$

so that $p_n(x) = \gamma_n P_n(x)$ with γ_n the leading coefficient of the orthonormal polynomial p_n. See [52], [49], [53], [54] for more information on unitary ensembles and other models from random matrix theory.

3.2. Multiple orthogonal polynomial ensembles. A multiple orthogonal polynomial (MOP) ensemble is a generalization of the probability density function (3.2). We note that (3.2) is a product of two determinants. Indeed, by the familiar formula for the Vandermonde determinant we have

$$\prod_{i<j}(x_j - x_i) = \det\left[x_j^{i-1}\right]_{i,j=1,\dots,n}.$$

Putting $w(x) = e^{-V(x)}$, we can then write the OP ensemble (3.2) as

$$\frac{1}{Z_n} \det\left[x_j^{i-1}\right]_{i,j=1,\dots,n} \det\left[x_j^{i-1} w(x_j)\right]_{i,j=1,\dots,n}.$$

Definition 3.1 [55]. A multiple orthogonal polynomial ensemble is a probability density function on $\mathbb{R}^n$ of the form

$$\mathscr{P}(x_1, \dots, x_n) = \frac{1}{Z_n} \det\left[x_j^{i-1}\right]_{i,j=1,\dots,n} \det\left[\varphi_i(x_j)\right]_{i,j=1,\dots,n}, \tag{3.8}$$

for certain functions $\varphi_1, \dots, \varphi_n \colon \mathbb{R} \to \mathbb{R}$ whose linear span is equal to

$$\operatorname{span}\left\{x^k w_j(x) \mid k = 0, \dots, n_j - 1,\ j = 1, \dots, p\right\}, \tag{3.9}$$

for p functions $w_1, \dots, w_p \colon \mathbb{R} \to \mathbb{R}$ and positive integers $n_1, \dots, n_p$ with $n = n_1 + \cdots + n_p$. We say that the MOP ensemble (3.8) is generated by the weight functions $w_1, \dots, w_p$ and the multi-index $\vec{n} = (n_1, \dots, n_p)$.

The weight functions $w_1, \dots, w_p$ and multi-index $\vec{n} = (n_1, \dots, n_p)$ generate a MOP ensemble if and only if

$$\det\left[x_j^{i-1}\right]_{i,j=1,\dots,n} \det\left[\varphi_i(x_j)\right]_{i,j=1,\dots,n} \geqslant 0 \quad \text{for every choice of } (x_1, \dots, x_n) \in \mathbb{R}^n,$$

or

$$\det\left[x_j^{i-1}\right]_{i,j=1,\dots,n} \det\left[\varphi_i(x_j)\right]_{i,j=1,\dots,n} \leqslant 0 \quad \text{for every choice of } (x_1, \dots, x_n) \in \mathbb{R}^n,$$

and in addition

$$Z_n = \int_{\mathbb{R}^n} \det\left[x_j^{i-1}\right]_{i,j=1,\dots,n} \det\left[\varphi_i(x_j)\right]_{i,j=1,\dots,n} dx_1 \cdots dx_n \in \mathbb{R} \setminus \{0\}. \tag{3.10}$$

By the continuum version of the Cauchy–Binet theorem (see, for example, [56]), we get from (3.9) that

$$Z_n = n! \det\left[\int_{\mathbb{R}^n} x^{i-1} \varphi_j(x)\,dx\right]_{i,j=1,\dots,n}. \tag{3.11}$$

By using elementary column operations on the determinant we can replace the functions φ_j in (3.11) by the functions in (3.9). Then the expression (3.11) for Z_n turns into a block Hankel determinant

$$Z_n = c_n \det \begin{bmatrix} H_1 & \cdots & H_p \end{bmatrix}$$

with p rectangular blocks, where

$$H_j = \left[\int_{-\infty}^{\infty} x^{i+k-2} w_j(x)\, dx \right]_{i=1,\ldots,n,\ k=1,\ldots,n_j}$$

is of size $n \times n_j$ and contains moments of the weight w_j.

Any probability density function of the form (3.8) can be regarded as a multiple orthogonal polynomial ensemble, simply by taking $p = n$, $w_j = \varphi_j$, and $n_j = 1$ for $j = 1, \ldots, n$. However, the main interest in this situation is when p is small compared to n.

3.3. Determinantal point process. A multiple orthogonal polynomial ensemble is a determinantal point process. That is, there is a kernel K_n such that (3.8) can be written as

$$\mathscr{P}(x_1, \ldots, x_n) = \frac{1}{n!} \det\big[K_n(x_i, x_j)\big]_{i,j=1,\ldots,n}, \tag{3.12}$$

and for every $k = 1, \ldots, n$,

$$\begin{aligned} R(x_1, \ldots, x_k) &:= \frac{n!}{(n-k)!} \int_{\mathbb{R}^{n-k}} \mathscr{P}(x_1, \ldots, x_n)\, dx_{k+1} \cdots dx_n \\ &= \det\big[K_n(x_i, x_j)\big]_{i,j=1,\ldots,k}. \end{aligned} \tag{3.13}$$

Thus, the formulae (3.4)–(3.6) continue to hold for the MOP ensemble (3.8). The kernel can be written as a double sum

$$K_n(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} \big[A_n^{-1}\big]_{j,i} x^{i-1} \varphi_j(y), \tag{3.14}$$

where $\big[A_n^{-1}\big]_{j,i}$ denotes the (j,i)th entry of the inverse of the matrix

$$A_n = \left[\int x^{i-1} \varphi_j(x)\, dx \right]_{i,j=1,\ldots,n}$$

(see also [57], [58]). Note that $Z_n = n! \det A_n$ by (3.11). Since $Z_n \neq 0$, the matrix A_n is invertible, and the kernel (3.14) is well defined.

The formulae (3.12)–(3.14) are special cases of more general formulae for biorthogonal ensembles [57], where instead of (3.8) one has a probability density function that is a product of two general determinants

$$\frac{1}{Z_n} \det\big[f_i(x_j)\big]_{i,j=1,\ldots,n} \det\big[g_i(x_j)\big]_{i,j=1,\ldots,n}. \tag{3.15}$$

We will not go into this more general situation here.

In the orthogonal polynomial case we can rewrite the kernel (3.3) by means of the Christoffel–Darboux formula in terms of the orthogonal polynomials of degrees n and $n-1$ only:

$$K_n(x,y) = \sqrt{e^{-V(x)}e^{-V(y)}}\,\gamma_{n-1}^2\,\frac{P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y)}{x-y}\,.$$

This is very useful in the computation of asymptotics. Indeed, in the large-n limit one would need only the orthogonal polynomials of large degree in order to find the behaviour of K_n as $n \to \infty$.

There is also a Christoffel–Darboux formula for the kernel (3.14) of a MOP ensemble. It was found first in [9] for the case $p = 2$ and later in [20] for $p \geqslant 2$. To describe this we recall that the multiple orthogonal polynomial $P_{\vec{n}}$ (see (1.5)) is a monic polynomial of degree $n = |\vec{n}\,| = n_1 + \cdots + n_p$ and is characterized by the conditions

$$\int_{-\infty}^{\infty} P_{\vec{n}}(x)x^k w_j(x)\,dx = 0, \qquad k = 0,\dots,n_j-1, \quad j = 1,\dots,p. \tag{3.16}$$

In the context of a MOP ensemble (3.8), $P_{\vec{n}}$ is uniquely determined, and can be expressed as the 'average characteristic polynomial'

$$P_{\vec{n}}(z) = \mathsf{E}\Big[\prod_{j=1}^{n}(z-x_j)\Big] \tag{3.17}$$

with respect to the ensemble (3.8). Thus,

$$P_{\vec{n}}(z) = \frac{1}{Z_n}\int_{\mathbb{R}_n}\prod_{j=1}^{n}(z-x_j)\prod_{i<j}(x_j-x_i)\det\big[\varphi_i(x_j)\big]_{i,j=1,\dots,n}\,dx_1\cdots dx_n, \tag{3.18}$$

which is an n-fold integral representation of $P_{\vec{n}}$.

The polynomials $P_{\vec{n}}$ are also called multiple orthogonal polynomials of type II. There is a dual notion of multiple orthogonal polynomials of type I. The MOPs of type I are polynomials $A_{\vec{n},j}$ for $j = 1,\dots,p$, where

$$\deg A_{\vec{n},j} \leqslant n_j - 1, \tag{3.19}$$

such that the function (linear form)

$$Q_{\vec{n}}(x) = \sum_{j=1}^{p} A_{\vec{n},j}(x)w_j(x) \tag{3.20}$$

satisfies the orthogonality conditions

$$\int x^k Q_{\vec{n}}(x)\,dx = \begin{cases} 0 & \text{for } k = 0,\dots,|\vec{n}\,|-2,\\ 1 & \text{for } k = |\vec{n}\,|-1. \end{cases} \tag{3.21}$$

Again, in the situation of a MOP ensemble (3.8) the MOPs of type I and the form (3.20) are uniquely determined. In addition, $Q_{\vec{n}}$ satisfies the equality

$$\int_{-\infty}^{\infty}\frac{Q_{\vec{n}}(x)}{z-x}\,dx = \mathsf{E}\Big[\prod_{j=1}^{n}(z-x_j)^{-1}\Big], \qquad z \in \mathbb{C}\setminus\mathbb{R},$$

which means that the Cauchy transform of $Q_{\vec{n}}$ is the average of the reciprocal of the characteristic polynomial of a random point set $x_1, \dots, x_n$ from the ensemble (3.8) (see also [59] and [15]).

The polynomial $P_{\vec{n}}$ and the linear forms $Q_{\vec{n}}$ appear in the Christoffel–Darboux formula for the correlation kernel (3.14), together with the analogous objects with neighbouring multi-indices $\vec{n} \pm \vec{e}_j$ $(j = 1, \dots, p)$, where $\vec{e}_j = (\delta_{i,j})_{i=1,\dots,p}$ is the p-dimensional vector with 1 at the jth position and 0 elsewhere. We assume that all these multi-indices are normal, so that the polynomials and linear forms exist. We also write

$$h_{\vec{n},j} = \int P_{\vec{n}}(x) x^{n_j} w_j(x)\, dx, \qquad j = 1, \dots, p, \tag{3.22}$$

which will be non-zero under our assumptions.

Theorem 3.1 [20], [60]. *Assume that the multi-indices $\vec{n}$ and $\vec{n} \pm \vec{e}_j$ are normal. Then the correlation kernel (3.14) of a MOP ensemble generated by p weights $w_1, \dots, w_p$ and multi-index $\vec{n} = (n_1, \dots, n_p)$ can be written as*

$$(x-y)K_n(x,y) = P_{\vec{n}}(x)Q_{\vec{n}}(y) - \sum_{j=1}^{p} \frac{h_{\vec{n},j}}{h_{\vec{n}-\vec{e}_j,j}} P_{\vec{n}-\vec{e}_j}(x) Q_{\vec{n}+\vec{e}_j}(y). \tag{3.23}$$

For $p = 1$ the formula (3.23) reduces to the Christoffel–Darboux formula for orthogonal polynomials. For an extension to multiple orthogonal polynomials of mixed type, see [61], [62], and [15].

Before giving the proof of (3.23) we first discuss the Riemann–Hilbert problem for multiple orthogonal polynomials, which also gives a natural reformulation of (3.23).

3.4. Riemann–Hilbert problem. The Riemann–Hilbert problem for orthogonal polynomials was first formulated by Fokas, Its, and Kitaev [63]. Its extension to multiple orthogonal polynomials is due to Van Assche, Geronimo, and Kuijlaars [19].

We assume that the weight functions $w_1, \dots, w_p$ are sufficiently smooth (for example, C^1-smoothness will do) and that a multi-index $\vec{n} = (n_1, \dots, n_p)$ is given. Then the Riemann–Hilbert problem asks for a $(p+1) \times (p+1)$ matrix-valued function

$$Y : \mathbb{C} \setminus \mathbb{R} \to \mathbb{C}^{(p+1)\times(p+1)}$$

satisfying the following conditions:

(i) Y is analytic in $\mathbb{C} \setminus \mathbb{R}$;

(ii) $Y(z)$ has limit values $Y_+(x)$ and $Y_-(x)$ as $z \to x \in \mathbb{R}$ from the upper and lower half-planes, respectively, and these limit values satisfy the equality

$$Y_+(x) = Y_-(x) \begin{pmatrix} 1 & w_1(x) & w_2(x) & \cdots & w_p(x) \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \qquad x \in \mathbb{R};$$

(iii) as $z \to \infty$, we have

$$Y(z) = \left(I_{p+1} + O(z^{-1})\right) \begin{pmatrix} z^n & 0 & 0 & \cdots & 0 \\ 0 & z^{-n_1} & 0 & \cdots & 0 \\ 0 & 0 & z^{-n_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & z^{-n_p} \end{pmatrix}.$$

The Riemann–Hilbert problem has a unique solution if and only if the MOP $P_{\vec{n}}$ of degree $n = |\vec{n}|$ satisfying (3.16) is uniquely determined. If the MOPs $P_{\vec{n}-\vec{e}_j}$ also exist, then the unique solution of the Riemann–Hilbert problem is given by

$$Y(z) = \begin{pmatrix} P_{\vec{n}}(z) & \cdots & \dfrac{1}{2\pi i}\displaystyle\int \dfrac{P_{\vec{n}}(x)w_p(x)}{x-z}\,dx \\ -\dfrac{2\pi i}{h_{\vec{n}-\vec{e}_1,1}} P_{\vec{n}-\vec{e}_1}(z) & \cdots & \dfrac{1}{h_{\vec{n}-\vec{e}_1,1}}\displaystyle\int \dfrac{P_{\vec{n}-\vec{e}_1}(x)w_p(x)}{x-z}\,dx \\ \vdots & \ddots & \vdots \\ -\dfrac{2\pi i}{h_{\vec{n}-\vec{e}_p,p}} P_{\vec{n}-\vec{e}_p}(z) & \cdots & \dfrac{1}{h_{\vec{n}-\vec{e}_p,p}}\displaystyle\int \dfrac{P_{\vec{n}-\vec{e}_p}(x)w_p(x)}{x-z}\,dx \end{pmatrix}. \tag{3.24}$$

It is easy to see that the inverse transpose matrix

$$X(z) = Y^{-t}(z)$$

satisfies the following Riemann-Hilbert problem:

(i) X is analytic in $\mathbb{C} \setminus \mathbb{R}$;

(ii) X has limit values $X_+(x)$ and $X_-(x)$ for $x \in \mathbb{R}$ that satisfy the equality

$$X_+(x) = X_-(x) \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -w_1(x) & 1 & 0 & \cdots & 0 \\ -w_2(x) & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -w_p(x) & 0 & 0 & \cdots & 1 \end{pmatrix}, \qquad x \in \mathbb{R};$$

(iii) as $z \to \infty$ we have

$$X(z) = \left(I_{p+1} + O(z^{-1})\right) \begin{pmatrix} z^{-n} & 0 & 0 & \cdots & 0 \\ 0 & z^{n_1} & 0 & \cdots & 0 \\ 0 & 0 & z^{n_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & z^{n_p} \end{pmatrix}.$$

The unique solution of this Riemann–Hilbert problem involves MOPs of type I and appears as follows:

$$X(z) = \begin{pmatrix} \displaystyle\int \frac{Q_{\vec{n}}(x)}{z-x}\,dx & 2\pi i A_{\vec{n},1}(z) & \cdots & 2\pi i A_{\vec{n},p}(z) \\ \displaystyle\frac{h_{\vec{n},1}}{2\pi i}\int \frac{Q_{\vec{n}+\vec{e}_1}(x)}{z-x}\,dx & h_{\vec{n},1}A_{\vec{n}+\vec{e}_1,1}(z) & \cdots & h_{\vec{n},1}A_{\vec{n}+\vec{e}_1,p}(z) \\ \displaystyle\frac{h_{\vec{n},2}}{2\pi i}\int \frac{Q_{\vec{n}+\vec{e}_2}(x)}{z-x}\,dx & h_{\vec{n},2}A_{\vec{n}+\vec{e}_2,1}(z) & \cdots & h_{\vec{n},2}A_{\vec{n}+\vec{e}_2,p}(z) \\ \vdots & \vdots & \ddots & \vdots \\ \displaystyle\frac{h_{\vec{n},p}}{2\pi i}\int \frac{Q_{\vec{n}+\vec{e}_p}(x)}{z-x}\,dx & h_{\vec{n},p}A_{\vec{n}+\vec{e}_p,1}(z) & \cdots & h_{\vec{n},p}A_{\vec{n}+\vec{e}_p,p}(z) \end{pmatrix}. \tag{3.25}$$

Comparing (3.23), (3.24), and (3.25), we then see that

$$\begin{aligned}(x-y)K_n(x,y) &= \frac{1}{2\pi i}\sum_{j=1}^{p} w_j(y)\big[Y_+^{-1}(y)Y_+(x)\big]_{j+1,1} \\ &= \frac{1}{2\pi i}\begin{pmatrix}0 & w_1(y) & \cdots & w_p(y)\end{pmatrix} Y_+^{-1}(y)Y_+(x)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\end{aligned} \tag{3.26}$$

which is a convenient way to rewrite the Christoffel-Darboux kernel (3.23).

3.5. Proof of Theorem 3.1. We prove Theorem 3.1 in the form (3.26). The proof we give is modelled after the proof in [62].

The kernel K_n in (3.14) is the kernel of a (non-orthogonal) projection operator onto a space of polynomials. That is, if we also use the same K_n to denote the operator

$$(K_n g)(x) = \int K_n(x,y)g(y)\,dy,$$

then $K_n g = g$ for polynomials g of degree $\leqslant n-1$. The kernel (null-space) of the operator K_n consists of all g such that $\int \varphi_j(x)g(x)\,dx = 0$ for every $j = 1,\dots,n$.

The proof reduces to establishing the same facts for the operator defined by the right-hand side of (3.26). Thus, if $L_n(x,y)$ denotes the right-hand side of (3.26) and

$$(L_n g)(x) = \int L_n(x,y)g(y)\,dy,$$

then we prove that

(a) $L_n g = g$ if g is a polynomial of degree $\leqslant n-1$,
(b) $L_n g = 0$ if g is such that

$$\int \varphi_j(x)g(x)\,dx = 0$$

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

Proof. (a) If g is a polynomial of degree $\leqslant n-1$, then by (3.26)

$$(L_n g)(x) = \left(\frac{1}{2\pi i}\int \frac{g(y)-g(x)}{x-y}\begin{pmatrix}0 & w_1(y) & \cdots & w_p(y)\end{pmatrix} Y_+^{-1}(y)\,dy\right) Y_+(x)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} + \frac{g(x)}{2\pi i}\int \begin{pmatrix}0 & w_1(y) & \cdots & w_p(y)\end{pmatrix} Y_+^{-1}(y)Y_+(x)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\frac{dy}{x-y}. \tag{3.27}$$

Since $Y^{-1} = X^t$ we have by (3.25) that the entries of Y^{-1} in rows 2 up to p are multiple orthogonal polynomials of type I. The entries in the row vector

$$\begin{pmatrix}0 & w_1(y) & \cdots & w_p(y)\end{pmatrix} Y_+^{-1}(y)$$

are linear forms as in (3.20), since

$$\left[\begin{pmatrix}0 & w_1(y) & \cdots & w_p(y)\end{pmatrix} Y_+^{-1}(y)\right]_k = \begin{cases} 2\pi i Q_{\vec{n}}(y), & k=1,\\ h_{\vec{n},k-1} Q_{\vec{n}+\vec{e}_{k-1}}(y), & k=2,\dots,p+1.\end{cases} \tag{3.28}$$

For each x we get that $\dfrac{g(y)-g(x)}{x-y}$ is a polynomial in y of degree $\leqslant n-2$. The first term on the right-hand side of (3.27) is then 0 because of (3.28) and the type-I multiple orthogonality conditions (3.21).

What remains is the second term on the right-hand side of (3.27), for which we get from (3.28) and (3.25) that

$$\begin{aligned}(L_n g)(x) &= g(x)\int \left(Q_{\vec{n}}(y)Y_{11}(x) + \sum_{k=2}^{p+1}\frac{h_{\vec{n},k-1}}{2\pi i} Q_{\vec{n}+\vec{e}_{k-1}}(y)Y_{k1}(x)\right)\frac{dy}{x-y}\\ &= g(x)\sum_{k=1}^{p} X_{k1,+}(x)Y_{k1}(x) = g(x),\end{aligned}$$

where the last equality follows because $X = Y^{-t}$. This proves part (a).

(b) If g is such that $\int \varphi_j(y)g(y)\,dy = 0$ for every $j = 1,\dots,n$, then

$$(L_n g)(x) = \left(\frac{1}{2\pi i}\int g(y)\begin{pmatrix}0 & w_1(y) & \cdots & w_p(y)\end{pmatrix}\frac{Y_+^{-1}(y)-Y_+^{-1}(x)}{x-y}\,dy\right) Y_+(x)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} + \frac{1}{2\pi i}\int g(y)\begin{pmatrix}0 & w_1(y) & \cdots & w_p(y)\end{pmatrix} Y_+^{-1}(x)Y_+(x)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\frac{dy}{x-y}, \tag{3.29}$$

where the second term on the right-hand side clearly vanishes. Since $X = Y^{-t}$, we see from (3.24) and (3.19) that for $k = 1,\dots,p$ the entries in row $k+1$ of

$\dfrac{Y_+^{-1}(y) - Y_+^{-1}(x)}{x-y}$ are polynomials in y of degrees $\leqslant n_k - 1$. Thus, all entries in the row vector

$$\begin{pmatrix} 0 & w_1(y) & \cdots & w_p(y) \end{pmatrix} \frac{Y_+^{-1}(y) - Y_+^{-1}(x)}{x-y}$$

are in the linear span of the functions (3.9), which by the definition of a MOP ensemble is equal to the linear span of the functions $\varphi_1, \ldots, \varphi_n$. Since $\int \varphi_j(y) g(y) dy = 0$ for every $j = 1, \ldots, n$, we then find that $L_n g = 0$, which proves part (b).

4. Random matrix model with an external source

The Hermitian random matrix model with an external source is a generalization of the unitary random matrix model (3.1) in which the unitary invariance is discarded. This model was introduced and first analyzed in the physics papers [64]– [66]. Let us consider a potential V of the form (3.1) and a fixed $n \times n$ Hermitian matrix A. Then the random matrix model with an external source is given by the probability measure

$$\frac{1}{\widetilde{Z}_n} e^{-\operatorname{Tr}\left(V(M) - AM\right)} dM \tag{4.1}$$

on the space of $n \times n$ Hermitian matrices M. Because of the unitary invariance of dM and of the trace, the measure (4.1) depends only on the eigenvalues of the external source A, and we can assume without loss of generality that A is a diagonal matrix, say

$$A = \operatorname{diag}(a_1, \ldots, a_n),$$

with $a_i \in \mathbb{R}$ for $i = 1, \ldots, n$.

4.1. External source model as a MOP ensemble. A spectral decomposition of the Hermitian matrix M (see [9])

$$M = UXU^{-1}, \qquad X = \operatorname{diag}(x_1, \ldots, x_n),$$

where U is unitary and $x_1, \ldots, x_n$ are the eigenvalues of M, can be viewed as a change of variables. The Jacobian of this change of variables is proportional to $\prod_{j>i}(x_j - x_i)^2$, which means that (4.1) gives rise to the following probability density function on the eigenvalues of M (after averaging over the eigenvectors):

$$\frac{1}{\widehat{Z}_n} \prod_{j>i} (x_j - x_i)^2 \prod_{j=1} e^{-V(x_j)} \int_{U(n)} e^{AUXU^{-1}} dU, \tag{4.2}$$

where dU is the Haar measure on the unitary group $U(n)$. The integral in (4.2) can be evaluated explicitly using the Harish Chandra/Itzkyson–Zuber formula [67]

$$\int_{U(n)} e^{AUXU^{-1}} dU = \left(\prod_{j=0}^{n-1} j! \right) \frac{\det\left[e^{a_i x_j} \right]_{i,j=1,\ldots,n}}{\prod_{j>i}(a_j - a_i) \prod_{j>i}(x_j - x_i)}, \tag{4.3}$$

which is valid in case all the a_i and all the x_j are mutually distinct. In case some of the a_i (or some of the x_j) coincide, one should take an appropriate limit of (4.3).

Inserting (4.3) into (4.2), one finds a probability density function

$$\mathscr{P}(x_1,\dots,x_n)=\frac{1}{Z_n(\vec{a}\,)}\prod_{j>i}(x_j-x_i)\det\left[e^{a_ix_j}\right]_{i,j=1,\dots,n}\prod_{j=1}^{n}e^{-V(x_j)} \tag{4.4}$$

with a new constant

$$Z_n(\vec{a}\,)=\widehat{Z}_n\left(\prod_{j=0}^{n-1}j!\right)^{-1}\prod_{j>i}(a_j-a_i)$$

that depends on the eigenvalues a_j of the external source. Note that (4.4) is of the form (3.8) with the functions $\varphi_i(x)=e^{-(V(x)-a_ix)}$. It is valid when all the a_i are mutually distinct.

Now suppose that not all the a_i are distinct, and in fact suppose that $a_1,\dots,a_p$ are the mutually distinct eigenvalues of A and that a_j appears with multiplicity n_j as an eigenvalue of A. Then by taking the confluent limit of (4.4) we arrive at a multiple orthogonal polynomial ensemble (3.8) generated by the p weight functions

$$w_j(x)=e^{-(V(x)-a_jx)},\qquad j=1,\dots,p, \tag{4.5}$$

and the multi-index $(n_1,\dots,n_p)$.

4.2. Large-n asymptotics. To understand the large-n asymptotics of the random matrix model with an external source (4.1), it is important to first recall the large-n asymptotics of the standard random matrix model (3.1) (without an external source). For interesting large-n behaviour it is appropriate to replace V by nV, which creates an asymptotic balance between the repulsion of the eigenvalues and the confining potential. Then the orthogonal polynomials appearing in the correlation kernel (3.3) become dependent on n. We write $p_{k,n}$ for the orthonormal polynomial of degree k with respect to the weight $e^{-nV(x)}$. The correlation kernel (3.3) is then

$$K_n(x,y)=\sqrt{e^{-nV(x)}e^{-nV(y)}}\sum_{k=0}^{n-1}p_{k,n}(x)p_{k,n}(y). \tag{4.6}$$

The large-n limit can now be considered in the global (macroscopic) or local (microscopic) regimes. The global regime deals with the limit of the global eigenvalue distribution. The mean density of the eigenvalues is

$$\rho_n(x)=\frac{1}{n}\,K_n(x,x),$$

and the limit of $\rho_n(x)$ as $n\to\infty$ is of interest. The limit indeed exists, and the limit density is characterized by the classical equilibrium problem from logarithmic potential theory with external fields [68]: minimize

$$\iint\log\frac{1}{|x-y|}\,d\mu(x)\,d\mu(y)+\int V(x)\,d\mu(x) \tag{4.7}$$

over all probability measures μ on $\mathbb{R}$. There is a unique minimizer μ^* which has a density with respect to Lebesgue measure (provided that V is a C^2-function), and

$$\lim_{n\to\infty} \frac{1}{n} K_n(x,x) = \rho(x) := \frac{d\mu^*}{dx}. \tag{4.8}$$

The local regime deals with the behaviour of eigenvalues on a small scale where one can distinguish the individual eigenvalues. A main characteristic of random matrices is the universality of the asymptotic behaviour of the eigenvalues in the large-n limit. That is, it does not depend on the potential V. In terms of the correlation kernel (4.5) this universality is expressed by the limit

$$\lim_{n\to\infty} \frac{1}{n\rho(x_0)} K_n\left(x_0 + \frac{x}{n\rho(x_0)}, x_0 + \frac{y}{n\rho(x_0)}\right) = \frac{\sin \pi(x-y)}{\pi(x-y)}, \tag{4.9}$$

which holds for every x_0 such that $\rho(x_0) > 0$. The right-hand side of (4.9) is known as the sine kernel (see, for instance, [45] and [49]).

At edge points of the support of μ^* one has a limit that is different from (4.9). We consider a typical case when x_0 is a right endpoint and the density ρ of μ^* goes to zero like a square root, that is,

$$\rho(x) = \frac{c}{\pi}(x_0 - x)^{1/2}\big(1 + o(1)\big) \quad \text{as } x \to x_0,$$

where $c > 0$. In this case

$$\lim_{n\to\infty} \frac{1}{(cn)^{2/3}} K_n\left(x_0 + \frac{x}{(cn)^{2/3}}, x_0 + \frac{y}{(cn)^{2/3}}\right) = \frac{\operatorname{Ai}(x)\operatorname{Ai}'(y) - \operatorname{Ai}'(x)\operatorname{Ai}(y)}{x-y}, \tag{4.10}$$

where Ai is the Airy function, which is the unique solution of the Airy differential equation $y''(x) = xy(x)$ that behaves like

$$\operatorname{Ai}(x) = \frac{1}{2\sqrt{\pi}\, x^{1/4}} e^{-\frac{2}{3}x^{3/2}}\big(1 + O(x^{-3/2})\big) \quad \text{as } x \to +\infty.$$

Our aim is to extend the above results to a random matrix model with an external source and to more general multiple orthogonal polynomial ensembles. Here one takes $p \geqslant 2$ fixed and weights $w_1, \dots, w_p$ that may vary with n. For each n one has a multi-index $\vec{n} = (n_1, \dots, n_p)$ with $n = |\vec{n}|$, and as $n \to \infty$ one typically assumes the existence of the limits

$$\lim \frac{n_j}{n} = r_j. \tag{4.11}$$

Then, depending on the situation, one would like to express the global regime (4.8) in terms of a measure μ^* supplying a minimum in some equilibrium problem such as (4.7). In the local regime one would like to recover the sine kernel (4.9) and the Airy kernel (4.10) as local scaling limits of the correlation kernel at typical points x_0. At special points one would like to identify new kinds of asymptotic behaviour that can be expressed in terms of various limit kernels.

The results in this direction are by no means definitive. In the rest of this section we report on some progress for the random matrix model with an external source, and in the remaining sections we present results on other models that

give rise to multiple orthogonal polynomials. A main tool for obtaining the asymptotics is the Deift/Zhou steepest-descent method for Riemann–Hilbert problems as applied to the Riemann–Hilbert problem for multiple orthogonal polynomials. The Riemann–Hilbert problem is posed for matrices of size 3×3 or larger, which leads to a number of interesting new features in the asymptotic analysis when compared to the problem for 2×2 Riemann–Hilbert problems. However, we will not discuss this any further in the present paper.

4.3. Gaussian model with an external source. In order to observe interesting large-n asymptotic behaviour in the random matrix model with an external source, we replace V by nV and A by nA in (4.1), that is, we consider the model

$$\frac{1}{\widetilde{Z}_n} e^{-n \operatorname{Tr}(V(M)-AM)}\, dM. \tag{4.12}$$

The Gaussian case $V(x) = x^2/2$ corresponds to the weights

$$w_j(x) = e^{-n(\frac{1}{2}x^2 - a_j x)}, \qquad j = 1, \dots, p, \tag{4.13}$$

which are Gaussian weights centered at a_j. The MOPs for these weights are known as multiple Hermite polynomials, and they have many special properties, such as explicit differential equations and recurrence coefficients [69]. The Gaussian case with a very special external source

$$A = \operatorname{diag}\big(\underbrace{a, \dots, a}_{n/2}, \underbrace{-a, \dots, -a}_{n/2}\big), \tag{4.14}$$

where $a > 0$ is a fixed parameter and n is even, was analyzed in the three papers [21], [11], [22]. Let $K_n(x, y; a)$ denote the correlation kernel for the eigenvalues of the MOP ensemble corresponding to the random matrix model (4.12) with $V(x) = x^2/2$ and the external source (4.14).

The global regime is governed by the algebraic equation (Pastur equation [70])

$$\xi^3 - z\xi^2 + (1 - a^2)\xi + a^2 z = 0, \tag{4.15}$$

which gives an algebraic function $\xi(z)$ defined on a three-sheeted Riemann surface. The branch points of (4.15) are $\pm z_1(a)$ and $\pm z_2(a)$, where

$$z_1(a) = \frac{\big(8a^4 + 20a^2 - 1 + (8a^2+1)^{3/2}\big)^{1/2}}{2\sqrt{2}\, a},$$
$$z_2(a) = \frac{\big(8a^4 + 20a^2 - 1 - (8a^2+1)^{3/2}\big)^{1/2}}{2\sqrt{2}\, a}.$$

The branch point $z_1(a)$ is always real and positive. For $a > 1$ the branch point $z_2(a)$ is also real and positive, with $0 < z_2(a) < z_1(a)$. For $a = 1$ we have $z_2(a) = 0$, and $z_2(a)$ is purely imaginary for $0 < a < 1$. The structure of the sheets of the Riemann surface is shown in Fig. 14 for the three cases $a > 1$, $a = 1$, and $0 < a < 1$.

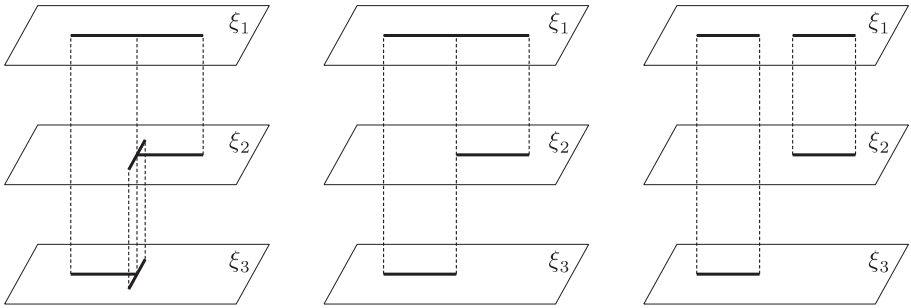

Figure 14. The structure of the Riemann surface for the equation (4.15) for the values $a > 1$ (left), $a = 1$ (middle), and $a < 1$ (right). In all cases the eigenvalues of M accumulate on the interval(s) of the first sheet with a density given by (4.16).

There is a solution $\xi_1(z)$ of (4.15) satisfying the condition

$$\xi_1(z;a) = z - \frac{1}{z} + O(z^{-2}) \quad \text{as } z \to \infty$$

and possessing an analytic continuation to the first sheet of the Riemann surface. The limit mean density of the eigenvalues in all cases is given by

$$\lim_{n\to\infty} \frac{1}{n} K_n(x,x;a) = \rho(x;a)\, \frac{1}{\pi} \operatorname{Im} \xi_{1,+}(x;a), \qquad x \in \mathbb{R}, \tag{4.16}$$

where $\xi_{1,+}(x;a)$ denotes the limit value of $\xi_1(z;a)$ as $z \to x \in \mathbb{R}$ with $\operatorname{Im} z > 0$.

In the local regime one finds the sine kernel (4.9) at points $x_0 \in \big(-z_1(a), z_1(a)\big)$ if $0 < a < 1$, and at points $x_0 \in \big(-z_1(a), -z_2(a)\big) \cup \big(z_2(a), z_1(a)\big)$ if $a > 1$. The Airy kernel limit (4.10) (appropriately modified in case of a left endpoint) is valid at the branch points $x_0 = \pm z_1(a)$ (in all cases) and at $x_0 = \pm z_2(a)$ in the case $a > 1$.

A new phenomenon in this model is the phase transition at $a = 1$. In the global regime we get by (4.16) that

$$\rho(x;1) = \frac{\sqrt{3}}{2} |x|^{1/3} + O(x) \quad \text{as } x \to 0. \tag{4.17}$$

The different exponent $1/3$ in the limit mean density also indicates a different local behaviour at $x = 0$. This is indeed the case, and one has the following double scaling limit [64], [71], [22] as described in the next theorem.

Theorem 4.1. *For every fixed $b \in \mathbb{R}$*

$$\frac{1}{n^{3/4}} K_n\left(\frac{x}{n^{3/4}}, \frac{y}{n^{3/4}}; 1 + \frac{b}{2\sqrt{n}}\right) = K^{\mathrm{Pe}}(x,y;b), \tag{4.18}$$

where K^{Pe} is the Pearcey kernel

$$K^{\mathrm{Pe}}(x,y;b) = \frac{p(x)q''(y) - p'(x)q'(y) + p''(x)q(y) - bp(x)q(y)}{x-y}, \tag{4.19}$$

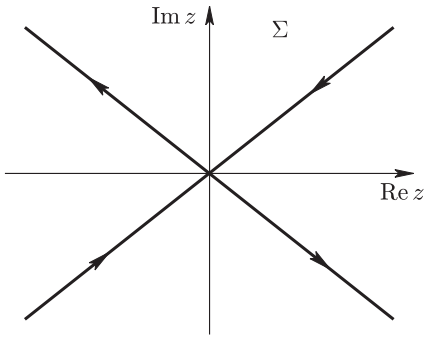

Figure 15. The contour Σ in the definition (4.20) of $q(y)$.

with

$$p(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{4}s^4-\frac{b}{2}s^2+isx}\,ds \quad \textit{and} \quad q(y)=\frac{1}{2\pi}\int_{\Sigma} e^{\frac{1}{4}t^4+\frac{b}{2}t^2+ity}\,dt. \tag{4.20}$$

The contour Σ *consists of the four rays* $\arg y=\pm\pi/4, \pm 3\pi/4$, *with the orientation shown in Fig.* 15.

The functions (4.20) are known as Pearcey integrals. They are solutions of the third-order differential equations

$$p'''(x)=xp(x)+bp'(x), \qquad q'''(y)=-yq(y)+bq'(y), \tag{4.21}$$

respectively. Brézin and Hikami [64] also gave the double-integral representation

$$K^{\mathrm{Pe}}(x,y;b)=\frac{1}{(2\pi i)^2}\int_{\Sigma}\int_{-i\infty}^{i\infty}\exp\left\{-\frac{1}{4}\,s^4+\frac{b}{2}\,s^2-ys+\frac{1}{4}\,t^4-\frac{b}{2}\,t^2+xt\right\}\frac{ds\,dt}{s-t} \tag{4.22}$$

for the Pearcey kernel (4.19).

4.4. Non-Gaussian model with an external source. A steepest-descent analysis in the Gaussian random matrix model with an external source was carried out in [21], [11], [22] on the basis of the algebraic equation (4.15). In the case of a higher-degree polynomial potential V in (4.12) and an external source (4.14) one still expects the existence of a third-order algebraic equation (spectral curve) that could be used in the asymptotic analysis. The corresponding Riemann surface could be of higher genus, and it seems to be very difficult to write it explicitly in general. See [72], however, for the quartic model $V(x)=x^4/4$ and sufficiently large a.

An alternative approach was developed in [14] (see also [12] and [13]), where the algebraic equation follows indirectly from a vector equilibrium problem for two measures. This approach works for the case of a random matrix model with an external source (4.12) when the potential is an even polynomial

$$V(x)=\sum_{j=1}^{d} v_j x^{2d}, \qquad v_d>0, \tag{4.23}$$

and the external source has the form (4.14) with $a > 0$. The vector equilibrium problem in [14] requires minimizing the following energy functional, which is a generalization of (4.7):

$$\iint \log \frac{1}{|x-y|}\, d\mu_1(x)\, d\mu_1(y) + \iint \log \frac{1}{|x-y|}\, d\mu_2(x)\, d\mu_2(y) - \iint \log \frac{1}{|x-y|}\, d\mu_1(x)\, d\mu_2(y) + \int \big(V(x) - a|x|\big)\, d\mu_1(x). \tag{4.24}$$

The energy functional (4.24) involves two measures μ_1 and μ_2, and it is assumed that

(a) μ_1 and μ_2 have finite logarithmic energy,
(b) μ_1 is a measure on $\mathbb{R}$ with total mass 1,
(c) μ_2 is a measure on $i\mathbb{R}$ with total mass $1/2$ and satisfies the constraint

$$\mu_2 \leqslant \sigma, \tag{4.25}$$

where σ is the mesure on $i\mathbb{R}$ with constant density

$$\frac{d\sigma}{|dz|} = \frac{a}{\pi}, \qquad z \in i\mathbb{R}. \tag{4.26}$$

A notable feature is the appearance of an upper constraint σ on the measure μ_2 (see (4.25)), while an external field $V(x) - a|x|$ acts on μ_1. Equilibrium problems with an upper constraint first appeared in the context of orthogonal polynomials with discrete orthogonality (see [73]–[75]). The interaction between μ_1 and μ_2 in (4.24) is of Nikishin type [34]. A vector potential equilibrium problem for three measures equivalent to (4.24)–(4.26) was considered in [12] and [13]. It is analogous to the equilibrium problems in §§ 1 and 2 (with the same interaction matrix and the same relations for the total masses of the measures, but with the addition of external fields determined by the external source potential $V(x) = x^4/4 - tx^2/2$).

Theorem 4.2. *The above energy functional has a unique minimizer (μ_1^*, μ_2^*) among all vectors of measures (μ_1, μ_2) satisfying the above conditions (a), (b), (c). The minimizer has the following properties.*

(i) *The support of μ_1^* is bounded and is a finite union of closed intervals,*

$$S(\mu_1^*) = \bigcup_{j=1}^{N} [a_j, b_j]. \tag{4.27}$$

(ii) *The support of μ_2^* is the full imaginary axis and there exists a $c \geqslant 0$ such that*

$$S(\sigma - \mu_2^*) = (-i\infty, -ic] \cup [ic, i\infty). \tag{4.28}$$

(iii) *μ_1^* has a density ρ with respect to Lebesgue measure. In regular cases (to be explained below) ρ is the limit mean density of the eigenvalues of the random matrix M in the external source model (4.12) as $n \to \infty$, that is,*

$$\rho(x) = \lim_{n\to\infty} \frac{1}{n} K_n(x, x),$$

where K_n is the correlation kernel for the eigenvalues in (4.12).

The convergence in part (iii) has been proved only for regular cases, although it is believed to be true in general.

Regularity is described in terms of variational conditions associated with the vector equilibrium problem. There is a constant ℓ such that

$$2V^{\mu_1^*}(x) = V^{\mu_2^*}(x) - V(x) + a|x| - \ell, \qquad x \in S(\mu_1^*), \tag{4.29}$$

$$2V^{\mu_1^*}(x) \geqslant V^{\mu_2^*}(x) - V(x) + a|x| - \ell, \qquad x \in \mathbb{R} \setminus S(\mu_1^*), \tag{4.30}$$

where V^μ denotes the logarithmic potential of μ (see (1.26)). We say that μ_1^* is regular if the density of μ_1^* is positive on the interior of each interval $[a_j, b_j]$ in $S(\mu_1^*)$ and vanishes like a square root at each endpoint, and if there is strict inequality in the variational condition (4.30). We say that μ_2^* is regular if $c > 0$ (see (4.28)), or if $c = 0$ and there is strict inequality in (4.25), that is,

$$\frac{d\mu_2^*(z)}{|dz|} < \frac{a}{\pi}, \qquad z \in i\mathbb{R}.$$

The usual scaling limits for the local regime, that is, the sine kernel (4.9) at interior points x_0 of $S(\mu_1^*)$ and the Airy kernel (4.10) at points $x_0 \in \{a_1, b_1, \ldots, a_N, b_N\}$, are also obtained as a result of the analysis in [14].

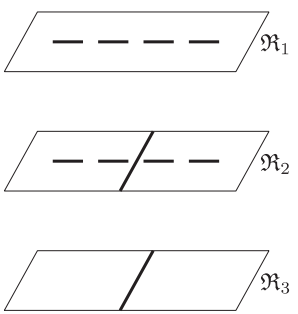

Figure 16. Riemann surface for Case **I**, when $S(\sigma - \mu_2^*) = i\mathbb{R}$.

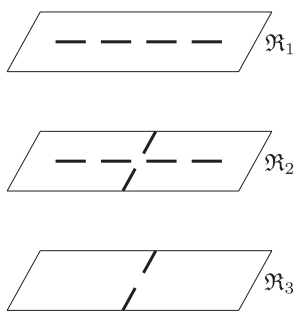

Figure 17. Riemann surface for Case **II**, when $0 \notin S(\mu_1^*) \cap S(\sigma - \mu_2^*)$.

A main tool in the analysis in [14] is a three-sheeted Riemann surface constructed from the solution of the vector equilibrium problem as follows. The supports $S(\mu_1^*)$ and $S(\sigma - \mu_2^*)$ are used to define the three sheets

$$\begin{aligned} \mathfrak{R}_1 &= \overline{\mathbb{C}} \setminus S(\mu_1^*), \\ \mathfrak{R}_2 &= \overline{\mathbb{C}} \setminus \big(S(\mu_1^*) \cup S(\sigma - \mu_2^*)\big), \\ \mathfrak{R}_3 &= \overline{\mathbb{C}} \setminus S(\sigma - \mu_2^*), \end{aligned} \tag{4.31}$$

which are then glued together along the cuts $S(\mu_1^*)$ and $S(\sigma - \mu_2^*)$ in the usual crosswise manner to produce a compact Riemann surface $\mathfrak{R}$. In a regular case 0 can belong to at most one of the cuts $S(\mu_1^*)$ or $S(\sigma - \mu_2^*)$, and we can distinguish the following three cases.

Case I: $0 \notin S(\mu_1^*)$ and $0 \in S(\sigma - \mu_2^*)$.
Case II: $0 \notin S(\mu_1^*)$ and $0 \notin S(\sigma - \mu_2^*)$.
Case III: $0 \in S(\mu_1^*)$ and $0 \notin S(\sigma - \mu_2^*)$.

The first two cases are illustrated in Figs. 16 and 17.

In all three cases we define

$$\begin{aligned}
\xi_1(z) &= V'(z) - \int \frac{d\mu_1^*(s)}{z-s}, && z \in \mathfrak{R}_1, \\
\xi_2(z) &= \pm a + \int \frac{d\mu_1^*(s)}{z-s} - \int \frac{d\mu_2^*(s)}{z-s}, && z \in \mathfrak{R}_2, \quad \pm \operatorname{Re} z > 0, \\
\xi_3(z) &= \mp a + \int \frac{d\mu_2^*(s)}{z-s}, && z \in \mathfrak{R}_3, \quad \pm \operatorname{Re} z > 0,
\end{aligned} \tag{4.32}$$

and it turns out that these functions can be continued analytically to meromorphic functions on the full compact Riemann surface whose only pole is at infinity on the first sheet. This is a consequence of the variational conditions associated with the vector equilibrium problem. The three functions ξ_j are then the three solutions of an algebraic equation

$$\xi^3 + p_2(z)\xi^2 + p_1(z)\xi + p_0(z) = 0$$

with polynomial coefficients. This is how one obtains the spectral curve from the vector equilibrium problem. In the case of the external source potential $V(x) = x^4/4 - tx^2/2$ explicit expressions for the coefficients of the equation of the spectral curve were obtained in [12].

The remaining part of the proof of Theorem 4.2 consists of a steepest-descent analysis of a Riemann–Hilbert problem for multiple orthogonal polynomials analogous to the problem in § 3.4, in the present case of size 3×3:

(i) $Y : \mathbb{C} \setminus \mathbb{R} \to \mathbb{C}^{3\times 3}$ is analytic;

(ii) $Y_+(x) = Y_-(x) \begin{pmatrix} 1 & e^{-n(V(x)-ax)} & e^{-n(V(x)+ax)} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ for $x \in \mathbb{R}$;

(iii) $Y(z) = \bigl(I_3 + O(z^{-1})\bigr) \begin{pmatrix} z^n & 0 & 0 \\ 0 & z^{-n/2} & 0 \\ 0 & 0 & z^{-n/2} \end{pmatrix}$ as $z \to \infty$.

This analysis consists of a sequence of explicit transformations $Y \mapsto X \mapsto T \mapsto S \mapsto R$ leading to a Riemann–Hilbert problem for a function R that is normalized at infinity (namely, $R(z) = I + O(1/z)$ as $z \to \infty$) and has a jump tending to the identity matrix as $n \to \infty$. As a consequence, we get that $R(z) \to I$ uniformly with respect to z as $n \to \infty$. The steps in the steepest-descent analysis are too involved to reproduce here.

The transitions between the three cases represent non-regular behaviour. For the case of a general even quartic potential

$$V(x) = \frac{1}{4}\, x^4 - \frac{t}{2}\, x^2$$

it is possible to describe the transitions between the cases explicitly. This gives the phase diagram in Fig. 18, which is taken from [14] (see also [13]).

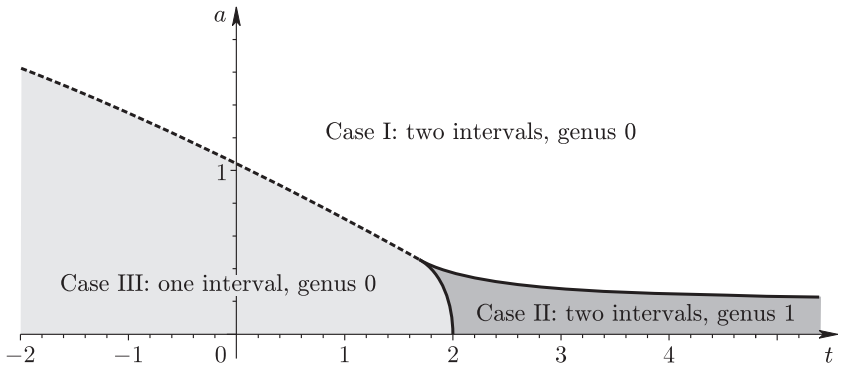

Figure 18. Phase diagram for the quartic potential $x^4/4 - tx^2/2$. The solid curve corresponds to the Painlevé II transition. The dotted curve corresponds to the Pearcey transition.

The transition from Case **I** to Case **III** is of the Pearcey type, which we already encountered in the Gaussian case. It leads to a double scaling limit involving the Pearcey kernel as in (4.18). The transition from Case **III** to Case **II** is of a different nature. Here also a gap in the support of μ_1^* closes, but now it also corresponds to a change in the genus of the Riemann surface. This transition is similar to the closing (or opening) of a gap in the one-matrix model, which is known to be described by a family of kernels involving Lax-pair solutions associated with the Hastings–McLeod solution of the Painlevé II equation (see [76], [77]). Also, the transition from Case **I** to Case **II** is of Painlevé II type, although now the opening of a gap is in the support of $\sigma - \mu_2^*$ on the imaginary axis. This affects only the (non-physical) second and third sheets, and therefore does not lead to any special critical behaviour in the local asymptotics of the eigenvalues.

5. Non-intersecting paths

Eigenvalues of random matrices can give rise to determinantal point processes which may reduce to MOP ensembles, as we have seen. Non-intersecting path ensembles are another major source for determinantal point processes. In very special cases these also reduce to MOP ensembles.

5.1. Non-intersecting Brownian motions. The setup for non-intersecting path ensembles is the following. Consider a one-dimensional diffusion process. This is a strong Markov process on the real line with continuous sample paths. We consider n independent copies with given starting positions $a_1 < a_2 < \cdots < a_n$ at time $t = 0$. A remarkable formula of Karlin and McGregor [78] then says that the probability density of the event of finding the paths at the positions $x_1, \ldots, x_n$ at time $t > 0$ without any two of them having intersected in the time interval $[0, t]$ is proportional to the determinant

$$\det\big[p_t(a_i, x_j)\big]_{i,j=1,\ldots,n}, \qquad x_1 < x_2 < \cdots < x_n, \tag{5.1}$$

where $p_t(a, x)$ denotes the transition probability density of the diffusion process.

The determinant (5.1) does not itself lead to a determinantal point process. However, this does happen if we also prescribe certain ending positions at some later time. Indeed, if we consider n independent copies with given starting positions $a_1 < \cdots < a_n$ and given ending positions $b_1 < \cdots < b_n$ at a time $T > 0$, then the event of finding the paths at prescribed points at a time $t \in (0, T)$ without any two of them having intersected in the time interval $[0, T]$ has a probability density that is proportional to the product of two determinants:

$$\det\big[p_t(a_i, x_j)\big]_{i,j=1,\dots,n} \det\big[p_{T-t}(x_j, b_i)\big]_{i,j=1,\dots,n}. \tag{5.2}$$

This is in fact a consequence of the strong Markov property.

Put otherwise, if we condition on the event that the paths do not intersect in the full time interval $[0, T]$, then the probability density function for the positions at time $t \in (0, T)$ is equal to

$$\frac{1}{Z_n} \det\big[p_t(a_i, x_j)\big]_{i,j=1,\dots,n} \det\big[p_{T-t}(x_j, b_i)\big]_{i,j=1,\dots,n} \tag{5.3}$$

with a certain normalization constant Z_n. The expression (5.3) is invariant under permutations of the x_j, and therefore by symmetrization we can also view (5.3) as a probability density function on the whole of $\mathbb{R}^n$ (with a different normalization constant), which is what we typically prefer to do.

The probability density function (5.3) is indeed a determinantal point process, more precisely, a biorthogonal ensemble (cf. (3.15), and see [57]). Only in exceptional cases does it reduce to the expression (3.8) for a MOP ensemble. The primary example in which there is a reduction to a MOP ensemble is the case of a Brownian motion (actually, Brownian bridges) with the transition probability density

$$p_t(x, y) = \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(x-y)^2}{2t}\right\}, \qquad t > 0. \tag{5.4}$$

In the fully confluent case when $a_i \to 0$ and $b_i \to 0$ for all i, the probability density function (5.3) has a limit which can be expressed as

$$\frac{1}{Z_n} \prod_{1 \leqslant j < k \leqslant n} (x_k - x_j)^2 \prod_{j=1}^{n} \exp\left\{-\frac{T}{2t(T-t)}\, x_j^2\right\}$$

with a different constant Z_n. This is the same as the probability density function (3.2) for the eigenvalues of a unitary invariant random matrix ensemble with potential $V(x) = \dfrac{T}{2t(T-t)}\, x_j^2$. This is the Gaussian Unitary Ensemble (GUE), up to a scaling factor.

If we again let $b_i \to 0$ for all i in (5.3) but take p different starting values $a_1, \dots, a_p$ with n_j of the paths starting at a_j, then (5.3)–(5.4) turns into a MOP ensemble with weights

$$w_j(x) = \exp\left\{-\frac{T}{2t(T-t)}\, x^2 + \frac{a_j}{t}\, x\right\}, \qquad j = 1, \dots, p, \tag{5.5}$$

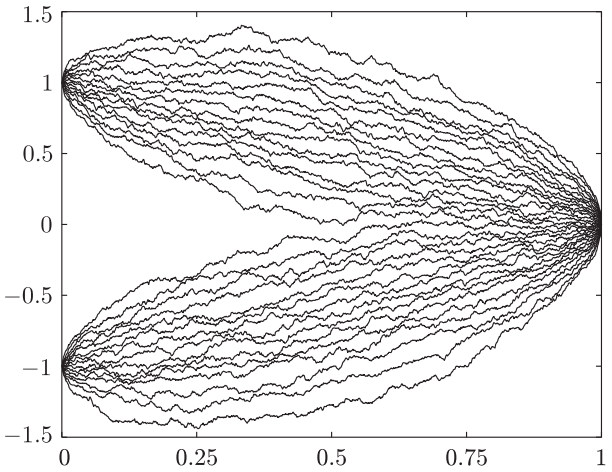

Figure 19. Non-intersecting Brownian bridges starting at two different values and ending at 0. At any time $t \in (0,1)$ the positions of the paths have the same distribution as the eigenvalues of an $n \times n$ Gaussian random matrix with an external source. The distribution is a MOP with two Gaussian weights of the form (5.5).

and multi-index $(n_1, \dots, n_p)$. These weights are, up to a scaling factor, equal to the weights in (4.13) arising from the Gaussian matrix model with an external source. The conclusion is that the eigenvalues in this random matrix model can also be viewed as the positions of n non-intersecting Brownian motions with p starting positions and a single ending point.

See Fig. 19 for an illustration with $p = 2$ starting points and $T = 1$. From the figure one sees that a limit shape is likely to appear in the large-n limit (after appropriate rescaling of the variance of the Brownian motion). This is indeed the case. There is a limiting heart-shaped region that contains the paths as $n \to \infty$. For small time t the paths stay in two separate groups. This corresponds to large values of a in the external source model. At a certain critical time the two groups come together and they continue as one group until the end. The transition at the critical time is of the Pearcey type as discussed in Theorem 4.1 in the context of the random matrix model with an external source.

5.2. Non-intersecting squared Bessel paths. The squared Bessel process is another one-dimensional diffusion process which in the confluent case gives rise to a MOP ensemble. The squared Bessel process is a Markov process on $[0, \infty)$ depending on a parameter $\alpha > -1$ and with transition probability density

$$\begin{aligned} p_t^{(\alpha)}(x,y) &= \frac{1}{2t}\left(\frac{y}{x}\right)^{\alpha/2} e^{-(x+y)/(2t)} I_\alpha\left(\frac{\sqrt{xy}}{t}\right), && x > 0, \quad y \geqslant 0, \\ p_t^{(\alpha)}(0,y) &= \frac{y^\alpha}{(2t)^{\alpha+1}\Gamma(\alpha+1)} e^{-y/(2t)}, && y \geqslant 0, \end{aligned} \tag{5.6}$$

where I_α denotes the modified Bessel function of the first kind of order α. Here the probability of a transition from x to y is greater for $y-x>0$ than for $y-x<0$, and for $x=0$ no transitions to $y<0$ are allowed.

We now consider n non-intersecting squared Bessel paths with starting positions $0<a_1<\cdots<a_n$ and ending positions $0<b_1<\cdots<b_n$ at time $T>0$. At any intermediate time $t\in(0,T)$ we have the probability density function (5.3) for the positions at time t. This will reduce to a MOP ensemble if we let $b_i\to 0$ for all i (or alternatively if $a_i\to 0$ for all i).

In the fully confluent limit where $a_i\to a>0$ and $b_j\to 0$ for all i and j, the probability density function (5.3) takes the form of a MOP ensemble (3.8) with weight functions

$$\begin{aligned} w_1(x) &= x^{\alpha/2}\exp\left\{-\frac{T}{2t(T-t)}\,x\right\}I_\alpha\left(\frac{\sqrt{ax}}{t}\right),\\ w_2(x) &= x^{(\alpha+1)/2}\exp\left\{-\frac{T}{2t(T-t)}\,x\right\}I_{\alpha+1}\left(\frac{\sqrt{ax}}{t}\right) \end{aligned} \tag{5.7}$$

and multi-index (n_1,n_2), where $n_1=\lceil n/2\rceil$ and $n_2=\lfloor n/2\rfloor$ (see [79]). In the limit as $a\to 0$ this further reduces to an orthogonal polynomial ensemble on $[0,\infty)$ with a Laguerre weight.

After appropriate time rescaling

$$t\mapsto\frac{t}{2n}\,,\qquad T\mapsto\frac{1}{2n}$$

the squared Bessel paths fill out a domain in the tx plane that was described explicitly in [79]. See Fig. 20 for an illustration. There is a critical time t^* such that for $t<t^*$ the paths stay away from the wall (the hard edge) at 0, while for $t>t^*$ the smallest paths come close to the wall as $n\to\infty$.

At any time $t\in(0,1)$ the paths have a limit mean distribution that is characterized by the following vector equilibrium problem (see [79]), which is of a similar nature to the one for the external source model (see (4.24)–(4.26)). The vector equilibrium measure must minimize

$$\begin{aligned} &\iint\log\frac{1}{|x-y|}\,d\mu_1(x)\,d\mu_1(y)+\iint\log\frac{1}{|x-y|}\,d\mu_2(x)\,d\mu_2(y)\\ &\quad-\iint\log\frac{1}{|x-y|}\,d\mu_1(x)\,d\mu_2(y)+\int\left(\frac{x}{t(1-t)}-\frac{2\sqrt{ax}}{t}\right)d\mu_1(x) \end{aligned} \tag{5.8}$$

among all vectors of measures (μ_1,μ_2) satisfying

(a) μ_1 and μ_2 have finite logarithmic energy,

(b) μ_1 is a measure on $[0,\infty)$ with total mass 1,

(c) μ_2 is a measure on $(-\infty,0]$ with total mass $1/2$ and satisfying the constraint

$$\mu_2\leqslant\sigma, \tag{5.9}$$

where σ is the measure on $(-\infty,0]$ with density

$$\frac{d\sigma}{dx}=\frac{\sqrt{a}}{\pi t}\,|x|^{-1/2},\qquad x<0. \tag{5.10}$$

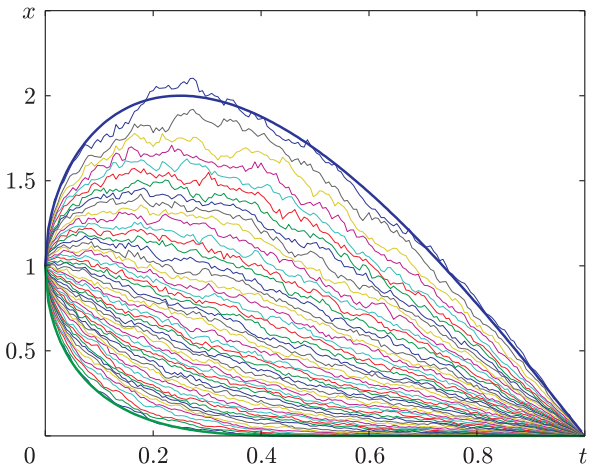

Figure 20. 50 non-intersecting squared Bessel paths starting at $a = 1$ and ending at $b = 0$. The solid curves denote the boundary of the domain that is filled by the squared Bessel paths as $n \to \infty$.

There is a unique minimizer (μ_1^*, μ_2^*), and μ_1^* corresponds to the limit mean density of paths at time t. The measures μ_j^*, $j = 1, 2$, are absolutely continuous with respect to Lebesgue measure, with densities that can be expressed in terms of solutions of the cubic equation (spectral curve)

$$z\xi^3 - \frac{2z}{t(1-t)}\,\xi^2 + \left(\frac{z}{t^2(1-t)^2} + \frac{1}{t(1-t)} - \frac{a}{t^2}\right)\xi - \frac{1}{t^2(1-t)^2} = 0. \tag{5.11}$$

The equation (5.11) has three solutions which we choose such that

$$\begin{aligned} \xi_1(z) &= \frac{1}{z} + O(z^{-2}), \\ \xi_2(z) &= \frac{1}{t(1-t)} - \frac{\sqrt{a}}{tz^{1/2}} - \frac{1}{2z} + O(z^{-3/2}), \\ \xi_3(z) &= \frac{1}{t(1-t)} + \frac{\sqrt{a}}{tz^{1/2}} - \frac{1}{2z} - O(z^{-3/2}) \end{aligned} \tag{5.12}$$

as $z \to \infty$. Then μ_1^* and μ_2^* are given by

$$\begin{aligned} \frac{d\mu_1^*}{dx} &= \frac{1}{\pi}\operatorname{Im}\xi_{2,+}(x), && x > 0, \\ \frac{d\mu_2^*}{dx} &= \frac{d\sigma}{dx} - \frac{1}{\pi}\operatorname{Im}\xi_{2,+}(x), && x < 0. \end{aligned} \tag{5.13}$$

The transition at the critical time t^* can be seen in the behaviour of the measures μ_1^* and μ_2^* as follows. For $t < t^*$ one has $S(\mu_1^*) = [p, q]$ with $0 < p < q < \infty$, and $S(\sigma - \mu_2^*) = (-\infty, 0]$, while for $t > t^*$ one has $S(\mu_1^*) = [0, q]$, and $S(\sigma - \mu_2^*) = (-\infty, -p]$ with $p > 0$. This means that the constraint (5.9) is active only for $t > t^*$.

The local behaviour at the critical time t^* was studied recently in [80]. See also [81] for a study of the case of non-intersecting squared Bessel paths with positive starting and ending positions.

6. The two-matrix model

The two-matrix model is another model from random matrix theory that gives rise to MOP ensembles. We discuss this model in the final section of this survey.

6.1. Definitions. The Hermitian two-matrix model is a probability measure of the form

$$\frac{1}{Z_n}\exp\bigl\{-n\operatorname{Tr}\bigl(V(M_1)+W(M_2)-\tau M_1M_2\bigr)\bigr\}\,dM_1\,dM_2, \tag{6.1}$$

defined on the space of pairs (M_1, M_2) of $n\times n$ Hermitian matrices. The number $\tau\in\mathbb{R}\setminus\{0\}$ is known as the coupling constant. There are two potentials V and W in (6.1), which typically are assumed to be polynomials of even degree with positive leading coefficients, so that (6.1) is indeed a probability measure.

The two-matrix model (6.1) with polynomial potentials V and W was introduced in [67], [82] as a model for quantum gravity and string theory. In this context the interest is in the double scaling limit for critical potentials. It is generally believed that the two-matrix model is able to describe all (p, q) conformal minimal models, whereas the one-matrix model is limited to $(p, 2)$ minimal models.

The eigenvalues of the matrices M_1 and M_2 in the two-matrix model give a determinantal point process with correlation kernels expressible in terms of so-called biorthogonal polynomials. These are two families $\{p_{k,n}(x)\}_{k=0}^{\infty}$ and $\{q_{l,n}(y)\}_{y=0}^{\infty}$ of monic polynomials, with $p_{k,n}$ of degree k and $q_{l,n}$ of degree l, satisfying the condition

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}p_{k,n}(x)q_{l,n}(y)\exp\{-n(V(x)+W(y)-\tau xy)\}\,dx\,dy=h_{k,n}^2\delta_{k,l}. \tag{6.2}$$

Ercolani and McLaughlin [83] showed that these polynomials are uniquely characterized by (6.2) and that they have simple and real zeros. Quite recently it was shown that the zeros interlace [84]. The correlation kernels are expressed in terms of these biorthogonal polynomials and their transformed functions

$$Q_{l,n}(x)=\exp\{-nV(x)\}\int_{-\infty}^{\infty}q_{l,n}(y)\exp\{-n(W(y)-\tau xy)\}\,dy,$$
$$P_{k,n}(y)=\exp\{-nW(y)\}\int_{-\infty}^{\infty}p_{k,n}(x)\exp\{-n(V(x)-\tau xy)\}\,dx$$

as follows:

$$\begin{aligned} K_{11}^{(n)}(x_1,x_2) &= \sum_{k=0}^{n-1} \frac{1}{h_{k,n}^2}\, p_{k,n}(x_1) Q_{k,n}(x_2), \\ K_{12}^{(n)}(x,y) &= \sum_{k=0}^{n-1} \frac{1}{h_{k,n}^2}\, p_{k,n}(x) q_{k,n}(y), \\ K_{21}^{(n)}(y,x) &= \sum_{k=0}^{n-1} \frac{1}{h_{k,n}^2}\, P_{k,n}(y) Q_{k,n}(x) - \exp\{-n(V(x)+W(y)-\tau xy)\}, \\ K_{22}^{(n)}(y_1,y_2) &= \sum_{k=0}^{n-1} \frac{1}{h_{k,n}^2}\, P_{k,n}(y_1) q_{k,n}(y_2). \end{aligned} \tag{6.3}$$

Then the joint probability density function for the eigenvalues $x_1,\dots,x_n$ of M_1 and the eigenvalues $y_1,\dots,y_n$ of M_2 is given by

$$\begin{aligned} &\mathscr{P}(x_1,\dots,x_n,y_1,\dots,y_n) \\ &\quad = \frac{1}{(n!)^2} \det \begin{pmatrix} \big(K_{11}^{(n)}(x_i,x_j)\big)_{i,j=1}^n & \big(K_{12}^{(n)}(x_i,y_j)\big)_{i,j=1}^n \\ \big(K_{21}^{(n)}(y_i,x_j)\big)_{i,j=1}^n & \big(K_{22}^{(n)}(y_i,y_j)\big)_{i,j=1}^n \end{pmatrix}, \end{aligned}$$

and the marginal densities take the form

$$\begin{aligned} &\underbrace{\int\cdots\int}_{n-k+n-l \text{ times}} \mathscr{P}(x_1,\dots,x_n,y_1,\dots,y_n)\, dx_{k+1}\cdots dx_n\, dy_{l+1}\cdots dy_n \\ &\quad = \frac{(n-k)!\,(n-l)!}{(n!)^2} \det \begin{pmatrix} \big(K_{11}^{(n)}(x_i,x_j)\big)_{i,j=1}^k & \big(K_{12}^{(n)}(x_i,y_j)\big)_{i,j=1}^{k,l} \\ \big(K_{21}^{(n)}(y_i,x_j)\big)_{i,j=1}^{l,k} & \big(K_{22}^{(n)}(y_i,y_j)\big)_{i,j=1}^l \end{pmatrix} \end{aligned} \tag{6.4}$$

(see [85]–[87]). In particular, by taking $l=0$, so that we average over all eigenvalues of M_2, we find that the eigenvalues of M_1 give a determinantal point process (3.12) with kernel $K_{11}^{(n)}$.

The biorthogonal polynomials characterized by (6.2) can be viewed as multiple orthogonal polynomials in the case of polynomial potentials V and W (see [10]). Indeed, if $\deg W = p+1$, then $p_{n,n}$ is a multiple orthogonal polynomial for the p weights

$$w_{j,n}(x) = e^{-nV(x)} \int_{-\infty}^{\infty} y^j \exp\{-n(W(y)-\tau xy)\}\, dy, \qquad j=0,\dots,p-1, \tag{6.5}$$

and the diagonal multi-index $\vec{n} = (n/p, n/p, \dots, n/p)$, provided that n is a multiple of p. If n is not a multiple of p, say $n \equiv r$ modulo p with $1 \leqslant r \leqslant p-1$, then the multi-index is

$$\vec{n} = \big(\underbrace{\lceil n/p\rceil,\dots,\lceil n/p\rceil}_{r \text{ times}}, \underbrace{\lfloor n/p\rfloor,\dots,\lfloor n/p\rfloor}_{p-r \text{ times}}\big).$$

Note that for odd j the weight (6.5) is not necessarily non-negative on $\mathbb{R}$. This does not play a role, however, since the MOPs with respect to the weights (6.5) are uniquely determined (by the result in [83]), and therefore the algebraic properties discussed in § 3 continue to hold. Thus, the weight functions (6.5) with the indicated multi-index $\vec{n}$ generate a MOP ensemble (3.8) which coincides with the joint distribution of the eigenvalues of M_1 (after averaging over those of M_2). Also, the MOPs are characterized by a Riemann–Hilbert problem of size $(p+1)\times(p+1)$ as in § 3.4. This Riemann–Hilbert problem was given in a different, but equivalent, form in [88]. The correlation kernel $K_{11}^{(n)}$ is expressed as before in terms of the solution of the Riemann–Hilbert problem by the formula (3.26).

6.2. The two-matrix model with quartic potential. If W is a quadratic polynomial, then the kernel $K_{11}^{(n)}$ is essentially the correlation kernel (3.3) of an OP ensemble. In this case the two-matrix model reduces again to the one-matrix model (if one considers only M_1).

The case of a quartic polynomial potential W was considered in the recent papers [89]–[91]. In [89] the Deift–Zhou steepest-descent method was applied to the Riemann–Hilbert problem for the case when

$$W(y) = \frac{1}{4}\, y^4$$

while V is a general even polynomial. This made possible a precise asymptotic analysis of the kernel $K_{11}^{(n)}$ as $n \to \infty$, leading in particular to the local universality results common in random matrix theory and involving the sine and Airy kernels. The paper [89] was restricted to the genus-zero case, but this restriction was removed in [91], which deals with higher genus. The genus here refers to an underlying Riemann surface (spectral curve) determined in [89] by means of a vector equilibrium problem in a way similar to what was discussed in § 4.4 for the matrix model with an external source.

The analysis was extended in [90] to the case of the more general even quartic potential

$$W(y) = \frac{1}{4}\, y^4 + \frac{\alpha}{2}\, y^2, \tag{6.6}$$

with $\alpha \in \mathbb{R}$. Here we summarize some of the results.

The main tool in [90] is a vector equilibrium problem for three measures. It requires minimizing the energy functional

$$\sum_{j=1}^{3} \iint \log \frac{1}{|x-y|}\, d\mu_j(x)\, d\mu_j(y) - \sum_{j=1}^{2} \iint \log \frac{1}{|x-y|}\, d\mu_j(x)\, d\mu_{j+1}(y) + \int V_1(x)\, d\mu_1(x) + \int V_3(x)\, d\mu_3(x), \tag{6.7}$$

where V_1 and V_3 are certain external fields (to be discussed below), among all vectors of measures (μ_1, μ_2, μ_3) satisfying the following conditions:

(a) the measures have finite logarithmic energy;
(b) μ_1 is a measure on $\mathbb{R}$ with total mass 1;
(c) μ_2 is a measure on $i\mathbb{R}$ with total mass 2/3;

(d) μ_3 is a measure on $\mathbb{R}$ with total mass $1/3$;

(e) $\mu_2 \leqslant \sigma$, where σ is a certain measure on the imaginary axis (see also below).

The external fields V_1 and V_3 and the upper constraint σ are explicit in the case $\alpha = 0$ in (6.6). Their description in the case of general α is somewhat involved. We shall go through it here for $\alpha < 0$, which is the case of a double-well potential W. See [90] for a discussion of all cases.

Thus, assume that $\alpha < 0$. We also assume without loss of generality that $\tau > 0$. The external field V_1 acting on μ_1 is defined by

$$V_1(x) = V(x) + \min_{s\in\mathbb{R}}\big(W(s) - \tau x s\big). \tag{6.8}$$

The minimum is attained at a value $s = s_1(x) \in \mathbb{R}$ for which $W'(s) = \tau x$, that is,

$$s^3 + \alpha s = \tau x. \tag{6.9}$$

For $\alpha < 0$ there can be more than one real solution of (6.9). The relevant value is the one that has the same sign as x (since $\tau > 0$). It is uniquely determined, except for the case $x = 0$.

For $\alpha < 0$ the external field $V_3(x)$ vanishes identically outside of the interval $\big(-x^*(\alpha), x^*(\alpha)\big)$, where

$$x^*(\alpha) = \frac{2}{\tau}\left(\frac{-\alpha}{3}\right)^{3/2}. \tag{6.10}$$

For $x \in \big(-x^*(\alpha), x^*(\alpha)\big)$ the equation (6.9) has three real solutions $s_1 = s_1(x)$, $s_2 = s_2(x)$, $s_3 = s_3(x)$, which we assume to be ordered so that

$$W(s_1) - \tau x s_1 \leqslant W(s_2) - \tau x s_2 \leqslant W(s_3) - \tau x s_3.$$

Then the external field V_3 acting on μ_3 is defined by

$$V_3(x) = \big(W(s_3(x)) - \tau x s_3(x)\big) - \big(W(s_2(x)) - \tau x s_2(x)\big) \quad \text{for } x \in \big(-x^*(\alpha), x^*(\alpha)\big), \tag{6.11}$$

and $V_3(x) = 0$ elsewhere. Thus, $V_3(x)$ is the difference between the local maximum and the other local minimum of the map $s \mapsto W(s) = \tau x s$, which minimum actually exists if and only if $x \in (-x^*(\alpha), x^*(\alpha))$.

The measure σ acting as a constraint on μ_2 is supported on the whole imaginary axis if $\alpha < 0$. In this case the saddle point equation

$$s^3 + \alpha s = \tau z, \quad \text{with } z \in i\mathbb{R}, \tag{6.12}$$

has one solution on the imaginary axis, one solution $s(z)$ in the right half-plane, and its reflection $-\overline{s(z)}$ in the imaginary axis. The measure σ is then defined by

$$\frac{d\sigma(z)}{|dz|} = \frac{\tau}{\pi}\operatorname{Re} s(z), \qquad z \in i\mathbb{R}, \tag{6.13}$$

where $s(z)$ is the solution of (6.12) with positive real part.

The following result from [90] is analogous to Theorem 4.2.

Theorem 6.1. *There is a unique minimizer $(\mu_1^*, \mu_2^*, \mu_3^*)$ in the above vector equilibrium problem, and it satisfies the following conditions.*

(i) *The support of μ_1^* is bounded and consists of a finite union of intervals*:

$$S(\mu_1^*) = \bigcup_{j=1}^{N} [a_j, b_j]. \tag{6.14}$$

(ii) *The support of μ_2^* coincides with the support of σ, and there exists a $c_2 \geqslant 0$ such that*

$$S(\sigma - \mu_2^*) = (-i\infty, -ic_2] \cup [ic_2, i\infty). \tag{6.15}$$

(iii) *There exists a constant $c_3 \geqslant 0$ such that*

$$S(\mu_3^*) = (-\infty, -c_3] \cup [c_3, \infty). \tag{6.16}$$

(iv) *μ_1^* has a density ρ with respect to Lebesgue measure. In regular cases ρ is the limit mean density of the eigenvalues of the random matrix M_1 in the 2-matrix model as $n \to \infty$, that is,*

$$\rho(x) = \lim_{n\to\infty} \frac{1}{n} K_{11}^{(n)}(x, x).$$

The regular cases are the same as in part (iii) of Theorem 4.2. In a regular case $S(\mu_1^*)$ and $S(\sigma - \mu_2^*)$ are disjoint, as are $S(\sigma - \mu_2^*)$ and $S(\mu_3^*)$. These sets are then used as cuts to define a Riemann surface $\mathfrak{R}$ with four sheets

$$\begin{aligned} \mathfrak{R}_1 &= \overline{\mathbb{C}} \setminus S(\mu_1^*), \\ \mathfrak{R}_2 &= \overline{\mathbb{C}} \setminus \big(S(\mu_1^*) \cup S(\sigma - \mu_2^*)\big), \\ \mathfrak{R}_3 &= \overline{\mathbb{C}} \setminus \big(S(\sigma - \mu_2^*) \cup S(\mu_3^*)\big), \\ \mathfrak{R}_4 &= \overline{\mathbb{C}} \setminus S(\mu_3^*) \end{aligned} \tag{6.17}$$

that are glued together in the usual crosswise manner. Then as was the case for the external source model (see (4.32)), we see that the function

$$\xi_1(z) = V'(z) - \int \frac{d\mu_1^*(s)}{z - s}, \qquad z \in \mathfrak{R}_1, \tag{6.18}$$

has a meromorphic continuation to the whole Riemann surface $\mathfrak{R}$, with poles only at the points at infinity. In fact, there is a simple pole at the point at infinity that is common to the second, third, and fourth sheets, which was not the case in (4.32).

These facts are the main ingredients for the steepest-descent analysis of the 4×4 matrix-valued Riemann–Hilbert problem.

Finally, we discuss the phase diagram for the case when $V(x) = x^2/2$ and $W(y) = y^4/4 + \alpha y^2/2$ (see [84], [90]). In this case the model depends on the two parameters $\alpha \in \mathbb{R}$ and $\tau > 0$. Cases with either $S(\mu_1^*) \cap S(\sigma - \mu_2^*) \neq \varnothing$ or $S(\mu_3^*) \cap S(\sigma - \mu_2^*) \neq \varnothing$ are not regular. These cases correspond to changes in the sheet structure of the Riemann surface, and in some cases to a change in its genus.

The four regions correspond to the following cases:

Case I: $0 \in S(\mu_1^*)$, $0 \notin S(\sigma - \mu_2^*)$, and $0 \in S(\mu_3^*)$.
Case II: $0 \notin S(\mu_1^*)$, $0 \notin S(\sigma - \mu_2^*)$, and $0 \in S(\mu_3^*)$.
Case III: $0 \notin S(\mu_1^*)$, $0 \in S(\sigma - \mu_2^*)$, and $0 \notin S(\mu_3^*)$.
Case IV: $0 \in S(\mu_1^*)$, $0 \notin S(\sigma - \mu_2^*)$, and $0 \notin S(\mu_3^*)$.

There is also a fifth possible case, which, however, does not occur if $V(x) = x^2/2$:

Case V: $0 \notin S(\mu_1^*)$, $0 \notin S(\sigma - \mu_2^*)$, and $0 \notin S(\mu_3^*)$.

The transitions from one case to another correspond to the opening or closing of a gap in one (or more) of the supports. The transitions are generically of the Pearcey or Painlevé II type (see [90]). All four cases come together at the special point $(\alpha, \tau) = (-1, 1)$. This is a multi-critical case in which new local behaviour can be expected, and it is currently being investigated.

Bibliography

[1] Ch. Hermite, "Sur la fonction exponentielle", *C. R. Acad. Sci. Paris* **77** (1873), 18–24, 74–79, 226–233, 285–293.

[2] А. А. Марков, "Два доказательства сходимости некоторых непрерывных дробей", *Избранные труды по теории непрерывных дробей и теории функций, наименее уклоняющихся от нуля*, Гостехиздат, М. 1948, с. 106–119; French original, A. Markoff, "Deux démonstrations de la convergence de certaines fractions continues", *Acta Math.* **19**:1 (1895), 93–104.

[3] В. Н. Сорокин, "Циклические графы и теорема Апери", *УМН* **57**:3 (2002), 99–134; English transl., V. N. Sorokin, "Cyclic graphs and Apéry's theorem", *Russian Math. Surveys* **57**:3 (2002), 535–571.

[4] W. Van Assche, "Multiple orthogonal polynomials, irrationality and transcendence", *Continued fractions: from analytic number theory to constructive approximation* (University of Missouri–Columbia, Columbia, MO, USA, 1998), Contemp. Math., vol. 236, Amer. Math. Soc., Providence, RI 1999, pp. 325–342.

[5] А. И. Аптекарев (ред.), *Рациональные приближения постоянной Эйлера и рекуррентные соотношения*, Сборник статей, Совр. пробл. матем., **9**, МИАН, М. 2007, 84 с. [A. I. Aptekatev (ed.), *Rational approximations of the Euler constant and recurrence relations*, Sovrem. Probl. Mat., vol. 9, Steklov Mathenatical Institute, Moscow 2007, 84 pp.]

[6] В. А. Калягин, "Аппроксимации Эрмита–Паде и спектральный анализ несимметричных операторов", *Матем. сб.* **185**:6 (1994), 79–100; English transl., V. A. Kalyagin, "Hermite–Padé approximants and spectral analysis of nonsymmetric operators", *Russian Acad. Sci. Sb. Math.* **82**:1 (1995), 199–216.

[7] A. Aptekarev and V. Kaliaguine (Kalyagin), "Complex rational approximation and difference operators", *Proceedings of the 3rd International Conference on Functional Analysis and Approximation Theory*, vol. I (Acquafredda di Maratea, Italy, 1996), Rend. Circ. Mat. Palermo (2) Suppl., vol. 52, Circolo Matematico di Palermo, Palermo 1998, pp. 3–21.

[8] A. I. Aptekarev, V. A. Kalyagin, and E. B. Saff, "Higher-order three-term recurrences and asymptotics of multiple orthogonal polynomials", *Constr. Approx.* **30**:2 (2009), 175–223.

[9] P. M. Bleher and A. B. J. Kuijlaars, "Random matrices with external source and multiple orthogonal polynomials", *Int. Math. Res. Not.*, 2004, no. 3, 109–129.

[10] A. B. J. Kuijlaars and K. T.-R. McLaughlin, "A Riemann–Hilbert problem for biorthogonal polynomials", *J. Comput. Appl. Math.* **178**:1-2 (2005), 313–320.

[11] A. I. Aptekarev, P. M. Bleher, and A. B. J. Kuijlaars, "Large n limit of Gaussian random matrices with external source. II", *Comm. Math. Phys.* **259**:2 (2005), 367–389.

[12] А. И. Аптекарев, В. Г. Лысов, Д. Н. Туляков, "Глобальный режим распределения собственных значений случайных матриц с ангармоническим потенциалом и внешним источником", *ТМФ* **159**:1 (2009), 34–57; English transl., A. I. Aptekarev, V. G. Lysov, and D. N. Tulyakov, "Global eigenvalue distribution regime of random matrices with an anharmonic potential and an external source", *Theoret. Math. Phys.* **159**:1 (2009), 448–468.

[13] А. И. Аптекарев, В. Г. Лысов, Д. Н. Туляков, "Случайные матрицы с внешним источником и асимптотика совместно ортогональных многочленов", *Матем. сб.* **202**:2 (2011), 3–56; English transl., A. I. Aptekarev, V. G. Lysov, and D. N. Tulyakov, "Random matrices with external source and the asymptotic behaviour of multiple orthogonal polynomials", *Sb. Math.* **202**:2 (2011), 155–206.

[14] P. Bleher, S. Delvaux, and A. B. J. Kuijlaars, "Random matrix model with external source and a constrained vector equilibrium problem", *Comm. Pure Appl. Math.* **64** (2011), 116–160.

[15] S. Delvaux, "Average characteristic polynomials for multiple orthogonal polynomial ensembles", *J. Approx. Theory* **162**:5 (2010), 1033–1067.

[16] А. А. Гончар, Е. А. Рахманов, В. Н. Сорокин, "Об аппроксимациях Эрмита–Паде для систем функций марковского типа", *Матем. сб.* **188**:5 (1997), 33–58; English transl., A. A. Gonchar, E. A. Rakhmanov, and V. N. Sorokin, "Hermite–Padé approximants for systems of Markov-type functions", *Sb. Math.* **188**:5 (1997), 671–696.

[17] А. И. Аптекарев, В. Г. Лысов, "Системы марковских функций, генерируемые графами, и асимптотика их аппроксимаций Эрмита–Паде", *Матем. сб.* **201**:2 (2010), 29–78; English transl., A. I. Aptekarev and V. G. Lysov, "Systems of Markov functions generated by graphs and the asymptotics of their Hermite–Padé approximants", *Sb. Math.* **201**:2 (2010), 183–234.

[18] A. I. Aptekarev, A. B. J. Kuijlaars, and W. Van Assche, "Asymptotics of Hermite–Padé rational approximants for two analytic functions with separated pairs of branch points (case of genus 0)", *Int. Math. Res. Pap. IMRP*, 2008, Art. ID rpm007, 128 pp.

[19] W. Van Assche, J. S. Geronimo, and A. B. J. Kuijlaars, "Riemann–Hilbert problems for multiple orthogonal polynomials", *Special Functions* 2000: *Current Perspective and Future Directions* (Tempe, AZ, USA, 2000), NATO Sci. Ser. II Math. Phys. Chem., vol. 30, Kluwer Acad. Publ., Dordrecht 2001, pp. 23–59.

[20] E. Daems and A. B. J. Kuijlaars, "A Christoffel–Darboux formula for multiple orthogonal polynomials", *J. Approx. Theory* **130**:2 (2004), 190–202; arXiv: math/0402031.

[21] P. M. Bleher and A. B. J. Kuijlaars, "Large n limit of Gaussian random matrices with external source. I", *Comm. Math. Phys.* **252**:1-3 (2004), 43–76.

[22] P. M. Bleher and A. B. J. Kuijlaars, "Large n limit of Gaussian random matrices with external source. III. Double scaling limit", *Comm. Math. Phys.* **270** (2007), 481–517.

[23] А. А. Гончар, Е. А. Рахманов, "О сходимости совместных аппроксимаций Паде для систем функций марковского типа", *Теория чисел, математический анализ и их приложения*, Сборник статей. Посвящается академику Ивану Матвеевичу Виноградову к его девяностолетию, Тр. МИАН СССР, **157**, 1981,

c. 31–48; English transl., A. A. Gonchar and E. A. Rakhmanov, "On convergence of simultaneous Padé approximants for systems of functions of Markov type", *Proc. Steklov Inst. Math.* **157** (1983), 31–50.

[24] А. А. Гончар, Е. А. Рахманов, "О задаче равновесия для векторных потенциалов", *УМН* **40**:4(244) (1985), 155–156; English transl., A. A. Gonchar and E. A. Rakhmanov, "On the equilibrium problem for vector potentials", *Russian Math. Surveys* **40**:4 (1985), 183–184.

[25] A. Angelesco, "Sur deux extensions des fractions continues algébriques", *C. R. Acad. Sci. Paris* **168** (1919), 262–265.

[26] Е. М. Никишин, "О совместных аппроксимациях Паде", *Матем. сб.* **113**:4 (1980), 499–519; English transl., E. M. Nikišin (Nikishin), "On simultaneous Padé approximants", *Math. USSR-Sb.* **41**:4 (1982), 409–425.

[27] K. Driver and H. Stahl, "Simultaneous rational approximants to Nikishin-systems. I, II", *Acta Sci. Math.* (*Szeged*) **60**:1-2 (1995), 245–263; **61**:1-4 (1995), 261–284.

[28] U. Fidalgo Prieto and G. López Lagomasino, "Nikishin systems are perfect", *Constr. Approx.* **34**:3 (2011), 297–356; arXiv: 1001.0554v1.

[29] А. И. Аптекарев, "Асимптотика полиномов совместной ортогональности в случае Анджелеско", *Матем. сб.* **136(178)**:1(5) (1988), 56–84; English transl., A. I. Aptekarev, "Asymptotics of simultaneously orthogonal polynomials in the Angelesco case", *Math. USSR-Sb.* **64**:1 (1989), 57–84.

[30] Ж. Бустаманте, Г. Лопес Лагомасино, "Аппроксимации Эрмита–Паде для систем Никишина аналитических функций", *Матем. сб.* **183**:11 (1992), 117–138; English transl., Zh. Bustamante and G. L. Lagomasino, "Hermite–Padé approximation for Nikishin systems of analytic functions", *Russian Acad. Sci. Sb. Math.* **77**:2 (1994), 367–384.

[31] Е. М. Никишин, "Об асимптотике линейных форм для совместных аппроксимаций Паде", *Изв. вузов. Матем.*, 1986, no. 2, 33–41; English transl., E. M. Nikishin, "The asymptotic behavior of linear forms for joint Padé approximations", *Soviet Math.* (*Iz. VUZ*) **30**:2 (1986), 43–52.

[32] А. И. Аптекарев, "Сильная асимптотика многочленов совместной ортогональности для систем Никишина", *Матем. сб.* **190**:5 (1999), 3–44; English transl., A. I. Aptekarev, "Strong asymptotics of multiply orthogonal polynomials for Nikishin systems", *Sb. Math.* **190**:5 (1999), 631–669.

[33] А. И. Аптекарев, "Асимптотика аппроксимаций Эрмита–Паде для пары функций с точками ветвления", *Докл. РАН* **422**:4 (2008), 443–445; English transl., A. I. Aptekarev, "Asymptotics of Hermite–Padé approximants for two functions with branch points", *Dokl. Math.* **78**:2 (2008), 717–719.

[34] Е. М. Никишин, В. Н. Сорокин, *Рациональные приближения и ортогональность*, Наука, М. 1988, 256 с.; English transl., E. M. Nikishin and V. N. Sorokin, *Rational approximations and orthogonality*, Transl. Math. Monogr., vol. 92, Amer. Math. Soc., Providence, RI 1991, viii+221 pp.

[35] L. C. Young, *Lectures on the calculus of variations and optimal control theory*, Saunders, Philadelphia–London–Toronto 1969, xi+331 pp.

[36] Е. А. Рахманов, "К асимптотике многочленов Эрмита–Паде для двух марковских функций", *Матем. сб.* **202**:1 (2011), 133–140; English transl., E. A. Rakhmanov, "The asymptotics of Hermite–Padé polynomials for two Markov-type functions", *Sb. Math.* **202**:1 (2011), 127–134.

[37] А. А. Гончар, Е. А. Рахманов, "Равновесная мера и распределение нулей экстремальных многочленов", *Матем. сб.* **125(167)**:1(9) (1984), 117–127; English transl., A. A. Gonchar and E. A. Rakhmanov, "Equilibrium measure and

the distribution of zeros of extremal polynomials", *Math. USSR-Sb.* **53**:1 (1986), 119–130.

[38] R. T. Baumel, J. L. Gammel, and J. Nuttall, "Asymptotic form of Hermite–Padé polynomials", *IMA J. Appl. Math.* **27**:3 (1981), 335–357.

[39] J. Nuttall, "Hermite–Padé approximants to functions meromorphic on a Riemann surface", *J. Approx. Theory* **32**:3 (1981), 233–240.

[40] J. Nuttall, "Asymptotics of diagonal Hermite–Padé polynomials", *J. Approx. Theory* **42**:4 (1984), 299–386.

[41] C. L. Siegel, *Topics in complex function theory*, vol. I: *Elliptic functions and uniformization theory*, Wiley, New York–London–Sydney 1969, ix+186 pp.; vol. II: *Automorphic functions and Abelian integrals*, Wiley, New York 1971, xii+193 pp.

[42] A. I. Aptekarev, V. A. Kalyagin, V. G. Lysov, and D. N. Toulyakov (Tulyakov), "Equilibrium of vector potentials and uniformization of the algebraic curves of genus 0", *J. Comput. Appl. Math.* **233**:3 (2009), 602–616.

[43] M. L. Mehta and M. Gaudin, "On the density of eigenvalues of a random matrix", *Nuclear Phys.* **18** (1960), 420–427.

[44] M. L. Mehta, *Random matrices*, 1st ed., Academic Press, New York 1967, x+259 pp.; 3rd ed., Pure Appl. Math. (Amst.), vol. 142, Elsevier/Academic Press, Amsterdam 2004, xviii+688 pp.

[45] P. Deift, *Orthogonal polynomials and random matrices: a Riemann–Hilbert approach*, Courant Lect. Notes Math., vol. 3, New York University, Courant Institute of Mathematical Sciences, New York; Amer. Math. Soc., Providence, RI 1999, viii+273 pp.

[46] P. Deift and X. Zhou, "A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the MKdV equation", *Ann. of Math.* (2) **137**:2 (1993), 295–368.

[47] В. Г. Лысов, "Сильная асимптотика аппроксимаций Эрмита–Паде для системы стилтьесовских функций с весом Лагерра", *Матем. сб.* **196**:12 (2005), 99–122; English transl., V. G. Lysov, "Strong asymptotics of the Hermite–Padé approximants for a system of Stieltjes functions with Laguerre weight", *Sb. Math.* **196**:12 (2005), 1815–1840.

[48] P. Deift, T. Kriecherbauer, and K. T.-R. McLaughlin, "New results on the equilibrium measure for logarithmic potentials in the presence of an external field", *J. Approx. Theory* **95**:3 (1998), 388–475.

[49] 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", *Comm. Pure Appl. Math.* **52**:11 (1999), 1335–1425.

[50] A. B. J. Kuijlaars, K. T.-R. McLaughlin, W. Van Assche, and M. Vanlessen, "The Riemann–Hilbert approach to strong asymptotics for orthogonal polynomials on $[-1, 1]$", *Adv. Math.* **188**:2 (2004), 337–398.

[51] А. И. Аптекарев, "Точные константы рациональных аппроксимаций аналитических функций", *Мат. сб.* **193**:1 (2002), 3–72; English transl., A. I. Aptekarev, "Sharp constants for rational approximations of analytic functions", *Sb. Math.* **193**:1 (2002), 1–72.

[52] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*, Cambridge Stud. Adv. Math., vol. 118, Cambridge University Press, Cambridge 2010, xiv+492 pp.

[53] P. Forrester, *Log-gases and random matrices*, London Math. Soc. Monogr. Ser., vol. 34, Princeton University Press, Princeton, NJ 2010, xiv+791 pp.

[54] G. Akemann, J. Baik, and Ph. Di Francesco (eds.), *The Oxford handbook of random matrix theory*, Oxford Univesity Press, Oxford 2011, xxxi+919 pp.

[55] A. B. J. Kuijlaars, "Multiple orthogonal polynomial ensembles", *Recent trends in orthogonal polynomials and approximation theory*, Contemp. Math., vol. 507 (J. Arvesú, F. Marcellán, A. Martínez-Finkelshtein, eds.), Amer. Math. Soc., Providence, RI 2010, pp. 155–176.

[56] P. Deift and D. Gioev, *Random matrix theory*: *invariant ensembles and universality*, Courant Lect. Notes Math., vol. 18, Amer. Math. Soc., Providence, RI 2009, x+217 pp.

[57] A. Borodin, "Biorthogonal ensembles", *Nuclear Phys. B* **536**:3 (1999), 704–732.

[58] K. Johansson, "Random matrices and determinantal processes", *Mathematical statistical physics*, Elsevier, Amsterdam 2006, pp. 1–55.

[59] P. Desrosiers and P. J. Forrester, "A note on biorthogonal ensembles", *J. Approx. Theory* **152**:2 (2008), 167–187.

[60] W. Van Assche, "Nearest neighbor recurrence relations for multiple orthogonal polynomials", *J. Approx. Theory* **163**:10 (2011), 1427–1448; arXiv: 1104.3778.

[61] C. Álvarez-Fernández, U. Fidalgo Prieto, and M. Mañas, "Multiple orthogonal polynomials of mixed type: Gauss–Borel factorization and the multi-component 2D Toda hierarchy", *Adv. Math.* **227**:4 (2011), 1451–1525.

[62] E. Daems and A. B. J. Kuijlaars, "Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions", *J. Approx. Theory* **146**:1 (2007), 91–114.

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

[64] E. Brézin and S. Hikami, "Universal singularity at the closure of a gap in a random matrix theory", *Phys. Rev. E* (3) **57**:4 (1998), 4140–4149.

[65] E. Brézin and S. Hikami, "Level spacing of random matrices in an external source", *Phys. Rev. E* (3) **58**:6, Part A (1998), 7176–7185.

[66] P. Zinn-Justin, "Random Hermitian matrices in an external field", *Nuclear Phys. B* **497** (1997), 725–732.

[67] C. Itzykson and J. B. Zuber, "The planar approximation. II", *J. Math. Phys.* **21**:3 (1980), 411–421.

[68] E. B. Saff and V. Totik, *Logarithmic potentials with external fields*, Grundlehren Math. Wiss., vol. 316, Springer-Verlag, Berlin 1997, xvi+505 pp.

[69] A. I. Aptekarev, A. Branquinho, and W. Van Assche, "Multiple orthogonal polynomials for classical weights", *Trans. Amer. Math. Soc.* **355**:10 (2003), 3887–3914.

[70] Л. А. Пастур, "О спектре случайных матриц", *ТМФ* **10**:1 (1972), 102–112; English transl., L. A. Pastur, "On the spectrum of random matrices", *Theoret. Math. Phys.* **10**:1 (1972), 67–74.

[71] C. A. Tracy and H. Widom, "The Pearcey process", *Comm. Math. Phys.* **263**:2 (2006), 381–400.

[72] K. T.-R. McLaughlin, "Asymptotic analysis of random matrices with external source and a family of algebraic curves", *Nonlinearity* **20**:7 (2007), 1547–1571.

[73] Е. А. Рахманов, "Равновесная мера и распределение нулей экстремальных многочленов дискретной переменной", *Матем. сб.* **187**:8 (1996), 109–124; English transl., E. A. Rakhmanov, "Equilibrium measure and the distribution of zeros of the extremal polynomials of a discrete variable", *Sb. Math.* **187**:8 (1996), 1213–1228.

[74] P. D. Dragnev and E. B. Saff, "Constrained energy problems with applications to orthogonal polynomials of a discrete variable", *J. Anal. Math.* **72**:1 (1997), 223–259.

[75] J. Baik, T. Kriecherbauer, K. T.-R. McLaughlin, and P. D. Miller, *Discrete orthogonal polynomials. Asymptotics and applications*, Ann. of Math. Stud., vol. 164, Princeton University Press, Princeton, NJ 2007, viii+170 pp.

[76] P. Bleher and A. Its, "Double scaling limit in the random matrix model: the Riemann–Hilbert approach", *Comm. Pure Appl. Math.* **56**:4 (2003), 433–516.

[77] T. Claeys and A. B. J. Kuijlaars, "Universality of the double scaling limit in random matrix models", *Comm. Pure Appl. Math.* **59**:11 (2006), 1573–1603.

[78] S. Karlin and J. McGregor, "Coincidence probabilities", *Pacific J. Math.* **9**:4 (1959), 1141–1164; http://projecteuclid.org/euclid.pjm/1103038889.

[79] A. B. J. Kuijlaars, A. Martínez-Finkelshtein, and F. Wielonsky, "Non-intersecting squared Bessel paths and multiple orthogonal polynomials for modified Bessel weights", *Comm. Math. Phys.* **286**:1 (2009), 217–275.

[80] A. B. J. Kuijlaars, A. Martínez-Finkelshtein, and F. Wielonsky, "Non-intersecting squared Bessel paths: critical time and double scaling limit", *Comm. Math. Phys.* **308** (2011), 227–279; arXiv: 1011.1278.

[81] S. Delvaux, A. B. J. Kuijlaars, P. Román, and L. Zhang, "Non-intersecting squared Bessel paths with one positive starting and ending point", *J. Anal. Math.*, 2011 (to appear); arXiv: 1105.2481.

[82] M. L. Mehta, "A method of integration over matrix variables", *Comm. Math. Phys.* **79**:3 (1981), 327–340.

[83] 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** (2001), 232–268.

[84] M. Duits, D. Geudens, and A. B. J. Kuijlaars, "A vector equilibrium problem for the two-matrix model in the quartic/quadratic case", *Nonlinearity* **24**:3 (2011), 951–993.

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

[86] M. L. Mehta and P. Shukla, "Two coupled matrices: eigenvalue correlations and spacing functions", *J. Phys. A* **27**:23 (1994), 7793–7803.

[87] S. Delvaux, "Average characteristic polynomials in the two-matrix model", *J. Math. Phys.* **52**:1 (2011), 013513, 23 pp.

[88] 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 (2003), 193–240.

[89] M. Duits and A. B. J. Kuijlaars, "Universality in the two matrix model: a Riemann–Hilbert steepest descent analysis", *Comm. Pure Appl. Math.* **62**:8 (2009), 1076–1153.

[90] M. Duits, A. B. J. Kuijlaars, and M. Y. Mo, *The Hermitian two matrix model with an even quartic potential*, Mem. Amer. Math. Soc., Amer. Math. Soc., Providence, RI 2011, v+105 pp.

[91] M. Y. Mo, "Universality in the two matrix model with a monomial quartic and a general even polynomial potential", *Comm. Math. Phys.* **291**:3 (2009), 863–894.

A. I. Aptekarev
Keldysh Institute of Applied Mathematics,
Russian Academy of Sciences
E-mail: aptekaa@keldysh.ru

Received 15/SEP/11

A. B. J. Kuijlaars
Katholieke Universiteit Leuven, Belgium
E-mail: arno.kuijlaars@wis.kuleuven.be