



Maurice Duits · Arno B. J. Kuijlaars

# The two-periodic Aztec diamond and matrix valued orthogonal polynomials

Received February 23, 2018

**Abstract.** We analyze domino tilings of the two-periodic Aztec diamond by means of matrix valued orthogonal polynomials that we obtain from a reformulation of the Aztec diamond as a non-intersecting path model with periodic transition matrices. In a more general framework we express the correlation kernel for the underlying determinantal point process as a double contour integral that contains the reproducing kernel of matrix valued orthogonal polynomials. We use the Riemann–Hilbert problem to simplify this formula for the case of the two-periodic Aztec diamond.

In the large size limit we recover the three phases of the model known as solid, liquid and gas. We describe the fine asymptotics for the gas phase and at the cusp points of the liquid-gas boundary, thereby complementing and extending results of Chhita and Johansson.

**Keywords.** Aztec diamond, random tilings, matrix valued orthogonal polynomials, Riemann–Hilbert problems

---

## 1. Introduction

We study domino tilings of the Aztec diamond with a two-periodic weighting. This model falls into a class of models for which existing techniques for studying fine asymptotics are not adequate and only recently first important progress has been made [9, 20]. We introduce a new approach based on matrix valued orthogonal polynomials that allows us to compute the determinantal correlations at finite size and their asymptotics as the size of the diamond gets large in a rather orderly way. We strongly believe that this approach will also prove to be a good starting point for other tiling models with a periodic weighting.

Random tilings of planar domains have been studied intensively in the past decade. Such models exhibit a rich structure including a limit shape and fluctuations that are expected to fall in various important universality classes (see [8, 22, 23, 44, 46, 47, 48, 49] for general references on the topic and [18, 59, 60] for recent contributions). When the correlation structure is determinantal, there is hope to understand the fine asymptotic structure by studying the asymptotic behavior of the correlation kernel. An important

---

M. Duits: Department of Mathematics, Royal Institute of Technology (KTH), Lindstedsvägen 25, 10044 Stockholm, Sweden; e-mail: [duits@kth.se](mailto:duits@kth.se)

A. B. J. Kuijlaars: Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200 B, 3001 Leuven, Belgium; e-mail: [arno.kuijlaars@kuleuven.be](mailto:arno.kuijlaars@kuleuven.be)

*Mathematics Subject Classification (2020):* Primary 60D05; Secondary 33C47, 82B26

source of examples of such models is the Schur process [57]. For these models the correlation kernel can be explicitly computed in terms of a double integral representation, opening up to the possibility of performing an asymptotic analysis by means of classical steepest descent (or stationary phase) techniques.

Of course, the Schur process is rather special and many models of interest fall outside this class. In particular this is true for random tilings or dimer models with doubly periodic weightings. Yet, these models have exciting new features and have therefore been discussed in the physics and mathematics literature [9, 20, 21, 30, 49, 56]. An important feature is the appearance of a so-called gas phase. For instance, in the two-periodic weighting for domino tiling of the Aztec diamond, the diamond can be partitioned into three regions: the solid, liquid and gas region [57] (as we will see in Figure 6 below). The gas region has not been observed in models that are in the Schur class. The 2-point correlations (for an associated particle process) in the gas region behave differently when compared to the liquid regions. Indeed, the correlation kernel decays exponentially with the distance  $d$  between the points, compared to  $\sim 1/d$  in the liquid region. At the liquid-solid boundary one expects the Airy process to appear, but the situation at the gas-liquid boundary is far more complicated [9, 20].

To the best of our knowledge, the two-periodic Aztec diamond is the only model with periodic weightings for which rigorous results on fine asymptotics exist [9, 20]. Inspired by a formula for the Kasteleyn matrix found by Chhita and Young [21], Chhita and Johansson [20] found a way to compute the asymptotic behavior of the Kasteleyn matrix as the size of the Aztec diamond goes to infinity. We will follow a different approach to studying such models with periodic weightings.

As we will recall in Section 3, the Aztec diamond can be described by non-intersecting paths (we refer to [46] and the references therein for more background on the relation between dimers, tilings, non-intersecting paths and all that). For a general class of discrete non-intersecting paths with  $p$ -periodic transition matrices (which includes  $p$ -periodic weightings for domino tilings of the Aztec diamond and  $p$ -periodic weightings for lozenge tilings of the hexagon), we show in Section 4 how the correlation kernel can be written as a double integral formula involving matrix valued polynomials that satisfy a non-hermitian orthogonality.

We believe that this general setup has a high potential for a rigorous asymptotic analysis. The key fact is that these matrix valued orthogonal polynomials can be characterized in terms of the solution of a  $2p \times 2p$  matrix valued Riemann–Hilbert problem. With the highly developed Riemann–Hilbert toolkit at hand, we may thus hope to compute the asymptotic behavior of the polynomials, and more importantly the correlation kernel. The formalism will be worked out in Section 4. It provides a new perspective even on the classical examples of uniform domino tilings of the Aztec diamond and lozenge tilings of a hexagon.

The main focus of the paper is to show how the Riemann–Hilbert approach can be exploited for an asymptotic analysis of the two-periodic Aztec diamond. Remarkably, in this case the result of the Riemann–Hilbert analysis is a surprisingly simple double integral formula for the correlation kernel. It is not an asymptotic result, but an exact formula valid for fixed finite  $N$ . This representation also appears to be more elementary

than the one given by Chhita and Johansson [20]. The Riemann–Hilbert analysis is given in Section 5. We analyze the double integral formula for the kernel asymptotically using classical steepest descent techniques in Section 6.

The model of the two-periodic Aztec diamond is explained and the main results are summarized in the next section.

## 2. Statement of results

In this section we introduce the two-periodic Aztec diamond and state our main results.

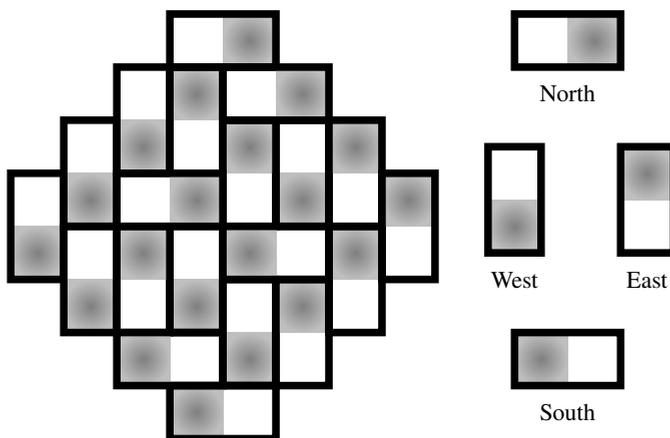
### 2.1. Definition of the model

The *Aztec diamond* is a region on the square lattice with a sawtooth boundary that can be covered by  $2 \times 1$  and  $1 \times 2$  rectangles, called *dominos*. The squares have a black/white checkerboard coloring, and a possible tiling of the Aztec diamond of size 4 is shown in Figure 1. There are four types of dominos, namely North, West, East, and South, that are also shown in the figure. The Aztec diamond model was first introduced in [35].

In the two-periodic Aztec diamond we assign a weight to each domino in a tiling, depending on its shape (horizontal or vertical) and its location in the Aztec diamond. We assume the Aztec diamond is of even size.

To describe the two-periodic weighting we introduce a coordinate system where  $(0, 0)$  is at the center of the Aztec diamond. The center of a horizontal domino has coordinates  $(x, y + 1/2)$  with  $x, y \in \mathbb{Z}$ . We then say that the horizontal domino is in *column*  $x$ . The center of a vertical domino has coordinates  $(x + 1/2, y)$  with  $x, y \in \mathbb{Z}$ , and we say that the vertical domino is in *row*  $y$ . The row and column numbers run from  $-N + 1$  to  $N - 1$ , where  $2N$  is the size of the Aztec diamond.

We fix two positive numbers  $a$  and  $b$  and define the weights as follows.



**Fig. 1.** Possible tiling of a  $4 \times 4$  Aztec diamond with four kinds of dominos.

**Definition 2.1.** The *weight of a domino*  $D$  in a tiling  $\mathcal{T}$  of the Aztec diamond is

$$w(D) = \begin{cases} a & \text{if } D \text{ is a horizontal domino in an even column,} \\ b & \text{if } D \text{ is a horizontal domino in an odd column,} \\ b & \text{if } D \text{ is a vertical domino in an even row,} \\ a & \text{if } D \text{ is a vertical domino in an odd row.} \end{cases} \tag{2.1}$$

The *weight of the tiling*  $\mathcal{T}$  is

$$w(\mathcal{T}) = \prod_{D \in \mathcal{T}} w(D), \tag{2.2}$$

and the probability for  $\mathcal{T}$  is

$$\mathcal{P}(\mathcal{T}) = w(\mathcal{T})/Z_N, \tag{2.3}$$

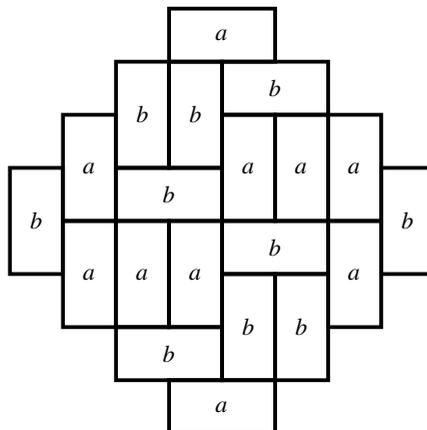
where  $Z_N = \sum_{\mathcal{T}'} w(\mathcal{T}')$  (sum over all possible tilings  $\mathcal{T}'$  of an Aztec diamond of size  $2N$ ) is the partition function.

In the example from Figure 1 the weights are shown in Figure 2. The weight of the tiling is  $w(\mathcal{T}) = a^{10}b^{10}$ .

The model is homogeneous in the sense that the probabilities (2.3) do not change if we multiply  $a$  and  $b$  by a common factor. We may and do assume  $ab = 1$ . In what follows it will be more convenient to work with

$$\alpha = a^2, \quad \beta = b^2 \tag{2.4}$$

instead of  $a$  and  $b$ . We have  $\alpha\beta = 1$ , and without loss of generality we assume  $\alpha \geq 1$ . If  $\alpha = \beta = 1$  then the model reduces to the uniform weighting on domino tilings, and so the true interest is in the case  $\alpha > 1$ , and this is what we assume from now on.



**Fig. 2.** Two-periodic weights of dominos in a tiling of the Aztec diamond. The vertical dominos in an even row and the horizontal dominos in an odd column have weight  $a$ . Other dominos have weight  $b$ .

2.2. Particle system and determinantal point process

By putting a particle in the black square of the West and South dominos, we obtain a random particle system. In our running example the particle systems is shown in Figure 3.

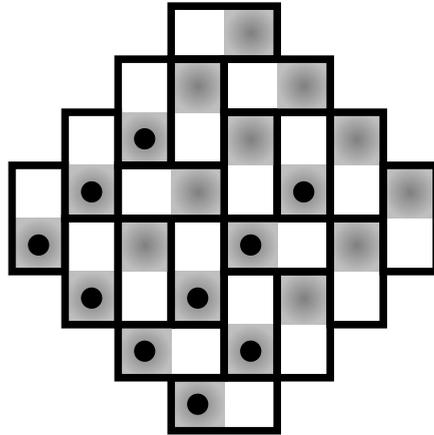


Fig. 3. Particles in a domino tiling.

We rotate the picture clockwise through 45 degrees and we change the coordinate system so that black squares are identified with the product set

$$\mathcal{B}_N = \{0, \dots, 2N\} \times \{0, \dots, 2N - 1\}. \tag{2.5}$$

Any possible tiling of the Aztec diamond gives rise to a subset  $\mathcal{X} = \mathcal{X}(\mathcal{T})$  of  $\mathcal{B}_N$  containing the squares that are occupied by a particle.

We use  $(m, n) \in \mathcal{B}_N$  to denote an element  $\mathcal{B}_N$  and we will refer to  $m$  as the *level* in  $\mathcal{B}_N$ . Any  $\mathcal{X}$  that comes from a tiling will have  $2N - m$  particles at level  $m$  for each  $m = 0, 1, \dots, 2N$ . Therefore the cardinality is  $|\mathcal{X}| = N(2N + 1)$ . There are also interlacing conditions that are satisfied when comparing the particles at level  $m$  with those at level  $m + 1$ .

The probability measure (2.3) on tilings gives rise to a probability measure on subsets  $\mathcal{X}$ , which turns out to be determinantal. This means that there exists a kernel

$$K_N : \mathcal{B}_N \times \mathcal{B}_N \rightarrow \mathbb{R} \tag{2.6}$$

with the property that for any subset  $S \subset \mathcal{B}_N$ ,

$$\mathcal{P}[S \subset \mathcal{X}] = \det [K_N(x, y)]_{x, y \in S}.$$

This is a discrete determinantal point process [14].

We found an explicit double contour integral formula for the kernel  $K_N$ . We take  $(m, n), (m', n') \in \mathcal{B}_N$ , and instead of  $K_N((m, n), (m', n'))$  we write  $K_N(m, n; m', n')$ .

We collect  $K_N(m, n; m', n')$  with some of its neighbors in a  $2 \times 2$  matrix

$$\mathbb{K}_N(m, n; m', n') = \begin{pmatrix} K_N(m, n; m', n') & K_N(m, n + 1; m', n') \\ K_N(m, n; m', n' + 1) & K_N(m, n + 1; m', n' + 1) \end{pmatrix}, \tag{2.7}$$

and this matrix appears in our formula (2.8) below.

**Theorem 2.2.** *Assume  $N$  is even and  $(m, n), (m', n') \in \mathcal{B}_N$  are such that  $m + n$  and  $m' + n'$  are even. Then*

$$\begin{aligned} \mathbb{K}_N(m, n; m', n') = & -\frac{\chi_{m>m'}}{2\pi i} \oint_{\gamma_{0,1}} A^{m-m'}(z) z^{(m'+n')/2-(m+n)/2} \frac{dz}{z} \\ & + \frac{1}{(2\pi i)^2} \oint_{\gamma_{0,1}} \frac{dz}{z} \oint_{\gamma_1} \frac{dw}{z-w} A^{N-m'}(w) F(w) A^{-N+m}(z) \\ & \times \frac{z^{N/2}(z-1)^N}{w^{N/2}(w-1)^N} \frac{w^{(m'+n')/2}}{z^{(m+n)/2}} \end{aligned} \tag{2.8}$$

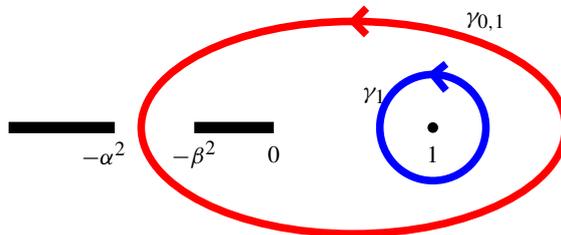
where

$$A(z) = \frac{1}{z-1} \begin{pmatrix} 2\alpha z & \alpha(z+1) \\ \beta z(z+1) & 2\beta z \end{pmatrix}, \tag{2.9}$$

$$F(z) = \frac{1}{2} I_2 + \frac{1}{2\sqrt{z(z+\alpha^2)(z+\beta^2)}} \begin{pmatrix} (\alpha-\beta)z & \alpha(z+1) \\ \beta z(z+1) & -(\alpha-\beta)z \end{pmatrix}, \tag{2.10}$$

where  $I_2$  denotes the  $2 \times 2$  identity matrix and we use the principal branch of the square root in (2.10). The notation  $\chi$  in (2.8) denotes the indicator function,  $\chi_{m>m'} = 1$  if  $m > m'$  and  $\chi_{m>m'} = 0$  otherwise.

The contour  $\gamma_{0,1}$  in (2.8) is a simple closed contour going around 0 and 1 in the positive direction. The contour  $\gamma_1$  is a simple closed contour in the right half-plane that goes around 1 in the positive direction, and it lies in the interior of  $\gamma_{0,1}$  (see Figure 4).



**Fig. 4.** Contours  $\gamma_{0,1}$  and  $\gamma_1$  used in the definition of  $\mathbb{K}_N$  in (2.8) in Theorem 2.2.

**Remark 2.3.** The square root factor in (2.10) is defined and analytic for complex  $z$  outside  $(-\infty, -\alpha^2] \cup [-\beta^2, 0]$  with the branch that is positive for real  $z > 0$ . This is one sheet of the Riemann surface  $\mathcal{R}$  associated with the cubic equation

$$y^2 = z(z + \alpha^2)(z + \beta^2), \tag{2.11}$$

which will play an important role in what follows. It is a two-sheeted surface consisting of two sheets  $\mathbb{C} \setminus ((-\infty, -\alpha^2] \cup [-\beta^2, 0])$  glued together along the two cuts  $(-\infty, -\alpha^2]$

and  $[-\beta^2, 0]$  in the usual crosswise manner. The surface has genus 1 unless  $\alpha = \beta = 1$  in which case the genus drops to 0.

The matrix valued function  $F(z)$  from (2.10) is considered on the first sheet. Its analytic continuation to the second sheet is given by  $I_2 - F(z)$ .

**Remark 2.4.** The eigenvalues of  $A(z)$  (see (2.9)), are equal to  $\frac{\rho_{1,2}(z)}{z-1}$  where

$$\rho_{1,2}(z) = (\alpha + \beta)z \pm \sqrt{z(z + \alpha^2)(z + \beta^2)} \tag{2.12}$$

are the eigenvalues of  $\begin{pmatrix} 2\alpha z & \alpha(z+1) \\ \beta z(z+1) & 2\beta z \end{pmatrix}$ . Thus (2.12) are the two branches of the meromorphic function  $\rho = (\alpha + \beta)z + y$  on the Riemann surface  $\mathcal{R}$  associated with (2.11), with  $\rho_1$  on the first sheet and  $\rho_2$  on the second sheet.

The eigenvectors can also be considered on the Riemann surface. There is a matrix  $E(z)$  whose columns are the eigenvectors such that

$$A(z) = \frac{1}{z-1} E(z) \begin{pmatrix} \rho_1(z) & 0 \\ 0 & \rho_2(z) \end{pmatrix} E^{-1}(z). \tag{2.13}$$

See (5.6) below for the precise formula for  $E(z)$ . It turns out that (see (2.10) for the definition of  $F(z)$ )

$$F(z) = E(z) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} E^{-1}(z). \tag{2.14}$$

Thus  $F(z)$  has eigenvalues 0 and 1, and  $F(z)$  commutes with  $A(z)$ .

We will also work with

$$W(z) = A^2(z)/z, \tag{2.15}$$

which in view of (2.13) has eigenvalue decomposition

$$W(z) = E(z)\Lambda(z)E^{-1}(z), \quad \Lambda(z) = \begin{pmatrix} \lambda_1(z) & 0 \\ 0 & \lambda_2(z) \end{pmatrix}, \tag{2.16}$$

with

$$\lambda_{1,2}(z) = \frac{\rho_{1,2}^2(z)}{z(z-1)^2}. \tag{2.17}$$

These eigenvalues are the two branches of a meromorphic function  $\lambda$  on the Riemann surface  $\mathcal{R}$ , with  $\lambda_1$  defined on the first sheet and  $\lambda_2$  on the second sheet.

### 2.3. Matrix valued orthogonal polynomials

The starting point of our approach to Theorem 2.2 is the non-intersecting path reformulation of the Aztec diamond and the Lindström–Gessel–Viennot lemma. This will be developed in Section 3. The novel ingredient in the further analysis is the use of matrix valued orthogonal polynomials (MVOP).

A matrix valued polynomial of degree  $k$  and size  $d$  is a function

$$P(z) = C_0 z^k + C_1 z^{k-1} + \dots + C_k$$

where  $C_0, \dots, C_k$  are matrices of size  $d \times d$ . Suppose  $w(z)$  is a  $d \times d$  weight matrix on a set  $\Gamma$  in the complex plane.

**Definition 2.5.** Suppose  $P_N$  is a matrix valued polynomial of degree  $N$  with an invertible leading coefficient. Then  $P_N$  is a *matrix valued orthogonal polynomial with weight matrix  $w$*  on  $\Gamma$  if

$$\int_{\Gamma} P_N(z)w(z)Q^t(z) dz = 0_d \quad (2.18)$$

for all matrix polynomials  $Q$  of degree  $\leq N - 1$ , where  $Q^t$  denotes the matrix transpose.

The integral in (2.18) is to be taken entrywise, and  $0_d$  denotes the  $d \times d$  zero matrix. We note that the order of the factors in the integrand in (2.18) is important since we are dealing with matrices.

For us the weight matrix will be  $\frac{1}{2\pi i} W^N(z)$  on the closed contour  $\gamma_1$  around 1. Thus  $d = 2$ , and  $W^N$  denotes the  $N$ th power of  $W$ , so that the weight matrix is varying with  $N$ . Recall that  $W$  is defined in (2.15), and explicitly we have

$$W(z) = \frac{1}{(z-1)^2} \begin{pmatrix} (z+1)^2 + 4\alpha^2 z & 2\alpha(\alpha+\beta)(z+1) \\ 2\beta(\alpha+\beta)z(z+1) & (z+1)^2 + 4\beta^2 z \end{pmatrix}. \quad (2.19)$$

The existing literature on MVOP mostly deals with the case of orthogonality on an interval of the real line, with a positive definite weight matrix  $w$  with all existing moments. In such a case the MVOP exists for every degree  $n$ , and they can be normalized in such a way that

$$\int_{\Gamma} P_n(z)w(z)P_n^t(z) dz = I_d, \quad j = 0, 1, \dots, n-1. \quad (2.20)$$

However, it is interesting to note that MVOP first appeared in connection with prediction theory where the orthogonality is on the unit circle (see [11] for a recent survey). The interest in MVOP on the real line has been steadily growing since the early 1990s. The analytic theory of MVOP on the real line is surveyed in [25], with [5] as one of the pioneering works. MVOP satisfy recurrence relations [34] and special cases satisfy differential equations [33]. Interesting examples of MVOP come from matrix valued spherical functions; see [40, 50] as well as many other papers.

We deviate from the usual set-up of MVOP in several ways:

- $\Gamma$  is a closed contour in the complex plane,
- the weight matrix  $w(z) = \frac{1}{2\pi i} W^N(z)$  is complex valued on  $\Gamma$  and varies with  $N$ ,
- the weight matrix is not symmetric or Hermitian (let alone positive definite), or have any other property that would imply existence and uniqueness of the MVOP.

Since there is no complex conjugation in (2.18), we are thus dealing with non-Hermitian matrix valued orthogonality with varying weights on a closed contour in the plane.

As already noted, existence and uniqueness of the MVOP are not guaranteed in this general setting. However, for the weight  $\frac{1}{2\pi i} W^N$  we can show that the monic MVOP up to degrees  $N$  all exist and are unique. However, since the weight matrix is not symmetric, we cannot normalize to obtain orthonormal MVOP as in (2.20). In our case the MVOP of degrees  $> N$  do not exist.

In Section 4 we consider a situation that is more general than the two-periodic Aztec diamond. It deals with a multi-level particle system that is determinantal, and transitions

between the levels are periodic. See Assumptions 4.1 and 4.2 for the precise assumptions. In this general setting we make a connection with matrix valued (bi)orthogonal polynomials and our main result in Section 4 is Theorem 4.7 that expresses the correlation kernel as a double contour integral containing a reproducing kernel for the matrix polynomials.

In the special situation of the two-periodic Aztec diamond the matrix  $\mathbb{K}_N$  with the correlation kernels as in (2.8) is given by

$$\begin{aligned}
 &-\frac{\chi_{m>m'}}{2\pi i} \oint_{\gamma_{0,1}} A^{m-m'}(z) z^{(m'+n')/2-(m+n)/2} \frac{dz}{z} \\
 &+ \frac{1}{(2\pi i)^2} \oint_{\gamma_{0,1}} \frac{dz}{z} \oint_{\gamma_{0,1}} \frac{dw}{z-w} A^{2N-m'}(w) R_N(w, z) A^m(z) \frac{w^{(m'+n')/2}}{z^{(m+n)/2} w^N} \tag{2.21}
 \end{aligned}$$

where  $R_N(w, z)$  is the reproducing kernel associated with the matrix polynomials of degrees  $\leq N - 1$ . That is,  $R_N(w, z)$  is a bivariate matrix valued polynomial of degree  $N - 1$  in both  $w$  and  $z$  such that

$$\frac{1}{2\pi i} \oint_{\gamma_1} R_N(w, z) W^N(z) Q^t(z) dz = Q^t(w)$$

for every matrix valued polynomial  $Q$  of degree  $\leq N - 1$ .

The MVOP of degree  $N$  is characterized by a Riemann–Hilbert problem, and the reproducing kernel  $R_N(w, z)$  can be expressed in terms of the solution of the Riemann–Hilbert problem. This is known from work of Delvaux [29] and we recall it in Section 4.6. Then we perform an analysis of the Riemann–Hilbert problem, and quite remarkably this produces the exact formula (2.8).

### 2.4. Classification of phases

The explicit formula (2.8) in Theorem 2.2 is suitable for asymptotic analysis as  $N \rightarrow \infty$ . See Figure 5 for a sampling of a large two-periodic Aztec diamond. In this figure three regions emerge where the tiling has different statistical behavior. We first describe how we can distinguish these three phases (solid, liquid and gas) in the model. The classification will depend on the location of saddle points for the double integral in (2.8).

We fix coordinates  $-1 < \xi_1 < 1$  and  $-1 < \xi_2 < 1$  and choose  $m, m' \approx (1 + \xi_1)N$  and  $n, n' \approx (1 + \xi_2)N$ . Then from the formula (2.8), we see that the  $z$ -integral of the double contour integral is dominated as  $N \rightarrow \infty$  by the expression

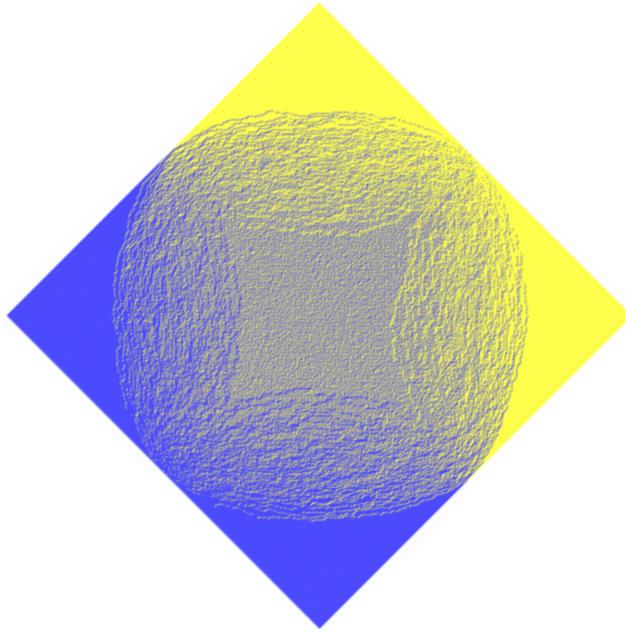
$$A^{\xi_1 N}(z) z^{N/2} (z - 1)^N z^{-(1+\xi_1/2+\xi_2/2)N} = W^{\xi_1 N/2}(z) (z - 1)^N z^{-(1+\xi_2)N/2}$$

where  $W$  is given by (2.15). In view of the eigenvalue decomposition (2.16) this is

$$E(z) \begin{pmatrix} e^{N\Phi_1(z)/2} & 0 \\ 0 & e^{N\Phi_2(z)/2} \end{pmatrix} E(z)^{-1}$$

with  $\Phi_j(z) = 2 \log(z - 1) - (1 + \xi_2) \log z + \xi_1 \log \lambda_j(z)$ ,  $j = 1, 2$ . Hence we are led to consider

$$\Phi(z) = 2 \log(z - 1) - (1 + \xi_2) \log z + \xi_1 \log \lambda(z) \tag{2.22}$$



**Fig. 5.** A sample of a two-periodic Aztec diamond of size 500. We colored the West and South dominos blue. The East and North dominos are colored yellow. The gas phase is visible in the middle. This figure is generated by a code that was kindly provided to us by Sunil Chhita.

as a function on the Riemann surface  $\mathcal{R}$ , depending on parameters  $\xi_1$  and  $\xi_2$ . It is multi-valued, but its differential

$$\Phi'(z)dz = \left( \frac{2}{z-1} - \frac{1+\xi_2}{z} + \xi_1 \frac{\lambda'(z)}{\lambda(z)} \right) dz \quad (2.23)$$

is a single-valued meromorphic differential with simple poles at  $z = 1^{(1)}$ ,  $z = 1^{(2)}$ ,  $z = 0$  and  $z = \infty$ ; see also Section 6.2. (For  $j = 1, 2$ , we use  $1^{(j)}$  to denote the value 1 on the  $j$ th sheet of the Riemann surface.) There are also four zeros, counting multiplicities, since the genus is 1.

**Definition 2.6.** The *saddle points* are the zeros of  $\Phi'(z)dz$ .

The real part  $\mathcal{R}_r$  of the Riemann surface consists of all real tuples  $(z, y)$  satisfying the algebraic equation (2.11) together with the point at infinity. The real part is the union of two cycles,

$$\mathcal{R}_r = \mathcal{C}_1 \cup \mathcal{C}_2, \quad (2.24)$$

where  $\mathcal{C}_1$  is the union of the intervals  $[-\alpha^2, -\beta^2]$  on the two sheets, and  $\mathcal{C}_2$  is the union of the two intervals  $[0, \infty]$  on both sheets.

It turns out that there are always at least two distinct saddle points on the cycle  $\mathcal{C}_1$  (see Proposition 6.4 below). The location of the other two saddle points determines the phase.

**Definition 2.7.** Let  $-1 < \xi_1, \xi_2 < 1$ .

- (a) If two simple saddles are in  $\mathcal{C}_2$ , then  $(\xi_1, \xi_2)$  is in the *solid phase*, and we write  $(\xi_1, \xi_2) \in \mathfrak{S}$ .
- (b) If two saddles are outside the real part of the Riemann surface, then  $(\xi_1, \xi_2)$  is in the *liquid phase*, and we write  $(\xi_1, \xi_2) \in \mathfrak{L}$ .
- (c) If all four saddles are simple and belong to  $\mathcal{C}_1$ , then  $(\xi_1, \xi_2)$  is in the *gas phase*, and we write  $(\xi_1, \xi_2) \in \mathfrak{G}$ .

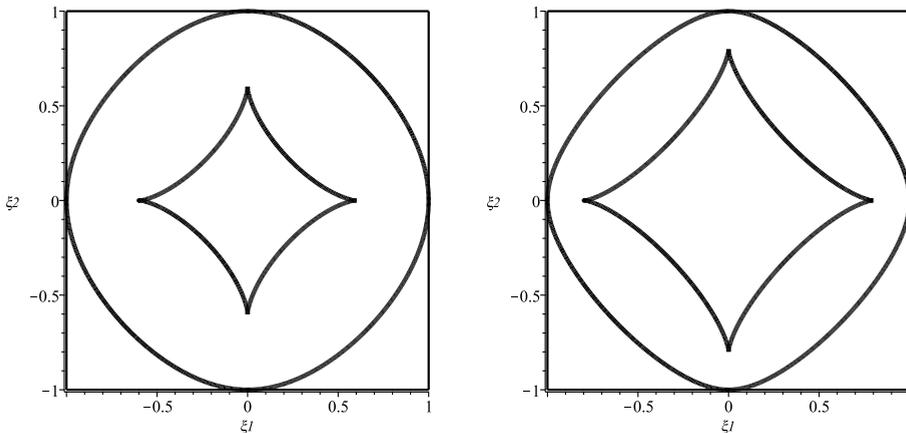
Transitions between phases take place when two or more saddle points coalesce.

- (d) If there is a double saddle point on  $\mathcal{C}_2$ , then  $(\xi_1, \xi_2)$  is on the *solid-liquid transition*.
- (e) If there is a double or triple saddle point on  $\mathcal{C}_1$ , then  $(\xi_1, \xi_2)$  is on the *liquid-gas transition*.

It is not possible to have a double saddle point outside the real part of the Riemann surface.

The condition for coalescing saddle points leads to an algebraic equation of degree 8 for  $\xi_1, \xi_2$  and it precisely coincides with the equation listed in [20, appendix] (see also (6.10) below).

The real section of the degree 8 algebraic equation has two components in case  $\alpha > 1$ , as shown in Figure 6. Both are contained in the square  $-1 \leq \xi_1, \xi_2 \leq 1$ . The outer component is a smooth closed curve that touches the square at the points  $(\pm 1, 0)$  and  $(0, \pm 1)$ . It is the boundary between the solid and liquid phases.



**Fig. 6.** Real section of the degree 8 algebraic curve in the  $\xi_1$ - $\xi_2$  plane for the cases  $\alpha = 2$  (left) and  $\alpha = 3$  (right). The outer component is the boundary between the solid and liquid phases and the inner component is the boundary between the liquid and gas phases.

The inner component is the boundary between the liquid and gas phases. It is a closed curve with four cusps at  $(\pm \frac{\alpha-\beta}{\alpha+\beta}, 0)$  and  $(0, \pm \frac{\alpha-\beta}{\alpha+\beta})$ .

2.5. Gas phase

Our next result gives the limit of  $\mathbb{K}_N$  in the gas phase. Recall that  $\mathbb{K}_N$  is defined by (2.7).

**Theorem 2.8.** *Assume  $(\xi_1, \xi_2) \in \mathfrak{G}$ . Suppose  $m, m', n, n'$  are integers that vary with  $N$  in such a way that*

$$m = (1 + \xi_1)N + o(N), \quad n = (1 + \xi_2)N + o(N) \tag{2.25}$$

as  $N \rightarrow \infty$ , while

$$m' - m = \Delta m, \quad n' - n = \Delta n \tag{2.26}$$

are fixed. Also assume that  $m + n$  and  $m' + n'$  are even. Then for  $N$  even,

$$\lim_{N \rightarrow \infty} \mathbb{K}_N(m, n; m', n') = \mathbb{K}_{\text{gas}}(m, n; m', n') \tag{2.27}$$

with

$$\mathbb{K}_{\text{gas}}(m, n; m', n') = \frac{1}{2\pi i} \oint_{\gamma} (F(z) - \chi_{\Delta m < 0} I_2) A^{-\Delta m}(z) z^{(\Delta m + \Delta n)/2} \frac{dz}{z} \tag{2.28}$$

where  $\gamma$  is a closed contour in  $\mathbb{C} \setminus ((-\infty, -\alpha^2] \cup [-\beta^2, 0])$  going around the interval  $[-\beta^2, 0]$  (which we can also view as a closed loop on the first sheet of the Riemann surface).

The limit (2.28) does not depend on  $\xi_1$  or  $\xi_2$  as long as these are in the gas region. The kernel  $\mathbb{K}_{\text{gas}}$  still depends on the parameters  $\alpha$  and  $\beta$  and therefore the two-periodic structure is still present in this limit.

The proof of Theorem 2.8 is in Section 6.5.

**Remark 2.9.** In (2.28) we can see the exponential decay of correlations that is characteristic for the gas phase as follows. We combine the factor  $z^{(\Delta m)/2}$  with  $A^{-\Delta m}(z)$ , since  $z^{(\Delta m)/2} A^{-\Delta m}(z) = W^{-(\Delta m)/2}(z)$  by (2.15). We find from this and (2.10) and (2.15) that

$$\begin{aligned} F(z)z^{(\Delta m)/2} A^{-\Delta m}(z) &= F(z)\lambda_1^{-(\Delta m)/2}(z), \\ (F(z) - I_2)z^{(\Delta m)/2} A^{-\Delta m}(z) &= (F(z) - I_2)\lambda_2^{-(\Delta m)/2}(z) \\ &= (F(z) - I_2)\lambda_1^{(\Delta m)/2}(z) \end{aligned}$$

since  $\lambda_2 = \lambda_1^{-1}$  (see Lemma 5.2(d) below). Thus (2.28) can be written as

$$\mathbb{K}_{\text{gas}}(m, n; m', n') = \begin{cases} \frac{1}{2\pi i} \oint_{\gamma} F(z)\lambda_1^{-(\Delta m)/2}(z)z^{(\Delta n)/2} \frac{dz}{z} & \text{if } \Delta m \geq 0, \\ \frac{1}{2\pi i} \oint_{\gamma} (F(z) - I_2)\lambda_1^{(\Delta m)/2}(z)z^{(\Delta n)/2} \frac{dz}{z} & \text{if } \Delta m < 0. \end{cases} \tag{2.29}$$

By analyticity and Cauchy’s theorem, we are free to deform the contour  $\gamma$  as long as it goes around the interval  $[-\beta^2, 0]$  and does not intersect  $(-\infty, -\alpha^2]$ . Since  $\beta < 1 < \alpha$  we can deform it to a circle centered at zero of radius  $< 1$  or to a circle of radius  $> 1$ .

If  $\Delta n \geq 0$  we deform to a circle  $|z| = r < 1$  and if  $\Delta n < 0$  we deform to a circle  $|z| = r > 1$ . In both cases the factor  $z^{(\Delta n)/2}$  is exponentially small as  $|\Delta n| \rightarrow \infty$ .

Since  $|\lambda_1| > 1$  on  $\gamma$  (as will follow from Lemma 5.2(d, f) below) the factor  $\lambda_1^{\pm(\Delta m)/2}(z)$  is also exponentially small as  $|\Delta m| \rightarrow \infty$ . It follows that the gas kernel (2.29) decays exponentially as  $|\Delta m| + |\Delta n| \rightarrow \infty$ .

### 2.6. Cusp points

On the boundary between the gas phase and the liquid phase, the gas kernel (2.28) is still the dominant contribution. This phenomenon was already observed by Chhita and Johansson [20] and further investigated by Beffara, Chhita and Johansson [9], who looked at the diagonal point  $\xi_1 = \xi_2$  on the boundary and proved that after averaging there is Airy-like behavior in the first subleading term.

We consider the cusp points, and show explicitly the appearance of Pearcey-like behavior in the subleading term of the kernel  $\mathbb{K}_N$ .

The four cusp points are located at  $(\xi_1, \xi_2) = (\pm \frac{\alpha - \beta}{\alpha + \beta}, 0)$  and  $(\xi_1, \xi_2) = (0, \pm \frac{\alpha - \beta}{\alpha + \beta})$  in the phase diagram. We focus on the top cusp point with coordinates  $(\xi_1^*, \xi_2^*) = (0, \frac{\alpha - \beta}{\alpha + \beta})$ . At this cusp point the triple saddle is located at the branch point  $-\alpha^2$ .

**Theorem 2.10.** *Suppose  $N, m + n$  and  $m' + n'$  are even. Write  $m = (1 + \xi_1)N, n = (1 + \xi_2)N, m' = (1 + \xi'_1)N, n' = (1 + \xi'_2)N$  and assume*

$$\begin{aligned} N^{3/4} \xi_1 &\rightarrow c_1 u, & N^{1/2}(\xi_2 - \xi_2^*) &\rightarrow c_2 v, & \xi_2^* &= \frac{\alpha - \beta}{\alpha + \beta}, \\ N^{3/4} \xi'_1 &\rightarrow c_1 u', & N^{1/2}(\xi'_2 - \xi_2^*) &\rightarrow c_2 v' \end{aligned} \tag{2.30}$$

as  $N \rightarrow \infty$ , with fixed  $u, u', v, v'$  and with constants

$$c_1 = \frac{2^{1/4}}{\sqrt{\alpha - \beta}}, \quad c_2 = \frac{\sqrt{2}}{\alpha + \beta}. \tag{2.31}$$

Then, in case  $m$  is even,

$$\begin{aligned} \lim_{N \rightarrow \infty} N^{1/4} (-1)^{(\Delta n - \Delta m)/2} \alpha^{-\Delta n} (\mathbb{K}_N(m, n; m', n') - \mathbb{K}_{\text{gas}}(m, n; m', n')) \\ = \frac{\sqrt{\alpha - \beta}}{2^{1/4}} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \frac{1}{(2\pi i)^2} \int_{\Sigma \cup (-\Sigma)} \int_{i\mathbb{R}} \frac{e^{\frac{1}{4}s^4 + \frac{1}{2}vs^2 + us}}{e^{\frac{1}{4}t^4 + \frac{1}{2}v't^2 + u't}} \frac{ds dt}{t - s} \end{aligned} \tag{2.32}$$

with contours  $\Sigma$  and  $-\Sigma$  as shown in Figure 7.

In case  $m$  is odd,

$$\begin{aligned} \lim_{N \rightarrow \infty} N^{1/4} (-1)^{(\Delta n - \Delta m)/2} \alpha^{-\Delta n} (\mathbb{K}_N(m, n; m', n') - \mathbb{K}_{\text{gas}}(m, n; m', n')) \\ = \frac{\sqrt{\alpha - \beta}}{2^{1/4}} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \frac{1}{(2\pi i)^2} \int_{\Sigma \cup (-\Sigma)^{-1}} \int_{i\mathbb{R}} \frac{e^{\frac{1}{4}s^4 + \frac{1}{2}vs^2 + us}}{e^{\frac{1}{4}t^4 + \frac{1}{2}v't^2 + u't}} \frac{ds dt}{t - s} \end{aligned} \tag{2.33}$$

where  $(-\Sigma)^{-1}$  indicates that the orientation on  $-\Sigma$  is reversed.

The proof of Theorem 2.10 is in Section 6.6.

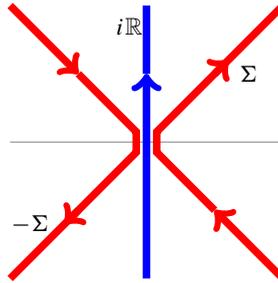


Fig. 7. The contours of integration for the Pearcey integrals in (2.32) and (2.33).

The double integral in (2.32),

$$\frac{1}{(2\pi i)^2} \int_{\Sigma \cup (-\Sigma)} \int_{i\mathbb{R}} \frac{e^{\frac{1}{4}s^4 + \frac{1}{2}vs^2 + us}}{e^{\frac{1}{4}t^4 + \frac{1}{2}v't^2 + u't}} \frac{ds dt}{t - s}, \tag{2.34}$$

is, up to a gaussian, known as the Pearcey kernel. It is one of the canonical kernels in random matrix theory that arises typically as a scaling limit near a cusp point. It was first described by Brézin and Hikami [17] in the context of random matrices with an external source (see also [13]). The Pearcey process was defined in [58, 63]. More recent contributions are for example [1, 10, 41]. Note that the actual Pearcey kernel includes a gaussian in addition to the double integral in (2.34). Remarkably, this term is hidden in the gas kernel and can be retrieved by a steepest descent analysis of that kernel.

**Theorem 2.11.** *Under the same assumptions as in Theorem 2.10 we have*

$$\begin{aligned} & \lim_{N \rightarrow \infty} N^{1/4} (-1)^{(\Delta n - |\Delta m|)/2} \alpha^{-\Delta n} \mathbb{K}_{\text{gas}}(m, n; m', n') \\ &= \begin{cases} \frac{\sqrt{\alpha - \beta}}{2^{1/4}} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \frac{1}{\sqrt{2\pi(v - v')}} e^{-\frac{(u-u')^2}{2(v-v')}} & \text{if } v > v', \\ 0 & \text{if } v < v' \text{ or} \\ & \text{if } v = v' \text{ and } u \neq u', \end{cases} \end{aligned} \tag{2.35}$$

as  $N \rightarrow \infty$ .

It is very curious that the double integral part of the Pearcey kernel appears in the scaling limit at the cusp point, but only in the subleading term. A similar phenomenon was already observed in [20] on the smooth parts of the liquid-gas boundary. The gas phase is dominant with subleading Airy behavior. Also here the gaussian part of the Airy kernel is hidden in the gas kernel [20, §3.2]. With some effort we can also find this from our approach.

### 3. Non-intersecting paths

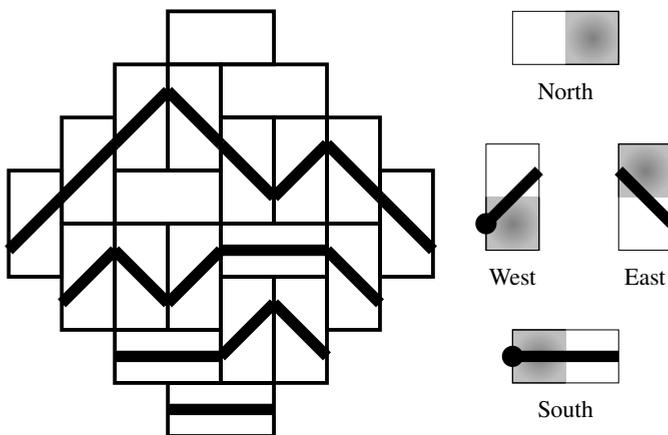
We discuss non-intersecting paths on the Aztec diamond. What follows in this section is not new, and can be found in several places; see e.g. [43, 44, 45] and the recent works

[7, 46]. Note, however, that we use a (random) double Aztec diamond to extend the paths instead of a deterministic extension in [43, 44]. Since the discussion is rather standard, we allow ourselves to be brief and refer to the arXiv version of the paper for more details.

### 3.1. Non-intersecting paths

The South, West and East dominos are marked by line segments as shown in Figure 8. The North dominos have no marking. There are also particles on the West and South dominos, but these will only play a role later on. We look at the line segment as part of paths that go from left to right and go either up (in a West domino), down (in an East domino), or horizontal (in a South domino). Each segment enters a domino in the black square, and exits it from the white square within the domino.

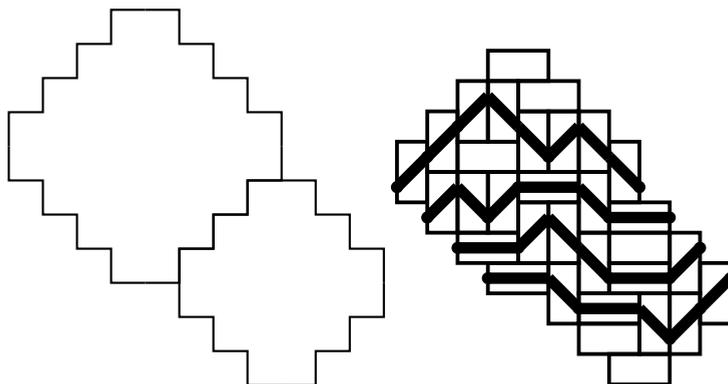
By including the markings on the tiles we obtain non-intersecting paths, starting at the lower left side of the Aztec diamond and ending at the lower right side. In the pictures that follow we forget about the black/white shading of the dominos.



**Fig. 8.** Line segments and particles on the dominos, which lead to non-intersecting paths in a domino tiling of the Aztec diamond.

Each path ends at the same height as where it started. Observe that the lengths of the paths vary greatly. To obtain a more symmetric picture, which will be useful for what follows, we attach to the right bottom side of the original Aztec diamond of size  $2N$  another one of size  $2N - 1$  as in Figure 9. It is not hard to see [32, Lemma 3.2] that any domino tiling of the double Aztec diamond splits into a domino tiling of the original Aztec diamond of size  $2N$  and a domino tiling of the attached Aztec diamond of size  $2N - 1$ . In other words: there are no dominos that are partly in the original Aztec diamond and partly in the newly attached Aztec diamond.

We thus cover the second Aztec diamond with dominos, independently from what we did in the original Aztec diamond. We also put in the markings with line segments.



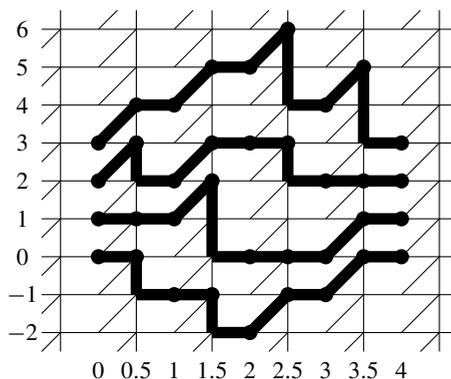
**Fig. 9.** The double Aztec diamond (left) and a tiling of the double Aztec diamond with non-intersecting paths and particles along the paths (right).

A possible tiling is shown in Figure 9 together with the corresponding non-intersecting paths.

The double Aztec diamond with partial overlap is considered in [2, 3], where the phenomenon of a tacnode is studied. For us, the two Aztec diamonds do not overlap and there is no tacnode phenomenon.

### 3.2. Modified paths on a graph

We are going to modify the paths in a standard way. First of all, we perform a shear transform  $(x, y) \mapsto (x + y, y)$  so that the paths end at the same height as they started. We also include some trivial steps so that each possible step up is preceded by a horizontal step of half a unit. See Figure 10.



**Fig. 10.** Modified paths on a directed graph. Horizontal and diagonal edges are oriented to the right and vertical edges are oriented downwards.

We then put particles on the paths as follows. The particles are on the integer lattice  $\mathbb{Z} \times \mathbb{Z}$ . We take coordinates so that the initial vertices are at  $(0, j)$  for  $j = 0, \dots, 2N - 1$  and the ending vertices are at  $(2N, j)$  for  $j = 0, \dots, 2N - 1$ . We then put particles at the intersection points of the paths and the vertical lines as follows. If the path has no vertical part, then it has a unique intersection point with the vertical line, and we put a particle there. If there is a vertical part of the path at that level, then we put the particle at the highest point (see Figure 10). Now there are  $4N + 1$  particles on each path.

The new paths have a two-step structure. Starting from an initial position, either we move horizontally to the right half a unit and stay at the same height, or we go diagonally up one unit in the vertical direction and horizontally half a unit. We call this a *Bernoulli step*. In both cases we end at a particle on the line with horizontal coordinate  $1/2$ . Then we make a number of vertical down-steps followed by a horizontal step half a unit to the right. The number of down-steps can be any non-negative integer, including zero. We call this a *geometric step*. Then we repeat the pattern. We do a Bernoulli step, a geometric step, a Bernoulli step, etc. The final step in each path is a geometric step, which should take us to the same height as where the path started.

Another requirement is that the resulting paths are non-intersecting. Any such path structure is in one-to-one correspondence with a unique domino tiling of the double Aztec diamond. The paths lie on an infinite directed graph that is also shown in Figure 10. We call it the *Aztec diamond graph*.

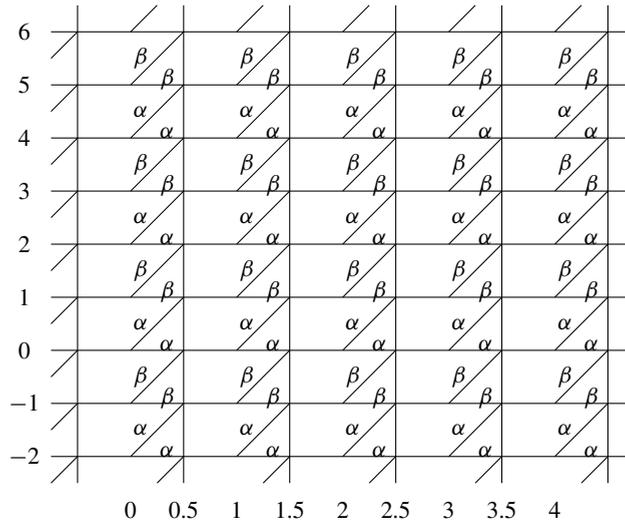
### 3.3. Weights

There is a one-to-one correspondence between tilings of the double Aztec diamond and non-intersecting paths on the Aztec diamond graph with prescribed initial and ending positions as described above.

In the two-periodic Aztec diamond we assign a weight (2.2) to a tiling with the corresponding probability (2.3). To be able to transfer this to the paths, we recall that in a Bernoulli step a diagonal up-step corresponds to a West domino, and a horizontal step corresponds to a South domino. The vertical steps in a geometric step correspond to East dominos. The horizontal step that closes a geometric step was added artificially and does not correspond to a domino.

We do not see the North dominos in the paths, and therefore we cannot transfer the weights on the dominos to weights on path segments directly. It is possible to assign weights to dominos in which North dominos have weight 1 and which is equivalent to (2.1) as it leads to the same probabilities (2.3) on tilings. This was also done in [20], but we present it in a different way here. A simple argument [32, Lemma 3.1] shows that each column has the same number of North and South dominos, and these dominos all have the same weight (2.1). We obtain the same weight (2.2) of a tiling if instead of assigning the same weight  $a$  or  $b$  to all the horizontal dominos in a column, we assign  $a^2$  or  $b^2$  to the South dominos and weight 1 to the North dominos in that column.

For symmetry reasons we apply the same operation to East and West dominos. Then instead of (2.1) we assign the following weight to a domino  $D$ , where we recall that



**Fig. 11.** Weights on the edges of the Aztec diamond graph corresponding to the weights (3.1) on dominos. Horizontal edges and diagonal edges from level  $m \in \mathbb{Z}$  to  $m + 1/2$  have weight  $\alpha$  in even numbered rows and weight  $\beta$  in odd numbered rows. The vertical edges and horizontal edges from level  $m + 1/2$  to  $m + 1$  have weight 1 and these weights are not shown in the figure.

$\alpha = a^2$  and  $\beta = b^2$ :

$$\widehat{w}(D) = \begin{cases} \alpha & \text{if } D \text{ is a South domino in an even column,} \\ \beta & \text{if } D \text{ is a South domino in an odd column,} \\ \beta & \text{if } D \text{ is a West domino in an even row,} \\ \alpha & \text{if } D \text{ is a West domino in an odd row,} \\ 1 & \text{if } D \text{ is a North or East domino.} \end{cases} \tag{3.1}$$

Since North dominos have weight 1, we can transfer the weights (3.1) on dominos to weights on the edges of the Aztec diamond graph. The result is shown in Figure 11. The weights alternate per row.

### 3.4. Transition matrices

We use the layered structure of the Aztec diamond graph to introduce transition matrices between levels. Here a level is just the horizontal coordinate. There are integer levels  $m$  and half-integer levels  $m + 1/2$  with  $m \in \mathbb{Z}$ .

The transition from level  $m$  to  $m + 1/2$  is a Bernoulli step. Because of the weights, the transition matrix is, for  $x, y \in \mathbb{Z}$ ,

$$T_{m,m+1/2}(x, y) = \begin{cases} \alpha & \text{if } x \text{ is even and } y \in \{x, x + 1\}, \\ \beta & \text{if } x \text{ is odd and } y \in \{x, x + 1\}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

Then  $T_{m,m+1/2}$  is two-periodic,  $T_{m,m+1/2}(x+2, y+2) = T_{m,m+1/2}(x, y)$  for all  $x, y \in \mathbb{Z}$ . As a matrix it is a block Laurent matrix (i.e., a block Toeplitz matrix that is infinite in both directions) with  $2 \times 2$  blocks. The diagonal block is  $\begin{pmatrix} \alpha & \alpha \\ 0 & \beta \end{pmatrix}$ , the block on the first upper diagonal is  $\begin{pmatrix} 0 & 0 \\ \beta & 0 \end{pmatrix}$ , and all other diagonals are zero. The associated symbol [16] is

$$A_{m,m+1/2}(z) = \begin{pmatrix} \alpha & \alpha \\ 0 & \beta \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \beta & 0 \end{pmatrix} z = \begin{pmatrix} \alpha & \alpha \\ \beta z & \beta \end{pmatrix} \tag{3.3}$$

with  $z \in \mathbb{C}$ .

To go from level  $m + 1/2$  to level  $m + 1$ , we make a number of vertical down-steps (possibly zero) and then a horizontal step. All weights are 1 and so the transition matrix is

$$T_{m+1/2,m+1}(x, y) = \begin{cases} 1 & \text{if } y \leq x, \\ 0 & \text{if } y > x. \end{cases} \tag{3.4}$$

This is a Laurent matrix, but we want to view it as a block Laurent matrix with  $2 \times 2$  blocks. The diagonal block is  $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ , all blocks below the main diagonal are  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$  and all blocks above the main diagonal are zero. The symbol is

$$A_{m+1/2,m+1}(z) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + \sum_{j=-\infty}^{-1} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} z^j = \frac{1}{z-1} \begin{pmatrix} z & 1 \\ z & z \end{pmatrix} \tag{3.5}$$

with  $|z| > 1$ .

Then (the product is matrix multiplication)

$$T = T_{m,m+1/2} T_{m+1/2,m+1} \tag{3.6}$$

is the transition matrix from level  $m$  to level  $m + 1$ , and  $T$  is two-periodic. The symbol for  $T$  is easily seen to be the product of (3.3) and (3.5):

$$A(z) = A_{m,m+1/2}(z) A_{m+1/2,m+1}(z) = \frac{1}{z-1} \begin{pmatrix} 2\alpha z & \alpha(z+1) \\ \beta z(z+1) & 2\beta z \end{pmatrix}, \tag{3.7}$$

which agrees with (2.9).

More generally, for any integers  $m < m'$  we have a transition matrix  $T^{m'-m}$  to go from level  $m$  to level  $m'$  with symbol  $A^{m'-m}$ . In particular  $T^{2N}$  is the transition matrix from level 0 to  $2N$  with symbol  $A^{2N}$ .

Now we want to invoke the Lindström–Gessel–Viennot lemma [38, 53] (see also [45, Theorem 3.1] for a proof), which gives an expression for the weighted number of non-intersecting paths on the graph with prescribed starting and ending positions. For us, the starting positions are  $(0, j)$ ,  $j = 0, \dots, 2N - 1$ , and the ending positions  $(2N, j)$ ,  $j = 0, \dots, 2N - 1$ . Since  $T^{2N}(j, k)$  is the sum of all weighted paths from  $(0, j)$  to  $(2N, k)$ , by the Lindström–Gessel–Viennot lemma the partition function is a determinant

$$Z_N = \det [T^{2N}(j, k)]_{j,k=0}^{2N-1}. \tag{3.8}$$

Because of the layered structure in the graph, we can also look at the positions of the particles at intermediate levels  $m$ . We restrict to integer values  $m$ , but we could also include the half-integer values.

Given an admissible  $2N$ -tuple of non-intersecting paths, we then find a point set configuration  $(x_j^m)_{j=0,m=1}^{2N-1,2N-1}$  where  $x_0^m < x_1^m < \dots < x_{N-1}^m$  are the vertical coordinates of the particles at level  $m$ . The probability measure on admissible tuples of non-intersecting paths yields a probability measure on particle configurations in  $\{1, \dots, 2N - 1\} \times \mathbb{Z}^N$ .

Then another application of the Lindström–Gessel–Viennot lemma shows that the joint probability for the particle configuration  $(x_j^m)_{j=0,m=1}^{2N-1,2N-1}$  is

$$\text{Prob}((x_j^m)_{j=0,m=1}^{2N-1,2N-1}) = \frac{1}{Z_N} \prod_{m=0}^{N-1} \det [T(x_j^m, x_k^{m+1})]_{j,k=0}^{2N-1}$$

$$\text{with } x_j^0 = x_j^{2N} = j \text{ for } j = 0, \dots, 2N - 1. \quad (3.9)$$

The point process (3.9) is determinantal. The correlation kernel is given by the Eynard–Mehta theorem [36] (see also [14, 15]). In the next section we will show how the kernel can be represented in terms of certain matrix valued orthogonal polynomials. In fact, the connection with matrix valued polynomials will hold in greater generality.

### 4. Determinantal point processes and MVOP

#### 4.1. The model

We analyze the following situation. We take an integer  $p \geq 1$ , and we consider transition matrices  $T_m : \mathbb{Z}^2 \rightarrow \mathbb{R}$  for  $m \in \mathbb{Z}$  that are  $p$ -periodic. This means that  $T_m(x + p, y + p) = T_m(x, y)$  for every  $m$  and  $x, y \in \mathbb{Z}$ .

The model also depends on integers  $N, L \in \mathbb{N}$  and  $M \in \mathbb{Z}$ . There will be  $L + 1$  levels numbered as  $0, 1, \dots, L$ . At each level  $m$  there are  $pN$  particles at integer positions denoted by  $x_0^m < x_1^m < \dots < x_{pN-1}^m$ . The initial and ending positions (at levels 0 and  $L$ ) are deterministic and are given by consecutive integers

$$x_j^0 = j, \quad x_j^L = pM + j, \quad j = 0, \dots, pN - 1. \quad (4.1)$$

Our assumption for this section is the following.

**Assumption 4.1.**  $(x_j^m)_{j=0,m=0}^{pN-1,L}$  is a multi-level particle system with joint probability

$$\text{Prob}((x_j^m)_{j=0,m=0}^{pN-1,L}) = \frac{1}{Z_N} \det [\delta_j(x_k^0)]_{j,k=0}^{pN-1}$$

$$\cdot \left( \prod_{m=0}^{L-1} \det [T_m(x_j^m, x_k^{m+1})]_{j,k=0}^{pN-1} \right) \cdot \det [\delta_{pM+j}(x_k^L)]_{j,k=0}^{pN-1}, \quad (4.2)$$

where the transition matrices  $T_m$  are  $p$ -periodic for every  $m$ . The constant  $Z_N$  in (4.2) is a normalizing constant and  $\delta_j(x) = \delta_{j,x}$  is the Kronecker delta.

The determinantal factors with the delta-functions in (4.2) ensure the boundary conditions (4.1).

Assumption 4.1 is satisfied for the two-periodic Aztec diamond by (3.9), provided we take  $p = 2, L = 2N, M = 0$  and  $T_m = T$  for each integer  $m$ , with  $T$  given by (3.6), using (3.2), (3.4). There are three crucial assumptions contained in Assumption 4.1.

- The transition matrices are  $p$ -periodic. It means that the  $T_m$  are block Laurent matrices with  $p \times p$  blocks.
- The initial and ending positions of the particles are at consecutive integers. This assumption allows us to make a connection with matrix valued polynomials. Note that we allow for a shift  $pM$  in the positions at level  $L$  compared to the initial positions.
- The transition matrices are such that (4.2) is a probability. That is, (4.2) is always non-negative and we can find a normalization constant  $Z_N$  such that all probabilities (4.2) add up to 1.

We have made two other assumptions, namely

- the number  $pN$  of particles at each level is a multiple of  $p$ , and
- the shift  $pM$  in the positions of the particles at the final level is also a multiple of  $p$ ,

but these are less essential. They are made for convenience and ease of notation and could be relaxed if needed.

For future analysis, we also assume

**Assumption 4.2.** The symbols for the block Laurent matrices  $T_m, m \in \mathbb{Z}$ , are analytic in a common annular domain  $R_1 < |z| < R_2$  in the complex plane.

For  $m < m'$  we use

$$T_{m,m'} = T_m \cdot T_{m+1} \cdots T_{m'-1} \tag{4.3}$$

for the transition matrix from level  $m$  to level  $m'$ . The matrix multiplication is well defined because of Assumption 4.2. Every  $T_{m,m'}$  is a block Laurent matrix with period  $p$ .

The Eynard–Mehta theorem [36] applies to (4.2). We present the Eynard–Mehta theorem as stated in [14]. We assume  $\phi_j, \psi_j$  for  $j = 0, 1, \dots, N - 1$  are given functions, and for functions  $\phi, \psi : \mathbb{Z} \rightarrow \mathbb{R}$  and  $T : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{R}$  we use

$$\begin{aligned} (\phi * T)(y) &= \sum_{x \in \mathbb{Z}} \phi(x)T(x, y), & (T * \psi)(x) &= \sum_{y \in \mathbb{Z}} T(x, y)\psi(y), \\ \phi * \psi &= \sum_{x \in \mathbb{Z}} \phi(x)\psi(x). \end{aligned}$$

**Theorem 4.3** (Eynard–Mehta). *A multi-level particle system of the form*

$$\begin{aligned} \text{Prob}((x_j^m)_{j=0,m=0}^{N-1,L}) &= \frac{1}{Z_N} \det [\phi_j(x_k^0)]_{j,k=0}^{N-1} \\ &\cdot \left( \prod_{m=0}^{L-1} \det [T_m(x_j^m, x_k^{m+1})]_{j,k=0}^{N-1} \right) \cdot \det [\psi_j(x_k^L)]_{j,k=0}^{N-1}, \end{aligned} \tag{4.4}$$

where  $\phi_j, \psi_j$  for  $j = 0, 1, \dots, N - 1$  are arbitrary given functions, is determinantal with correlation kernel

$$K(m, x; m', y) = -\chi_{m > m'} T_{m',m}(y, x) + \sum_{i,j=0}^{N-1} (\phi_i * T_{0,m})(x) [\mathbf{G}^{-t}]_{i,j} (T_{m',L} * \psi_j)(y) \tag{4.5}$$

where the Gram matrix  $\mathbf{G}$  is defined by

$$\mathbf{G} = (G_{i,j})_{i,j=0}^{N-1}, \quad G_{i,j} = \phi_i * T_{0,L} * \psi_j. \tag{4.6}$$

We apply Theorem 4.3 to (4.2) and we find the following correlation kernel:

**Corollary 4.4.** *The multi-level particle system (4.2) is determinantal with correlation kernel*

$$K(m, x; m', y) = -\chi_{m > m'} T_{m',m}(y, x) + \sum_{i,j=0}^{pN-1} T_{0,m}(i, x) [\mathbf{G}^{-t}]_{i,j} T_{m',L}(y, pM + j) \tag{4.7}$$

with

$$\mathbf{G} = (G_{i,j})_{i,j=0}^{pN-1}, \quad G_{i,j} = T_{0,L}(i, pM + j). \tag{4.8}$$

*Proof.* This follows from Theorem 4.3 since we have  $(\delta_i * T_{0,m})(x) = T_{0,m}(i, x)$  and  $(T_{m',L} * \delta_{pM+j})(y) = T_{m',L}(y, pM + j)$ . □

The Gram matrix  $\mathbf{G}$  in (4.6) is a finite section of the block Laurent matrix  $T_{0,L}$ . It has size  $pN \times pN$  and we also view it as a block Toeplitz matrix of size  $N \times N$  with blocks of size  $p \times p$ . It is part of the conclusion of the Eynard–Mehta Theorem 4.3 that  $\mathbf{G}$  is invertible, and so the invertibility is in particular a consequence of Assumptions 4.1 and 4.2.

#### 4.2. Symbols and matrix biorthogonality

Associated with the block Laurent matrices  $T_m$  and  $T_{m,m'}$  we have the symbols  $A_m(z)$  and  $A_{m,m'}(z)$ . According to Assumption 4.2 all symbols are analytic in an annular domain  $R_1 < |z| < R_2$ . We have the identity

$$A_{m,n}(z) = A_m(z) A_{m+1}(z) \cdots A_{m'-1}(z), \quad m < m', \tag{4.9}$$

for  $R_1 < |z| < R_2$ . The series that define the symbol do not commute (in general) and thus the order of the factors in the product is important.

We let  $\gamma$  be a circle of radius  $R \in (R_1, R_2)$  with counterclockwise orientation. By Cauchy’s theorem, we can recover the Laurent matrix entries from the symbols, and we have

$$[T_{m,m'}(px + i, py + j)]_{i,j=0}^{p-1} = \frac{1}{2\pi i} \oint_{\gamma} A_{m,m'}(z) z^{x-y} \frac{dz}{z}. \tag{4.10}$$

In particular

$$[T_{0,L}(px + i, pM + py + j)]_{i,j=0}^{p-1} = \frac{1}{2\pi i} \oint_{\gamma} A_{0,L}(z) z^{x-y-M} \frac{dz}{z}, \tag{4.11}$$

and if we restrict to  $0 \leq x, y \leq N - 1$  then the blocks (4.11) are the  $p \times p$  blocks in the Gram matrix  $\mathbf{G}$  (see (4.8)).

We consider the matrix-valued weight

$$W_{0,L}(z) = \frac{A_{0,L}(z)}{z^{M+N}}, \quad z \in \gamma, \tag{4.12}$$

on the contour  $\gamma$ . Clearly  $W_{0,L}$  also depends on  $M$  and  $N$  but we do not include this in the notation. It introduces a bilinear pairing between  $p \times p$  matrix valued functions,

$$\langle P, Q \rangle = \frac{1}{2\pi i} \oint_{\gamma} P(z) W_{0,L}(z) Q^t(z) dz \tag{4.13}$$

where  $Q^t$  denotes the matrix transpose (no complex conjugation). The integral is taken entrywise, and so  $\langle P, Q \rangle$  is again a  $p \times p$  matrix.

A matrix valued function is a *polynomial of degree  $\leq d$*  if all its entries are polynomials of degree  $\leq d$ .

For invertible matrices  $\mathbf{P}$  and  $\mathbf{Q}$  of size  $pN \times pN$  we define

$$\begin{pmatrix} P_0(z) \\ P_1(z) \\ \vdots \\ P_{N-1}(z) \end{pmatrix} = \mathbf{P} \begin{pmatrix} I_p \\ zI_p \\ \vdots \\ z^{N-1}I_p \end{pmatrix}, \quad \begin{pmatrix} Q_0(z) \\ Q_1(z) \\ \vdots \\ Q_{N-1}(z) \end{pmatrix} = \mathbf{Q} \begin{pmatrix} z^{N-1}I_p \\ \vdots \\ zI_p \\ I_p \end{pmatrix}. \tag{4.14}$$

Then  $P_j$  and  $Q_j$  for  $j = 0, 1, \dots, N - 1$  are matrix valued polynomials of degrees  $\leq N - 1$ .

**Proposition 4.5.** *Let  $\mathbf{P}$  and  $\mathbf{Q}$  be invertible  $pN \times pN$  matrices, and let  $P_j, Q_j$  be the matrix valued polynomials as in (4.14). Let  $\mathbf{G}$  be the Gram matrix from (4.8). The following are equivalent:*

- (a)  $\mathbf{G}^{-1} = \mathbf{Q}'\mathbf{P}$ .
- (b) For each  $j, k = 0, 1, \dots, N - 1$ ,

$$\frac{1}{2\pi i} \oint_{\gamma} P_j(z) W_{0,L}(z) Q_k^t(z) dz = \delta_{j,k} I_p \tag{4.15}$$

where  $W_{0,L}(z) = A_{0,L}(z)/z^{M+N}$  as in (4.12).

*Proof.* Consider

$$X = \frac{1}{2\pi i} \oint_{\gamma} \begin{pmatrix} P_0(z) \\ P_1(z) \\ \vdots \\ P_{N-1}(z) \end{pmatrix} W_{0,L}(z) \begin{pmatrix} Q_0^t(z) & Q_1^t(z) & \cdots & Q_{N-1}^t(z) \end{pmatrix} dz.$$

Then, by (4.12) and the definition (4.14) of the matrix valued polynomials,

$$\mathbf{P}^{-1} X (\mathbf{Q}^{-1})^t = \frac{1}{2\pi i} \oint_{\gamma} \begin{pmatrix} I_p \\ zI_p \\ \vdots \\ z^{N-1}I_p \end{pmatrix} \frac{A_{0,L}(z)}{z^M} \begin{pmatrix} I_p & z^{-1}I_p & \cdots & z^{-N+1}I_p \end{pmatrix} \frac{dz}{z}.$$

This is a block Toeplitz matrix with  $p \times p$  blocks. For  $0 \leq x, y \leq N - 1$ , the  $xy$ th block is

$$\frac{1}{2\pi i} \oint_{\gamma} A_{0,L}(z) z^{x-y-M} \frac{dz}{z},$$

which by (4.8) and (4.11) is equal to the  $xy$ th block of  $\mathbf{G}$ . Thus  $\mathbf{P}^{-1} X (\mathbf{Q}^{-1})^t = \mathbf{G}$ , which means that  $\mathbf{G}^{-1} = \mathbf{Q}^t \mathbf{P}$  if and only if  $X = I_{pN}$ . This proves the proposition, since  $X = I_{pN}$  is equivalent to the biorthogonality (4.15).  $\square$

The property (4.15) is a matrix valued biorthogonality between the two sequences  $(P_j)_{j=0}^{N-1}$  and  $(Q_j)_{j=0}^{N-1}$ . The matrix valued biorthogonal polynomials are clearly not unique but depend on the particular factorization of  $\mathbf{G}^{-1}$ .

### 4.3. Reproducing kernel

Let  $\mathbf{G}^{-1} = \mathbf{Q}^t \mathbf{P}$  be any factorization of  $\mathbf{G}^{-1}$  and let  $P_j, Q_j$  be the matrix polynomials as in (4.14). We consider

$$R_N(w, z) = \sum_{j=0}^{N-1} Q_j^t(w) P_j(z), \tag{4.16}$$

which is a bivariate polynomial of degree  $\leq N - 1$  in both  $w$  and  $z$ .

#### Lemma 4.6.

(a) For every matrix valued polynomial  $P$  of degree  $\leq N - 1$  we have

$$\frac{1}{2\pi i} \oint_{\gamma} P(w) W_{0,L}(w) R_N(w, z) dw = P(z). \tag{4.17}$$

(b) For every matrix valued polynomial  $Q$  of degree  $\leq N - 1$  we have

$$\frac{1}{2\pi i} \oint_{\gamma} R_N(w, z) W_{0,L}(z) Q^t(z) dz = Q^t(w). \tag{4.18}$$

(c) *Either of the properties (a) and (b) characterizes (4.16) in the sense that if a bivariate polynomial  $\widehat{R}_N(w, z)$  of degree  $\leq N - 1$  in both  $w$  and  $z$  satisfies either (a) or (b), then  $\widehat{R}_N(w, z) = R_N(w, z)$  for every  $w, z \in \mathbb{C}$ .*

*Proof.* Parts (a) and (b) are immediate from the biorthogonality (4.15) and the fact that any matrix valued polynomials  $P$  of degree  $\leq N - 1$  can be written as  $P(z) = \sum_{k=0}^{N-1} A_k P_k(z) = \sum_{k=0}^{N-1} B_k Q_k(z)$  for suitable constant matrices  $A_0, \dots, A_{N-1}, B_0, \dots, B_{N-1}$ .

Let  $\widehat{R}_N(w, z)$  be as in part (c), and suppose that

$$\frac{1}{2\pi i} \oint_{\gamma} \widehat{R}_N(w, z) W_{0,L}(z) Q^t(z) dz = Q^t(w) \tag{4.19}$$

for every matrix valued polynomial  $Q$  of degree  $\leq N - 1$ . For a fixed  $w$  we note that  $z \mapsto \widehat{R}_N(w, z)$  is a matrix valued polynomial of degree  $\leq N - 1$  and it can be written as a linear combination of  $P_0, \dots, P_{N-1}$  with matrix coefficients. The matrix coefficients depend on  $w$ , and we get for some  $A_j(w), j = 0, \dots, N - 1$ ,

$$\widehat{R}_N(w, z) = \sum_{j=0}^{N-1} A_j(w) P_j(z). \tag{4.20}$$

Then taking (4.19) with  $Q = Q_k$ , and using the biorthogonality (4.15), we get

$$Q_k^t(w) = \sum_{j=0}^{N-1} A_j(w) \frac{1}{2\pi i} \oint_{\gamma} P_j(z) W_{0,L}(z) Q_k^t(z) dz = A_k(w)$$

for every  $k = 0, \dots, N - 1$ . Thus  $\widehat{R}_N(w, z) = R_N(w, z)$  by (4.20) and (4.16). This proves part (c) in case the reproducing property of (b) is satisfied. The other case follows similarly. □

Due to (4.17), (4.18), we call  $R_N(w, z)$  the *reproducing kernel* for the pairing (4.13).

From Lemma 4.6(c) we find in particular that the sum in (4.16) does not depend on the particular choice of factorization  $\mathbf{G}^{-1} = \mathbf{Q}^t \mathbf{P}$ . It does depend on  $\mathbf{G}$ , as can also be seen from the expression obtained from (4.14) and (4.16),

$$R_N(w, z) = \begin{pmatrix} w^{N-1} I_p & \cdots & w I_p & I_p \end{pmatrix} \mathbf{G}^{-1} \begin{pmatrix} I_p \\ z I_p \\ \vdots \\ z^{N-1} I_p \end{pmatrix}, \tag{4.21}$$

which clearly only depends on  $\mathbf{G}$ .

#### 4.4. Main theorem

Now we are ready for the main theorem in this section.

**Theorem 4.7.** *Assume the transition matrices  $T_m$  are  $p$ -periodic and that the above Assumptions 4.1 and 4.2 are satisfied. Then the multi-level particle system (4.2) is determinantal with correlation kernel  $K$  given by*

$$\begin{aligned}
 [K(m, px + j; m', py + i)]_{i,j=0}^{p-1} &= -\frac{\chi_{m>m'}}{2\pi i} \oint_{\gamma} A_{m',m}(z) z^{y-x} \frac{dz}{z} \\
 &+ \frac{1}{(2\pi i)^2} \oint_{\gamma} \oint_{\gamma} A_{m',L}(w) R_N(w, z) A_{0,m}(z) \frac{w^y}{z^{x+1} w^{M+N}} dz dw, \\
 &x, y \in \mathbb{Z}, 0 < m, m' < L, \tag{4.22}
 \end{aligned}$$

where  $R_N(w, z)$  is the reproducing kernel (4.16) built out of matrix valued biorthogonal polynomials associated with the weight  $W_{0,L}(z) = A_{0,L}(z)/z^{M+N}$  on  $\gamma$ .

*Proof.* We know that the particle system is determinantal with kernel given by (4.7).

The first term in (4.7) gives rise to the first term on the right-hand side of (4.22) in view of (4.10) (note that  $x$  and  $y$  are interchanged in  $T_{m',m}(y, x)$  in (4.7)).

Let  $K_0(m, x; m', y)$  be the second term on the right-hand side of (4.7). Instead of the summation indices  $i$  and  $j$  in the double sum we use  $pv + k$  and  $pv' + k'$  with  $0 \leq v, v' \leq N - 1, 0 \leq k, k' \leq p - 1$ . Then from (4.7),

$$\begin{aligned}
 &K_0(m, px + j; m', py + i) \\
 &= \sum_{v,v'=0}^{N-1} \sum_{k,k'=0}^{p-1} T_{m',L}(py + i, pM + pv' + k') [\mathbf{G}^{-1}]_{pv'+k', pv+k} T_{0,m}(pv + k, px + j). \tag{4.23}
 \end{aligned}$$

Using (4.10) we can write this in block form

$$\begin{aligned}
 [K_0(m, px + j; m', py + i)]_{i,j=0}^{p-1} &= \left( \frac{1}{2\pi i} \oint_{\gamma} A_{m',L}(w) w^{y-M-v'} \frac{dw}{w} \right)_{0 \leq v' \leq N-1} \mathbf{G}^{-1} \\
 &\cdot \left( \frac{1}{2\pi i} \oint_{\gamma} A_{0,m}(z) z^{v-x} \frac{dz}{z} \right)_{0 \leq v \leq N-1}^t \tag{4.24}
 \end{aligned}$$

where the first factor on the right-hand side of (4.24) is a block row vector of length  $N$  with  $p \times p$  blocks, and the last factor is a similar block column vector. We combine the integrals to obtain a double integral and then we use (4.21) to find the double integral in (4.22). This completes the proof.  $\square$

#### 4.5. Matrix valued orthogonal polynomials

The goal of this subsection and the next is to express the reproducing kernel (4.16) (or (4.21)) in terms of matrix valued orthogonal polynomials (MVOP) and use a Christoffel–Darboux formula for the sum (4.16). Such a formula is known for MVOP in various forms [4, 6, 25, 39]. We are however in a non-standard situation, with a non-Hermitian orthogonality, and the MVOP need not exist for every degree. Fortunately, the degree  $N$  MVOP will exist in the present situation, as we will explain in this subsection.

If we can find a factorization  $\mathbf{G}^{-1} = \mathbf{Q}'\mathbf{P}$  leading to matrix valued polynomials (4.14) with  $P_j(z) = Q_j(z)$  and  $\deg P_j = j$  for every  $j = 0, 1, \dots, N - 1$ , then we would have a finite sequence of MVOP that are in fact orthonormal:

$$\frac{1}{2\pi i} \oint_{\gamma} P_j(z)W_{0,L}(z)P_k^t(z) dz = \delta_{j,k}I_p, \quad j, k = 0, \dots, N - 1, \tag{4.25}$$

and then

$$R_N(w, z) = \sum_{j=0}^{N-1} P_j^t(w)P_j(z). \tag{4.26}$$

From (4.26) it would follow that  $R_N(z, w) = R_N^t(w, z)$  and this is an identity that is not necessarily satisfied. Thus we cannot expect that the orthonormal MVOP exist. Instead we focus on monic MVOP.

The orthogonality (4.25) is non-Hermitian orthogonality, and it is not associated with a positive definite scalar product. Also existence and uniqueness of the monic MVOP is not guaranteed in general. However, the MVOP of degree  $N$  does exist, and this is a consequence of the fact that  $\mathbf{G}$  is invertible.

**Lemma 4.8.** *There is a unique monic matrix valued polynomial  $P_N(z) = z^N I_p + \dots$  of degree  $N$  such that*

$$\frac{1}{2\pi i} \oint_{\gamma} P_N(z)W_{0,L}(z)z^k dz = 0_p, \quad k = 0, 1, \dots, N - 1. \tag{4.27}$$

*Proof.* The conditions (4.27) give us  $p^2N$  linear equations for the  $p^2N$  unknown coefficients of a monic matrix valued polynomial of degree  $N$ . The linear system has matrix  $\mathbf{G}$ , provided we number the coefficients and the conditions appropriately, and since  $\mathbf{G}$  is invertible, the existence and uniqueness of  $P_N$  follows.

More explicitly, write  $P_N(z) = I_p z^N + \sum_{j=0}^{N-1} C_j z^j$  with  $p \times p$  matrices  $C_j$  to be determined. The orthogonality conditions

$$\frac{1}{2\pi i} \oint_{\gamma} P_N(z)W_{0,L}(z)z^{N-1-k} dz = 0, \quad k = 0, 1, \dots, N - 1,$$

yield

$$\sum_{j=0}^{N-1} C_j \frac{1}{2\pi i} \oint_{\gamma} W_{0,L}(z)z^{N+j-k} \frac{dz}{z} = -\frac{1}{2\pi i} \oint_{\gamma} W_{0,L}(z)z^{2N-k} \frac{dz}{z}$$

for  $k = 0, 1, \dots, N - 1$ . Since  $W_{0,L}(z) = A_{0,L}(z)/z^{M+N}$ , the left-hand side is

$$\sum_{j=0}^{N-1} C_j \frac{1}{2\pi i} \oint_{\gamma} \frac{A_{0,L}(z)}{z^M} z^{j-k} \frac{dz}{z} = \sum_{j=0}^{N-1} C_j G_{j,k}$$

where  $G_{j,k}$  denotes the  $(j, k)$ th block of the block Toeplitz matrix  $\mathbf{G}$  (see also (4.11)). Varying  $k = 0, 1, \dots, N - 1$ , we see that

$$(C_0 \quad \dots \quad C_{N-1}) \mathbf{G} = -\frac{1}{2\pi i} \oint_{\gamma} W_{0,L}(z) z^{2N} (I_p \quad z^{-1}I_p \quad \dots \quad z^{-N+1}I_p) \frac{dz}{z}.$$

The matrix  $\mathbf{G}$  is invertible, and thus the matrix coefficients  $C_0, \dots, C_{N-1}$  are uniquely determined, and the monic MVOP of degree  $N$  exists uniquely.  $\square$

#### 4.6. Riemann–Hilbert problem and Christoffel–Darboux formula

The MVOP of degree  $N$  is characterized by a Riemann–Hilbert problem of size  $2p \times 2p$ . The RH problem asks for a  $2p \times 2p$  matrix valued function  $Y : \mathbb{C} \setminus \gamma \rightarrow \mathbb{C}^{2p \times 2p}$  satisfying

- $Y$  is analytic,
- $Y_+ = Y_- \begin{pmatrix} I_p & W_{0,L} \\ 0_p & I_p \end{pmatrix}$  on  $\gamma$  with counterclockwise orientation,
- $Y(z) = (I_{2p} + O(z^{-1})) \begin{pmatrix} z^N I_p & 0_p \\ 0_p & z^{-N} I_p \end{pmatrix}$  as  $z \rightarrow \infty$ .

In the scalar valued case, i.e.  $p = 1$ , the RH problem is due to Fokas, Its, and Kitaev [37]. The matrix valued extension can be found in [19, 29, 39]. It is similar to the RH problem for multiple orthogonal polynomials [64].

The RH problem has a unique solution, since by Lemma 4.8 the monic MVOP of degree  $P_N$  exists and is unique. The solution is

$$Y(z) = \begin{pmatrix} P_N(z) & \frac{1}{2\pi i} \oint_{\gamma} \frac{P_N(s)W_{0,L}(s)}{s-z} ds \\ Q_{N-1}(z) & \frac{1}{2\pi i} \oint_{\gamma} \frac{Q_{N-1}(s)W_{0,L}(s)}{s-z} ds \end{pmatrix}, \quad z \in \mathbb{C} \setminus \gamma, \tag{4.28}$$

where  $Q_{N-1}$  is a matrix valued polynomial of degree  $\leq N - 1$  such that

$$\frac{1}{2\pi i} \oint_{\gamma} Q_{N-1}(z)W_{0,L}(z)z^k dz = \begin{cases} 0_p, & k = 0, 1, \dots, N - 2, \\ -I_p, & k = N - 1. \end{cases} \tag{4.29}$$

One can show that  $Q_{N-1}$  also uniquely exists, since the conditions (4.29) give a system of  $p^2N$  linear equations for the  $p^2N$  coefficients of  $Q_{N-1}$ , and the matrix of this system can be identified with  $\mathbf{G}$ . Since  $\mathbf{G}$  is invertible, there is a unique solution. If the leading coefficient of  $Q_{N-1}$  would be invertible (which is typically the case, but it is not guaranteed in general) then the monic MVOP  $P_{N-1}$  of degree  $N - 1$  would exist as well and  $Q_{N-1} = -H_{N-1}^{-1}P_{N-1}$  for some invertible  $H_{N-1}$ .

In the following result we express the reproducing kernel  $R_N(w, z)$  in terms of the solution of the RH problem. It can be viewed as a Christoffel–Darboux formula and in this form it is due to Delvaux [29]. It is similar to the Christoffel–Darboux formulas for multiple orthogonal polynomials [12, 24] which also use the RH problem.

**Proposition 4.9.** *We have*

$$R_N(w, z) = \frac{1}{z-w} \begin{pmatrix} 0_p & I_p \end{pmatrix} Y^{-1}(w)Y(z) \begin{pmatrix} I_p \\ 0_p \end{pmatrix}. \quad (4.30)$$

*Proof.* This is due to Delvaux [29, Proposition 1.10] (see also [39]).  $\square$

We insert (4.30) into formula (4.22) and find a convenient formula for the correlation kernel in terms of the solution of the RH problem. A possible asymptotic analysis of the kernel would consist of two parts. First we do an analysis of the RH problem that would give us the asymptotic behavior of the kernel (4.30). Then this is followed by an asymptotic analysis of the double contour integral in (4.22) by means of classical methods of steepest descent. We are able to do this for the two-periodic Aztec diamond.

#### 4.7. Two-periodic Aztec diamond

We saw in Section 3 that the two-periodic Aztec diamond gives rise to the multi-level particle system (3.9). This satisfies Assumption 4.1 if we take  $p = 2$  and  $L = 2N$  and  $M = 0$ . The transition matrices are independent of  $m$  (see (3.6)) and the matrix symbol is given by (3.7). The weight matrix is  $W_{0,L}(z) = A_{0,L}(z)/z^{M+N} = W^N(z)$  with

$$\begin{aligned} W(z) &= \frac{A^2(z)}{z} = \frac{1}{z(z-1)^2} \begin{pmatrix} 2\alpha z & \alpha(z+1) \\ \beta z(z+1) & 2\beta z \end{pmatrix}^2 \\ &= \frac{1}{(z-1)^2} \begin{pmatrix} (z+1)^2 + 4\alpha^2 z & 2\alpha(\alpha+\beta)(z+1) \\ 2\beta(\alpha+\beta)z(z+1) & (z+1)^2 + 4\beta^2 z \end{pmatrix} \end{aligned} \quad (4.31)$$

as in (2.15) and (2.19). Observe that  $W$  has no pole at the origin.

Theorem 4.7 applies and it gives the form of the correlation kernel, in  $2 \times 2$  matrix form, that will be stated in (5.4) below. It is equivalent to the form already announced in (2.21).

The correlation kernel contains the reproducing kernel  $R_N(w, z)$  with respect to the varying weight  $W^N$ , and by Proposition 4.9 the latter kernel is expressed in terms of the RH problem for the MVOP of degree  $N$ . By Lemma 4.8 we conclude that the degree  $N$  monic MVOP with respect to the weight  $W^N$  exists. While it does not matter for what follows, one can show that the MVOP of lower degrees also exist [32, Lemma 4.10].

In the next section we continue with the analysis of the RH problem and we show that the correlation kernel (2.21) can be rewritten as (2.8).

## 5. Analysis of the RH problem

We consider an Aztec diamond of size  $2N$  with two-periodic weighting.

### 5.1. Correlation kernel

In the two-periodic Aztec diamond we find the matrix symbol

$$A(z) = \frac{1}{z-1} \begin{pmatrix} 2\alpha z & \alpha(z+1) \\ \beta z(z+1) & 2\beta z \end{pmatrix}, \quad \alpha\beta = 1, \quad (5.1)$$

and the matrix valued weight is  $W^N$  with  $W$  given by (4.31). Note that  $W^N$  is a rational function with a pole at  $z = 1$  only.

The contour  $\gamma$  in the RH problem from Section 4.6 goes around 0 and lies in the domain  $|z| > 1$ . By analyticity, since  $W$  only has a pole at  $z = 1$ , we are free to deform the contour to a circle around 1. We use  $\gamma_1$  to denote the circle of radius  $r < 1$  around 1. We obtain the following RH problem for  $Y : \mathbb{C} \setminus \gamma_1 \rightarrow \mathbb{C}^{4 \times 4}$ :

- $Y$  is analytic,
- $Y$  has the jump

$$Y_+(z) = Y_-(z) \begin{pmatrix} I_2 & W^N(z) \\ 0_2 & I_2 \end{pmatrix}, \quad z \in \gamma_1, \tag{5.2}$$

- $Y$  has asymptotic behavior

$$Y(z) = (I_4 + O(z^{-1})) \begin{pmatrix} z^N I_2 & 0_2 \\ 0_2 & z^{-N} I_2 \end{pmatrix} \quad \text{as } z \rightarrow \infty. \tag{5.3}$$

Because of Theorem 4.7 and (4.30) we find the following correlation kernel for arbitrary integer levels  $m, m'$  with  $0 < m, m' < 2N = L$  and  $M = 0$ :

$$\begin{aligned} & \begin{pmatrix} K_N(m, 2x; m', 2y) & K_N(m, 2x + 1; m', 2y) \\ K_N(m, 2x; m', 2y + 1) & K_N(m, 2x + 1; m', 2y + 1) \end{pmatrix} \\ &= -\frac{\chi_{m > m'}}{2\pi i} \oint_{\gamma} A^{m-m'}(z) z^{y-x-1} dz \\ &+ \frac{1}{(2\pi i)^2} \oint_{\gamma_{0,1}} \oint_{\gamma_{0,1}} A^{2N-m'}(w) \begin{pmatrix} 0_2 & I_2 \end{pmatrix} Y^{-1}(w) Y(z) \begin{pmatrix} I_2 \\ 0_2 \end{pmatrix} A^m(z) \frac{w^y}{z^x w^N} \frac{dz dw}{z(z-w)}. \end{aligned} \tag{5.4}$$

The contour  $\gamma_{0,1}$  in (5.4) is a circle of radius  $> 1 + r$  around the origin, as before. The radius is large enough for  $\gamma_1$  to lie inside  $\gamma_{0,1}$ .

The analysis of the correlation kernel (5.4) consists of two parts. First we apply a RH analysis to the RH problem for  $Y$  and then we use this for an asymptotic analysis of the double integral. The RH analysis is remarkably simple. It is not an asymptotic analysis, since the outcome is an exact new formula for the correlation kernel.

**Theorem 5.1.** *Assume  $2y \geq m'$  and  $N$  is even. Then the correlation kernel (5.4) is equal to*

$$\begin{aligned} & -\frac{\chi_{m > m'}}{2\pi i} \oint_{\gamma_{0,1}} A^{m-m'}(z) z^{y-x} \frac{dz}{z} \\ & + \frac{1}{(2\pi i)^2} \oint_{z \in \gamma_{0,1}} \oint_{w \in \gamma_1} A^{N-m'}(w) F(w) A^{-N+m}(z) \frac{z^{N/2}}{w^{N/2}} \frac{(z-1)^N}{(w-1)^N} \frac{w^y}{z^x} \frac{dz dw}{z(z-w)} \end{aligned} \tag{5.5}$$

where  $F$  is given by (2.10).

Passing from the non-intersecting path model back to the domino tilings of the Aztec diamond, we should make the change of variables  $m \mapsto m, 2x \mapsto m + n, m' \mapsto m'$  and  $2y \mapsto m' + n'$ . Inserting these values in (5.5) we obtain the correlation kernel (2.8) and so Theorem 2.2 follows immediately from Theorem 5.1.

The rest of Section 5 is devoted to the proof of Theorem 5.1. We follow the general scheme of the analysis of RH problems, known as the Deift–Zhou steepest descent analysis [28], which was first applied to orthogonal polynomials in [26, 27]. Extensions to larger size RH problems are for example in [13, 31]; see also the survey [51] and the references therein. However, the RH analysis in this section is not an asymptotic analysis, as it produces the exact formula (5.5).

### 5.2. Eigenvalues and eigenvectors on the Riemann surface

We use the eigenvalues  $\rho_{1,2}$  of  $\begin{pmatrix} 2\alpha z & \alpha(z+1) \\ \beta z(z+1) & 2\beta z \end{pmatrix}$  and the eigenvalues  $\lambda_{1,2}$  of  $W$  as already introduced in (2.12) and (2.17). The corresponding eigenvectors are in the columns of the matrix

$$E(z) = \begin{pmatrix} \alpha(z+1) & \alpha(z+1) \\ \rho_1(z) - 2\alpha z & \rho_2(z) - 2\alpha z \end{pmatrix}, \tag{5.6}$$

and we have the decompositions (2.13), (2.14) and (2.16).

The eigenvalues and eigenvectors are defined and analytic in the complex plane cut along the two intervals  $(-\infty, -\alpha^2]$  and  $[-\beta^2, 0]$  where we have  $\lambda_{1,\pm} = \lambda_{2,\mp}, \rho_{1,\pm} = \rho_{2,\mp}$ , and

$$E_+ = E_- \sigma_1 \quad \text{on } (-\infty, -\alpha^2] \cup [-\beta^2, 0] \tag{5.7}$$

with  $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ .

As already mentioned in Remark 2.3, we use the two-sheeted Riemann surface  $\mathcal{R}$  associated with the equation (2.11). The Riemann surface has genus 1, unless  $\alpha = \beta = 1$ , in which case the genus is zero.

We use  $z$  for a generic coordinate on  $\mathcal{R}$ , and if we want to emphasize that  $z$  is on the  $j$ th sheet, we write  $z^{(j)}$  for  $j = 1, 2$ . We write  $\lambda$  for the function on  $\mathcal{R}$  given by

$$\lambda(z) = \lambda_j(z) \quad \text{if } z = z^{(j)} \text{ is on the } j\text{th sheet} \tag{5.8}$$

(see (2.17)), and similarly for  $\rho$ . These are meromorphic on  $\mathcal{R}$ , namely  $\rho = (\alpha + \beta)z + y$  and  $\lambda = \frac{\rho^2}{z(z-1)^2}$  (see (2.11), (2.12) and (2.17)).

**Lemma 5.2.** (a)  $\rho$  has a simple zero at  $z = 0$ , a double zero at  $z = 1^{(2)}$  (the point  $z = 1$  on the second sheet), a triple pole at  $z = \infty$ , and no other zeros or poles.

(b)  $\lambda$  has a double zero at  $z = 1^{(2)}$ , a double pole at  $z = 1^{(1)}$ , and no other zeros or poles.

(c) The function

$$\rho(z) - 2\alpha z = (\beta - \alpha)z + y \tag{5.9}$$

has a zero at  $z = 0$ , and a double zero at  $z = -1^{(1)}$  (if  $\alpha > \beta$ ).

(d)  $\lambda_1(z)\lambda_2(z) = 1$  for every  $z$  and  $\lambda(\infty) = 1$ .

(e) For real  $x$  we have

$$|\lambda_{1,\pm}(x)| = |\lambda_{2,\pm}(x)| = 1, \quad x \in (-\infty, -\alpha^2] \cup [-\beta^2, 0], \quad (5.10)$$

$$\lambda_1(x) > 1 > \lambda_2(x) > 0, \quad x \in (0, \infty), \quad (5.11)$$

$$\lambda_1(x) < -1 < \lambda_2(x) < 0, \quad x \in (-\alpha^2, -\beta^2). \quad (5.12)$$

(f)  $|\lambda_1(z)| > |\lambda_2(z)|$  for every  $z \in \mathbb{C} \setminus ((-\infty, -\alpha^2] \cup [-\beta^2, 0])$ .

*Proof.* Parts (a), (b), and (c) are easy to verify from the definitions. We note that part (d) comes from the fact that

$$\det W(z) = 1 \quad (5.13)$$

for every  $z$ , which follows from (4.31) by a direct calculation, and therefore  $\lambda_1(z)\lambda_2(z) = \det W(z) = 1$  for every  $z$ . Also from (4.31),

$$\lim_{z \rightarrow \infty} W(z) = \begin{pmatrix} 1 & 0 \\ 2\alpha(\alpha + \beta) & 1 \end{pmatrix} =: W_\infty \quad \text{as } z \rightarrow \infty \quad (5.14)$$

and so for its eigenvalues we have  $\lambda_{1,2}(z) \rightarrow 1$  as  $z \rightarrow \infty$ .

For  $x \in (-\infty, -\alpha^2] \cup [-\beta^2, 0]$  we have  $\lambda_{1,\pm}(x) = \lambda_{2,\mp}(x)$  and  $\lambda_{1,\pm}(x) = 1/\lambda_{2,\pm}(x)$  because of part (d). Then the identity  $\lambda_{1,+}(x)\lambda_{1,-}(x) = \lambda_{2,+}(x)\lambda_{2,-}(x) = 1$  follows, which gives (5.10).

The functions  $\log |\lambda_1|$  and  $\log |\lambda_2|$  are harmonic on  $\mathbb{C} \setminus ((-\infty, -\alpha^2] \cup [-\beta^2, 0] \cup \{1\})$ , they are both zero on  $(-\infty, -\alpha^2] \cup [-\beta^2, 0]$ , have the value 1 at infinity, while  $\lim_{z \rightarrow 1} \log |\lambda_1(z)| = +\infty$ ,  $\lim_{z \rightarrow 1} \log |\lambda_2(z)| = -\infty$  because of part (b). Then by the minimum principle for harmonic functions,  $\log |\lambda_1(z)| > \log |\lambda_2(z)|$  for every  $z \in \mathbb{C} \setminus ((-\infty, -\alpha^2] \cup [-\beta^2, 0])$ . This establishes part (f), and also the inequalities (5.11) and (5.12) of part (e) since  $\lambda_1(x)$  and  $\lambda_2(x)$  are real and positive for  $x \in (0, \infty)$  and real and negative for  $x \in (-\alpha^2, -\beta^2)$  (see (2.17)).  $\square$

### 5.3. First transformation $Y \mapsto X$ of the RH problem

We use the matrix of eigenvalues (5.6) in the first transformation of the RH problem. We define

$$X = Y \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}, \quad (5.15)$$

which satisfies the following RH problem.

- $X$  is analytic on  $\mathbb{C} \setminus (\gamma_1 \cup (-\infty, -\alpha^2] \cup [-\beta^2, 0])$ ,
- $X$  has the jumps

$$X_+ = \begin{cases} X_- \begin{pmatrix} I_2 & \Lambda^N \\ 0_2 & I_2 \end{pmatrix} & \text{on } \gamma_1, \\ X_- \begin{pmatrix} \sigma_1 & 0_2 \\ 0_2 & \sigma_1 \end{pmatrix} & \text{on } (-\infty, -\alpha^2] \cup [-\beta^2, 0], \end{cases} \quad (5.16)$$

- $X$  has asymptotic behavior

$$X(z) = (I_4 + O(1/z)) \begin{pmatrix} z^N E(z) & 0_2 \\ 0_2 & z^{-N} E(z) \end{pmatrix} \quad \text{as } z \rightarrow \infty. \tag{5.17}$$

This is easy to verify from the RH problem for  $Y$ , the definition (5.15), and the properties (2.16) and (5.7).

Since  $\det Y(z) = 1$  and

$$\det E(z) = -2\alpha(z + 1)\sqrt{z(z + \alpha^2)(z + \beta^2)}, \tag{5.18}$$

which is easy to check from (5.6), we have by (5.15)

$$\det X = (\det E)^2 = 4\alpha^2 z(z + 1)^2(z + \alpha^2)(z + \beta^2). \tag{5.19}$$

### 5.4. Second transformation $X \mapsto U$

Remarkably, we do not need equilibrium measures or  $g$ -functions for the next transformation.

From Lemma 5.2(b) we know that both  $z \mapsto (z - 1)^2\lambda_1(z)$  and  $z \mapsto (z - 1)^{-2}\lambda_2(z)$  have a removable singularity at  $z = 1$ , and hence they are analytic in  $\mathbb{C} \setminus ((-\infty, -\alpha^2] \cup [-\beta^2, 0])$  without any zeros. We recall that  $N$  is even and we put

$$U = LX \times \begin{cases} \text{diag}\left(\frac{\lambda_1^{N/2}}{(z - 1)^N}, \frac{\lambda_2^{N/2}}{(z - 1)^N}, \frac{(z - 1)^N}{\lambda_1^{N/2}}, \frac{(z - 1)^N}{\lambda_2^{N/2}}\right) & \text{for } |z - 1| > r, \\ \text{diag}\left((z - 1)^N \lambda_1^{N/2}, \frac{\lambda_2^{N/2}}{(z - 1)^N}, \frac{1}{(z - 1)^N \lambda_1^{N/2}}, \frac{(z - 1)^N}{\lambda_2^{N/2}}\right) & \text{for } |z - 1| < r, \end{cases} \tag{5.20}$$

where  $L$  is the constant matrix

$$L = \begin{pmatrix} W_\infty^{-N/2} & 0_2 \\ 0_2 & W_\infty^{N/2} \end{pmatrix} \tag{5.21}$$

with  $W_\infty$  as in (5.14). Then  $U$  is defined on  $\mathbb{C} \setminus (\gamma_1 \cup (-\infty, -\alpha^2] \cup [-\beta^2, 0])$ , and from the definition (5.20) and the RH problem for  $X$  we obtain

- $U$  is analytic,
- $U$  has the jumps

$$U_+ = \begin{cases} U_- \begin{pmatrix} (z - 1)^{2N} & 0 & 1 & 0 \\ 0 & 1 & 0 & (z - 1)^{2N} \\ 0 & 0 & (z - 1)^{-2N} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \text{on } \gamma_1, \\ U_- \begin{pmatrix} \sigma_1 & 0_2 \\ 0_2 & \sigma_1 \end{pmatrix} & \text{on } (-\infty, -\alpha^2] \cup [-\beta^2, 0], \end{cases} \tag{5.22}$$

- $U$  has asymptotic behavior

$$\begin{aligned}
 U(z) &= L(I_4 + O(1/z)) \begin{pmatrix} E(z)\Lambda^{N/2}(z) & 0_2 \\ 0_2 & E(z)\Lambda^{-N/2}(z) \end{pmatrix} \\
 &= (I_4 + O(1/z)) \begin{pmatrix} E(z) & 0_2 \\ 0_2 & E(z) \end{pmatrix} \quad \text{as } z \rightarrow \infty.
 \end{aligned} \tag{5.23}$$

To obtain the jump (5.22) on  $(-\infty, -\alpha^2] \cup [-\beta^2, 0]$  we also have to use the fact that  $\lambda_{1,\pm} = \lambda_{2,\mp}$  on these cuts.

The asymptotic condition (5.23) requires some explanation. The first equality in (5.23) is clear from the definition (5.20) of  $U$  for  $|z-1| > r$ , and the asymptotic behavior (5.16) of  $X$ . By (2.16) we have  $E(z)\Lambda^{\pm N/2}(z) = W^{\pm N/2}(z)E(z)$ , so that by (4.31), (5.14), and (5.21) we get

$$\begin{aligned}
 L \begin{pmatrix} E(z)\Lambda^{N/2}(z) & 0_2 \\ 0_2 & E(z)\Lambda^{-N/2}(z) \end{pmatrix} &= L \begin{pmatrix} W^{N/2}(z)E(z) & 0_2 \\ 0_2 & W^{-N/2}(z)E(z) \end{pmatrix} \\
 &= (I_4 + O(1/z)) \begin{pmatrix} E(z) & 0_2 \\ 0_2 & E(z) \end{pmatrix},
 \end{aligned}$$

and this leads to the second equality in (5.23).

### 5.5. Third transformation $U \mapsto T$

In the third transformation we turn the entries  $(z-1)^{\pm 2N}$  in the jump matrix on  $\gamma_1$  into an off-diagonal entry. It corresponds to the opening of lenses in a steepest descent analysis. We also remove the 24-entry in the jump matrix on  $\gamma_1$ .

We define

$$T(z) = \begin{cases} U(z) & \text{for } |z-1| > r, \\ U(z) \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -(z-1)^{2N} \\ 1 & 0 & (z-1)^{2N} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \text{for } |z-1| < r. \end{cases} \tag{5.24}$$

Straightforward calculations, where we just use (5.24) and the RH problem for  $U$ , show that  $T$  satisfies

- $T : \mathbb{C} \setminus (\gamma_1 \cup (-\infty, -\alpha^2] \cup [-\beta^2, 0]) \rightarrow \mathbb{C}^{4 \times 4}$  is analytic,
- $T$  has the jumps

$$T_{\pm} = \begin{cases} T_{-} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ (z-1)^{-2N} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \text{on } \gamma_1, \\ T_{-} \begin{pmatrix} \sigma_1 & 0_2 \\ 0_2 & \sigma_1 \end{pmatrix} & \text{on } (-\infty, -\alpha^2] \cup [-\beta^2, 0], \end{cases} \tag{5.25}$$

- $T$  has asymptotic behavior

$$T(z) = (I_4 + O(1/z)) \begin{pmatrix} E(z) & 0_2 \\ 0_2 & E(z) \end{pmatrix} \quad \text{as } z \rightarrow \infty. \tag{5.26}$$

5.6. Fourth transformation  $T \mapsto S$

We next remove the jumps on the negative real axis. We use  $\begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}$  as global parametrix, since it has the same jump on  $(-\infty, -\alpha^2] \cup [-\beta^2, 0]$  as  $T$  has (see (5.25)). We define

$$S = T \begin{pmatrix} E^{-1} & 0_2 \\ 0_2 & E^{-1} \end{pmatrix}, \tag{5.27}$$

and then  $S$  has no jump on  $(-\infty, -\alpha^2) \cup (-\beta^2, 0)$ , that is,  $S_+ = S_-$  on these two intervals.

Since  $E$  is not invertible at  $z = 0, z = -\alpha^2, z = -\beta^2$  (see also (5.18)), we could have introduced singularities at these points. Therefore we look at the combined transformations  $Y \mapsto X \mapsto U \mapsto T \mapsto S$  in order to express  $S$  directly in terms of  $Y$ . For  $z$  outside  $\gamma_1$  we have, by (5.15), (5.20), (5.24) and (5.27),

$$\begin{aligned} S &= U \begin{pmatrix} E^{-1} & 0_2 \\ 0_2 & E^{-1} \end{pmatrix} = LX \begin{pmatrix} (z-1)^{-N} \Lambda^{N/2} E^{-1} & 0_2 \\ 0_2 & (z-1)^N \Lambda^{-N/2} E^{-1} \end{pmatrix} \\ &= LY \begin{pmatrix} (z-1)^{-N} E \Lambda^{N/2} E^{-1} & 0_2 \\ 0_2 & (z-1)^N E \Lambda^{-N/2} E^{-1} \end{pmatrix}. \end{aligned}$$

Since  $E \Lambda E^{-1} = W$  by (2.16), we simply have (recall  $N$  is even)

$$S = LY \begin{pmatrix} (z-1)^{-N} W^{N/2} & 0_2 \\ 0_2 & (z-1)^N W^{-N/2} \end{pmatrix}, \quad |z-1| > r. \tag{5.28}$$

This shows indeed that (5.27) does not introduce any singularities, since  $\det W(z) = 1$  for every  $z$ , and  $W(z)$  and  $W^{-1}(z)$  have poles at  $z = 1$  only.

Thus  $S$  has analytic continuation across  $(-\infty, -\alpha^2]$  and  $[-\beta^2, 0)$  and satisfies the following RH problem that we obtain from (5.27) and the RH problem for  $T$ .

- $S : \mathbb{C} \setminus \gamma_1 \rightarrow \mathbb{C}^{4 \times 4}$  is analytic,
- $S$  has the jump

$$S_+(z) = S_-(z) \begin{pmatrix} I_2 & 0_2 \\ (z-1)^{-2N} F(z) & I_2 \end{pmatrix} \quad \text{for } z \in \gamma_1, \tag{5.29}$$

where  $F(z) = E(z) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} E^{-1}(z)$  is as in (2.10) and (2.14),

- $S$  has asymptotic behavior  $S(z) = I_4 + O(1/z)$  as  $z \rightarrow \infty$ .

The RH problem is now normalized at infinity. Note also that the transformation (5.27) restores the property  $\det S = 1$ , since  $\det T = \det X = (\det E)^2$  (see (5.19)).

The jump matrix in (5.29) is lower triangular, and the RH problem for  $S$  is normalized at infinity. This means that we can solve the RH problem explicitly by a contour integral. We find

$$S(z) = \begin{pmatrix} I_2 & 0_2 \\ \frac{1}{2\pi i} \oint_{\gamma_1} \frac{F(s)}{(s-1)^{2N}(s-z)} ds & I_2 \end{pmatrix}, \quad z \in \mathbb{C} \setminus \gamma_1. \tag{5.30}$$

5.7. Proof of Theorem 5.1

We analyze the effect of the transformations on the correlation kernel (5.4). From (5.28) and (2.15) we have, for  $z, w$  outside  $\gamma_1$ ,

$$\begin{aligned} & (0_2 \quad I_2) Y^{-1}(w)Y(z) \begin{pmatrix} I_2 \\ 0_2 \end{pmatrix} \\ &= (w-1)^N(z-1)^N W^{-N/2}(w) (0_2 \quad I_2) S^{-1}(w)S(z) \begin{pmatrix} I_2 \\ 0_2 \end{pmatrix} W^{-N/2}(z) \\ &= (w-1)^N(z-1)^N w^{N/2}z^{N/2} A^{-N}(w) (0_2 \quad I_2) S^{-1}(w)S(z) \begin{pmatrix} I_2 \\ 0_2 \end{pmatrix} A^{-N}(z). \end{aligned}$$

Thus

$$\begin{aligned} & A^{2N-m'}(w) (0_2 \quad I_2) Y^{-1}(w)Y(z) \begin{pmatrix} I_2 \\ 0_2 \end{pmatrix} A^m(z)w^{-N} \\ &= (w-1)^N(z-1)^N w^{-N/2}z^{N/2} A^{N-m'}(w) (0_2 \quad I_2) S^{-1}(w)S(z) \begin{pmatrix} I_2 \\ 0_2 \end{pmatrix} A^{-N+m}(z), \end{aligned} \tag{5.31}$$

which is part of the expression that appears in the double integral in (5.4). Because of (5.30) and

$$S^{-1}(w) = \begin{pmatrix} I_2 & 0_2 \\ -\frac{1}{2\pi i} \oint_{\gamma_1} \frac{F(s)}{(s-1)^{2N}(s-w)} ds & I_2 \end{pmatrix}$$

we have

$$\begin{aligned} & (0_2 \quad I_2) S^{-1}(w)S(z) \begin{pmatrix} I_2 \\ 0_2 \end{pmatrix} \\ &= \frac{1}{2\pi i} \oint_{\gamma_1} \frac{F(s)}{(s-1)^{2N}(s-z)} ds - \frac{1}{2\pi i} \oint_{\gamma_1} \frac{F(s)}{(s-1)^{2N}(s-w)} ds \\ &= -\frac{z-w}{2\pi i} \oint_{\gamma_1} \frac{F(s)}{(s-1)^{2N}(s-w)(s-z)} ds. \end{aligned} \tag{5.32}$$

Using (5.31) and (5.32) we see that the double integral in (5.4) is equal to

$$-\frac{1}{(2\pi i)^2} \oint_{\gamma_{0,1}} \oint_{\gamma_{0,1}} A^{N-m'}(w) \left( \frac{1}{2\pi i} \oint_{\gamma_1} \frac{F(s)}{(s-1)^{2N}(s-w)(s-z)} ds \right) \times A^{-N+m}(z) \frac{z^{N/2}}{w^{N/2}} (w-1)^N (z-1)^N \frac{w^y}{z^{x+1}} dz dw. \tag{5.33}$$

We change the order of integration in (5.33) and evaluate the  $w$ -integral first. By a residue calculation,

$$\frac{1}{2\pi i} \oint_{\gamma_{0,1}} A^{N-m'}(w)(w-1)^N w^{y-N/2} \frac{1}{w-s} dw = A^{N-m'}(s)(s-1)^N s^{y-N/2}, \quad s \in \gamma_1. \tag{5.34}$$

Indeed, the singularities at  $w = 0$  and  $w = 1$  in the integrand on the left-hand side of (5.34) are removable (we use (5.1),  $N$  is even, and  $2y \geq m'$ ). The only singularity is at  $w = s$  and (5.34) indeed follows by Cauchy's formula since  $s \in \gamma_1$  lies inside  $\gamma_{0,1}$ .

Using (5.34) in (5.33) and changing the integration variable  $s$  to  $w$ , we obtain the double integral in (5.5). The single integral in (5.5) is of course immediate from (5.4). This completes the proof of Theorem 5.1. □

### 6. Asymptotic analysis

In the final section of the paper we analyze the formula (2.8) in a scaling limit where  $N \rightarrow \infty$  and the coordinates  $(m, n)$  and  $(m', n')$  scale linearly with  $N$ . We are going to distinguish the three phases of the model, and prove Theorems 2.8, 2.10, and 2.11.

#### 6.1. Preliminaries

We first rewrite the formula (2.8) in a form that already contains the gas phase kernel (2.28) and double integrals with the phase functions  $\Phi_1$  and  $\Phi_2$  from (2.22) (see Corollary 6.3). We may and do assume that the contour  $\gamma_{0,1}$  is a contour in  $\mathbb{C} \setminus ((-\infty, -\alpha^2] \cup [-\beta^2, 0])$  going around the interval  $[-\beta^2, 0]$  once in positive direction.

**Lemma 6.1.** *The integrand in the double integral in (2.8) is  $O(w^{-N+n'/2-1/2})$  as  $w \rightarrow \infty$ .*

*Proof.* From the formulas (2.9)–(2.10) we easily get  $A(w) = O(w)$ ,  $F(w) = O(w^{1/2})$ ,  $A^2(w) = O(w)$  and  $A(w)F(w) = O(w)$  as  $w \rightarrow \infty$ . This implies that  $A^{N-m'}(w)F(w) = O(w^{(N-m'+1)/2})$  as  $w \rightarrow \infty$ . Use this in the integrand in (2.8) and the lemma follows. □

Note that  $n'$  could go up to  $2N - 1$ , and then  $O(w^{-N+n'/2-1/2}) = O(w^{-1})$ . However, then we are close to the boundary of the Aztec diamond, and we do not consider this in what follows, since we focus on the gas phase. So we assume  $n' \leq 2N - 2$ , and then the integrand in the double integral in (2.8) is  $O(w^{-3/2})$  as  $w \rightarrow \infty$ .

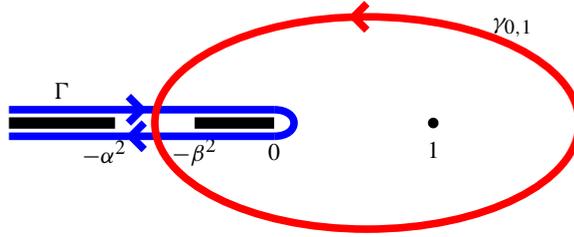


Fig. 12. Contours  $\gamma_{0,1}$  and  $\Gamma$ .

Then for a fixed  $z \in \gamma_{0,1}$ , we deform the contour  $\gamma_1$  to a contour  $\Gamma$  going around the negative real axis, starting at  $-\infty$  in the upper half-plane and ending at  $-\infty$  in the lower half-plane, as in Figure 12. Since the integrand is  $O(w^{-3/2})$ , by Lemma 6.1, there is no contribution from infinity, but there is a residue contribution from the pole at  $w = z$ . These residues combine to give the  $z$ -integral (we use the fact that  $F(z)$  and  $A(z)$  commute)

$$\frac{1}{2\pi i} \oint_{\gamma_{0,1}} F(z)A^{m-m'}(z) \frac{z^{(m'+n')/2}}{z^{(m+n)/2}} \frac{dz}{z}$$

Together with the single integral in (2.8) this gives the limit (2.28) that we expect to get in the gas phase. We proved the following.

**Proposition 6.2.** *Suppose  $N$  is even and  $(m, n), (m', n') \in \mathcal{B}_N$  with  $m + n$  and  $m' + n'$  even and  $n' \leq 2N - 2$ . Then*

$$\begin{aligned} \mathbb{K}_N(m, n; m', n') &= \mathbb{K}_{\text{gas}}(m, n; m', n') \\ &+ \frac{1}{(2\pi i)^2} \oint_{\gamma_{0,1}} \frac{dz}{z} \int_{\Gamma} \frac{dw}{z-w} A^{N-m'}(w)F(w)A^{-N+m}(z) \frac{z^{N/2}(z-1)^N}{w^{N/2}(w-1)^N} \frac{w^{(m'+n')/2}}{z^{(m+n)/2}}. \end{aligned} \tag{6.1}$$

Thus to establish Theorem 2.8 we have to prove that in the gas phase the double integral in (6.1) tends to 0 as  $N \rightarrow \infty$  at an exponential rate.

We can rewrite (6.1) where we assume that  $m = (1 + \xi_1)N, n = (1 + \xi_2)N, m' = (1 + \xi'_1)N$  and  $n' = (1 + \xi'_2)N$ . We use  $\Phi_1$  and  $\Phi_2$  as in (2.22), and to emphasize that these functions depend on  $\xi_1$  and  $\xi_2$ , we write  $\Phi_1(z; \xi_1, \xi_2)$  and  $\Phi_2(z; \xi_1, \xi_2)$ .

**Corollary 6.3.** *Suppose  $m = (1 + \xi_1)N, n = (1 + \xi_2)N, m' = (1 + \xi'_1)N$ , and  $n' = (1 + \xi'_2)N$  with  $-1 < \xi_1, \xi_2, \xi'_1, \xi'_2 < 1$ . Assume  $N, m + n$  and  $m' + n'$  are even. Then*

$$\begin{aligned} \mathbb{K}_N(m, n; m', n') &= \mathbb{K}_{\text{gas}}(m, n; m', n') \\ &+ \frac{1}{(2\pi i)^2} \oint_{\gamma_{0,1}} \frac{dz}{z} \int_{\Gamma} \frac{dw}{z-w} F(w)F(z)e^{N(\Phi_1(z;\xi_1,\xi_2)-\Phi_1(w;\xi'_1,\xi'_2))/2} \\ &+ \frac{1}{(2\pi i)^2} \oint_{\gamma_{0,1}} \frac{dz}{z} \int_{\Gamma} \frac{dw}{z-w} F(w)(I_2 - F(z))e^{N(\Phi_2(z;\xi_1,\xi_2)-\Phi_1(w;\xi'_1,\xi'_2))/2}, \end{aligned} \tag{6.2}$$

with contours as in Figure 12.

*Proof.* Because of (2.13)–(2.14) we have  $A^{N-m'}(w)F(w) = F(w)\left(\frac{\rho_1(w)}{w-1}\right)^{N-m'}$ . Thus in view of (2.17) and (2.22) we get

$$\begin{aligned} A^{N-m'}(w)F(w)\frac{w^{(m'+n')/2}}{w^{N/2}(w-1)^N} &= F(w)\lambda_1^{(N-m')/2}(w)\frac{w^{n'/2}}{(w-1)^N} \\ &= F(w)e^{-N\Phi_1(w;\xi'_1,\xi'_2)/2}, \end{aligned} \tag{6.3}$$

and similarly

$$A^{-N+m}(z)\frac{z^{N/2}(z-1)^N}{z^{(m+n)/2}} = F(z)e^{N\Phi_1(z;\xi_1,\xi_2)/2} + (I_2 - F(z))e^{N\Phi_2(z;\xi_1,\xi_2)/2}. \tag{6.4}$$

Using (6.3) and (6.4) in (6.1) we arrive at (6.2). □

### 6.2. Saddle points

The large  $N$  behavior of the  $z$ -integrals in (6.2) is dominated by the factors  $e^{N\Phi_1(z)}$  and  $e^{N\Phi_2(z)/2}$  that are exponential in  $N$ . Similarly the  $w$  part is dominated by  $e^{-N\Phi_1(w)/2}$ .

We study the saddle points, which in Definition 2.6 were already introduced as the zeros of the meromorphic differential  $\Phi'(z)dz$  from (2.23) defined on the Riemann surface  $\mathcal{R}$  associated with (2.11). Of course,  $\Phi$  depends on  $\xi_1, \xi_2$ , and thus the saddle points depend on these parameters. Throughout we restrict to  $-1 < \xi_1, \xi_2 < 1$ . The differential has simple poles at  $1^{(1)}, 1^{(2)}, 0$  and  $\infty$  with residues given in the following table.

pole	residue of of $\frac{dz}{z-1}$	residue of $\frac{dz}{z}$	residue of $\frac{\lambda'}{\lambda}dz$	residue of $\Phi'dz$
$1^{(1)}$	1	0	-2	$-2\xi_1 + 2$
$1^{(2)}$	1	0	2	$2\xi_1 + 2$
0	0	2	0	$-2\xi_2 - 2$
$\infty$	-2	-2	0	$2\xi_2 - 2$

The residues of  $\frac{\lambda'}{\lambda}dz$  at  $z = 1^{(1)}$  and  $z = 1^{(2)}$  come from the double pole and double zero that  $\lambda$  has at these points (see Lemma 5.2(b)). The residues add up to zero, as it should be.

We assume  $\alpha > 1$  so that the genus of  $\mathcal{R}$  is one. Then there are also four zeros of  $\Phi'dz$  counting multiplicities. Recall that the real part of the Riemann surface consists of the cycles  $\mathcal{C}_1$  and  $\mathcal{C}_2$  as in (2.24).

**Proposition 6.4.** *For every  $\xi_1, \xi_2 \in (-1, 1)$  there are at least two distinct saddle points on the cycle  $\mathcal{C}_1$ .*

*Proof.* If  $\mathcal{C}$  is a path from  $P$  to  $Q$  on the Riemann surface avoiding the poles, then by (2.22),

$$\int_{\mathcal{C}} \Phi'(z) dz = [2 \log(z-1) - (1 + \xi_2) \log z + \xi_1 \log \lambda(z)]_P^Q$$

for a choice of continuous branches of the logarithms along the path. Since  $\xi_1$  and  $\xi_2$  are real, it follows that the real part is well-defined, it depends on  $P$  and  $Q$ , but is otherwise independent of the path. Thus

$$\operatorname{Re}\left(\oint_{\mathcal{C}} \Phi'(z) dz\right) = 0 \tag{6.5}$$

for a closed path  $\mathcal{C}$ .

Observe that there are no poles on the cycle  $\mathcal{C}_1$ , and  $\Phi'$  is real there. If there were no two distinct zeros on  $\mathcal{C}_1$ , then there would be no sign change, and the integral would be non-zero and real, which would contradict the condition (6.5).  $\square$

The saddle points are explicit in case  $\xi_1 = 0$ , since then by (2.23),

$$\Phi'(z)dz = \left(\frac{2}{z-1} - \frac{1+\xi_2}{z}\right)dz. \tag{6.6}$$

The equation  $\frac{2}{z-1} = \frac{1+\xi_2}{z}$  has the unique solution

$$z_c(\xi_2) = -\frac{1+\xi_2}{1-\xi_2}. \tag{6.7}$$

This gives us two saddle points, namely the two points on  $\mathcal{R}$  with (6.7) as  $z$ -coordinate. The other two saddles come from the branch points  $-\alpha^2, -\beta^2$ , which are zeros of the differential  $dz$ . The branch point  $z = 0$  is also a zero of  $dz$ , but this zero gets cancelled by the (double) pole of  $\frac{1+\xi_2}{z}$  in (6.6).

For special values of  $\xi_2$  the saddles at  $z = z_c(\xi_2)$  coincide with the saddle at  $-\alpha^2$  or  $-\beta^2$ . This happens for the values  $\pm\xi_2^*$  with

$$\xi_2^* = \frac{\alpha - \beta}{\alpha + \beta} \in (0, 1). \tag{6.8}$$

Then depending on the value of  $\xi_2$ , we are in the liquid or gas phase, or on the liquid-gas transition, as defined in Definition 2.7.

**Lemma 6.5.** *Suppose  $\xi_1 = 0$  and  $-1 < \xi_2 < 1$ .*

- (a)  $(0, \xi_2)$  cannot be in the solid phase.
- (b) If  $\xi_2 \in (-1, -\xi_2^*) \cup (\xi_2^*, 1)$  then  $(0, \xi_2) \in \mathfrak{L}$ .
- (c) If  $-\xi_2^* < \xi_2 < \xi_2^*$  then  $(0, \xi_2) \in \mathfrak{G}$ .
- (d) If  $\xi_2 = \pm\xi_2^*$  then  $(0, \xi_2)$  is on the liquid-gas transition.

*Proof.* (a) It is clear from (6.7) that  $z_c(\xi_2) < 0$  and so there are no saddles on the positive real axis.

(b) If  $-1 < \xi_2 < -\xi_2^*$  then  $z_c(\xi_2) \in (-\beta^2, 0)$ , and if  $\xi_2^* < \xi_2 < 1$  then  $z_c(\xi_2) \in (-\infty, -\alpha^2)$ . Even though  $z_c(\xi_2)$  is real, the two saddles with  $z$ -coordinate equal to  $z_c(\xi_2)$  are not on the real part of the Riemann surface, and thus in both cases we are in the liquid phase.

(c) If  $-\xi_2^* < \xi_2 < \xi_2^*$  then  $z_c(\xi_2) \in (-\alpha^2, -\beta^2)$ . Then the saddles with  $z$ -coordinate equal to  $z_c(\xi_2)$  are on the cycle  $C_1$ . The branch points  $-\alpha^2$  and  $-\beta^2$  are the other two saddles and they are also on the cycle. Thus all four saddles are on the cycle  $C_1$  and they are distinct, and we are in the gas phase.

(d) If  $\xi_2 = -\xi_2^*$  then  $z_c(\xi_2) = -\beta^2$  and if  $\xi_2 = \xi_2^*$  then  $z_c(\xi_2) = -\alpha^2$ . In both cases there is a triple saddle point at one of the branch points, and we are on the liquid-gas transition.  $\square$

### 6.3. Algebraic equation

The condition of coalescing saddle points leads to an algebraic equation for  $\xi_1$  and  $\xi_2$ . We are able to calculate it with the help of Maple.

First of all, the saddle point equation  $\Phi'(z)dz = 0$  (see (2.23)) leads us to consider

$$\frac{2}{z-1} - \frac{1+\xi_2}{z} + \xi_1 \frac{\lambda'(z)}{\lambda(z)} = 0,$$

which after clearing denominators, and using (2.17) and (2.12), gives a polynomial equation in  $z$  and  $y = \sqrt{z(z+\alpha^2)(z+\beta^2)}$ . We eliminate the square root to obtain a polynomial equation in  $z$  of degree 4, which is

$$\begin{aligned} (1-\xi_2)^2 z^4 &+ ((\alpha^2 + \beta^2)((1-\xi_2)^2 - \xi_1^2) + 2(1-\xi_2^2 - \xi_1^2))z^3 \\ &+ (2(\alpha^2 + \beta^2)(1-\xi_2^2 - \xi_1^2) + 2 - 4\xi_1^2 + 2\xi_2^2)z^2 \\ &+ ((\alpha^2 + \beta^2)((1+\xi_2)^2 - \xi_1^2) + 2(1-\xi_2^2 - \xi_1^2))z + (1+\xi_2)^2 = 0. \end{aligned} \tag{6.9}$$

By definition, the saddles are the four zeros of the polynomial (6.9).

The discriminant with respect to  $z$  of (6.9) is a polynomial in  $\xi_1$  and  $\xi_2$  that has trivial factors  $\xi_1^2$  and  $\xi_2^2$ . The remaining factor is a degree 8 polynomial, which is symmetric in the two variables. Setting this to zero, we obtain the following equation for coalescing saddles:

$$\begin{aligned} (\alpha^2 + 1)^6 (\xi_1^8 + \xi_2^8) &- 4(\alpha^2 + 1)^4 (\alpha^2 + 2\alpha - 1)(\alpha^2 - 2\alpha - 1) (\xi_1^6 \xi_2^2 + \xi_1^2 \xi_2^6) \\ &- 4(\alpha^2 + 1)^4 (\alpha^4 - \alpha^2 + 1) (\xi_1^6 + \xi_2^6) + 2(\alpha^2 + 1)^2 (3\alpha^8 - 20\alpha^6 + 82\alpha^4 - 20\alpha^2 + 3) \xi_1^4 \xi_2^4 \\ &+ 4(\alpha^2 + 1)^2 (\alpha^8 + 17\alpha^6 - 48\alpha^4 + 17\alpha^2 + 1) (\xi_1^4 \xi_2^2 + \xi_1^2 \xi_2^4) \\ &+ 6(\alpha^4 - 1)^2 (\alpha^4 + 1) (\xi_1^4 + \xi_2^4) + 4(\alpha^2 - 1)^2 (\alpha^8 - 22\alpha^6 - 42\alpha^4 - 22\alpha^2 + 1) \xi_1^2 \xi_2^2 \\ &- 4(\alpha^2 - 1)^4 (\alpha^2 + \alpha + 1)(\alpha^2 - \alpha + 1) (\xi_1^2 + \xi_2^2) + (\alpha^2 - 1)^6 = 0. \end{aligned} \tag{6.10}$$

Up to a multiplicative constant, the equation (6.10) coincides with the one given by Chhita and Johansson [20, Appendix A]. See also [61, Section 8] for an equation that corresponds to (6.10) with  $\alpha = 2$  up to a change of variables. For  $\alpha = 1$ , (6.10) reduces (up to a numerical factor) to  $(1 - \xi_1^2 - \xi_2^2)(\xi_1^2 + \xi_2^2)^3 = 0$ , and the real section is the unit circle.

For  $\alpha > 1$ , the real section of (6.10) has two components, as shown in Figure 6, both contained in the square  $-1 \leq \xi_1, \xi_2 \leq 1$ . The outer component is a smooth closed curve that touches the square in the points  $(\pm 1, 0)$  and  $(0, \pm 1)$ . The inner component is a closed curve with four cusps at  $(\pm \frac{\alpha-\beta}{\alpha+\beta}, 0)$  and  $(0, \pm \frac{\alpha-\beta}{\alpha+\beta})$ .

**Proposition 6.6.** *Let  $-1 < \xi_1, \xi_2 < 1$ .*

- (a)  $(\xi_1, \xi_2) \in \mathfrak{G}$  (gas phase) if and only if  $(\xi_1, \xi_2)$  is inside the inner component of the algebraic curve.
- (b)  $(\xi_1, \xi_2) \in \mathfrak{L}$  (liquid phase) if and only if  $(\xi_1, \xi_2)$  is outside the inner component and inside the outer component.
- (c)  $(\xi_1, \xi_2) \in \mathfrak{S}$  (solid phase) if and only if  $(\xi_1, \xi_2)$  is outside the outer component.

*Proof.* If  $\xi_1 = 0$  then all statements of the proposition follow from Lemma 6.5.

The proof in the general cases follows by a continuity argument, since the saddles depend continuously on the parameters  $\xi_1, \xi_2$ , and a saddle can only leave the real part of the Riemann surface if it coalesces with another saddle and then the pair can move away from the real part. This transition can thus only occur for  $\xi_1, \xi_2$  satisfying the algebraic equation (6.10). Note that this argument also applies to the point at infinity, since by definition (2.24) the point at infinity is included in the real part.

Combining this with Lemma 6.5 we find that any point  $(\xi_1, \xi_2)$  inside the inner component belongs to the gas phase, and any point in between the inner and outer component belongs to the liquid phase.

To treat the solid phase, we check that any  $(\xi_1, \xi_2)$  close enough to a corner point is in the solid phase. We can see this from the equation (6.10). If  $\xi_1 = \pm 1$  and  $\xi_2 = -1$  then (6.9) has solutions  $z = 0, z = 1$  (and two other solutions that are on the cycle  $\mathcal{C}_1$ ), and if  $\xi_1 = \pm 1$  and  $\xi_2 = 1$  then (6.9) has solutions  $z = \infty$  and  $z = 1$ . Thus for each of the four corner points there are two distinct saddles on  $\mathcal{C}_2$ . Then by continuity this continues to be the case if  $(\xi_1, \xi_2)$  is close enough to one of the corner points, and it continues to be so until the two saddles on  $\mathcal{C}_2$  coalesce, and this happens on the outer component.

This completes the proof of Proposition 6.6. □

#### 6.4. Gas phase: steepest descent paths

In the gas phase all four saddles are located on the cycle  $\mathcal{C}_1$ , and they are all simple. To prepare for the proof of Theorem 2.8, we need more precise information on the location of the saddles.

**Lemma 6.7.** *Suppose  $(\xi_1, \xi_2) \in \mathfrak{G}$ . Then the function*

$$\operatorname{Re} \Phi(z) = 2 \log |z - 1| - (1 + \xi_2) \log |z| + \xi_1 \log |\lambda(z)|, \quad z \in \mathcal{C}_1, \quad (6.11)$$

*attains a local minimum at two of the saddles, say  $z_{s,1}$  and  $z_{s,2}$ , where  $z_{s,j}$  is on the  $j$ th sheet for  $j = 1, 2$ .*

*It attains a local maximum on  $\mathcal{C}_1$  at the other two saddles  $z_{s,3} < z_{s,4}$ . If  $\xi_1 > 0$  then  $z_{s,3}$  and  $z_{s,4}$  are on the first sheet and  $-\alpha^2 < z_{s,3} < z_{s,1} < z_{s,4} < -\beta^2$ ; if  $\xi_1 < 0$  then  $z_{s,3}$  and  $z_{s,4}$  are on the second sheet and  $-\alpha^2 < z_{s,3} < z_{s,2} < z_{s,4} < -\beta^2$ ; and if  $\xi_1 = 0$  then  $z_{s,3} = -\alpha^2$  and  $z_{s,4} = -\beta^2$  are at the branch points.*

*Proof.* The conclusion is easy to verify if  $\xi_1 = 0$ , since  $z \mapsto 2 \log |z - 1| - (1 + \xi_2) \log |z|$  attains a local minimum at  $z = z_c(\xi_2)$  given by (6.7). Thus  $z_{s,j} = (z_c(\xi_2))^{(j)}$  for  $j = 1, 2$ , and the other saddles  $z_{s,3}$  and  $z_{s,4}$  are at the branch points.

For  $\xi_1 \neq 0$ , we notice that both  $\lambda_1$  and  $\lambda_2$  are real and negative on the interval  $[-\alpha^2, -\beta^2]$  (see (5.12)), with a square root behavior at endpoints (which follows from (2.17) and the square roots in (2.12)). Thus  $\lambda'_{1,2}$  become infinite at the endpoints, and a closer inspection of (2.12), (2.17) shows that

$$\lim_{z \rightarrow -\alpha^2+} \frac{\lambda'_1(z)}{\lambda_1(z)} = +\infty, \quad \lim_{z \rightarrow -\beta^2-} \frac{\lambda'_1(z)}{\lambda_1(z)} = -\infty, \tag{6.12}$$

$$\lim_{z \rightarrow -\alpha^2+} \frac{\lambda'_2(z)}{\lambda_2(z)} = -\infty, \quad \lim_{z \rightarrow -\beta^2-} \frac{\lambda'_2(z)}{\lambda_2(z)} = +\infty. \tag{6.13}$$

Thus if  $\xi_1 \neq 0$ , both functions

$$\Phi'_j(z) = \frac{2}{z-1} - \frac{1+\xi_2}{z} + \xi_1 \frac{\lambda'(z)}{\lambda_j(z)}, \quad j = 1, 2, \tag{6.14}$$

are infinite at the endpoints of the interval  $[-\alpha^2, -\beta^2]$  but with opposite signs. By continuity, each function has an odd number of zeros. There are exactly four simple saddles on the cycle  $\mathcal{C}_1$  as we are in the gas phase, and therefore one of  $\Phi'_j$ ,  $j = 1, 2$ , has three simple zeros and the other one has one simple zero.

We already noted that for  $\xi_1 = 0$ ,

$$\text{Re } \Phi_j(z) = 2 \log |z-1| - (1+\xi_2) \log |z| + \xi_1 \log |\lambda_j(z)| \tag{6.15}$$

attains a local minimum at an interior saddle for  $j = 1, 2$ . Because of analytic dependence on parameters this continues to be the case for  $\xi_1 \neq 0$ , and in fact, since there is no coalescence of saddle points, it remains true for every  $(\xi_1, \xi_2) \in \mathfrak{G}$ . Thus saddles  $z_{s,1}$  and  $z_{s,2}$  where  $\text{Re } \Phi$  has a local minimum exist, and  $z_{s,j}$  is on the  $j$ th sheet.

Now, if  $\xi_1 > 0$  then from (6.12) and (6.14) it follows that  $\lim_{z \rightarrow -\alpha^2+} \Phi'_1(z) = +\infty$ ,  $\lim_{z \rightarrow -\beta^2-} \Phi'_1(z) = -\infty$ , and so  $\text{Re } \Phi_1$  increases on an interval  $[-\alpha^2, -\alpha^2 + \delta]$  and decreases on  $[-\beta^2 - \delta, -\beta^2]$  for some  $\delta > 0$ . Since there is a local minimum at  $z_{s,1}$  on the first sheet, it should be that  $\text{Re } \Phi_1$  has two local maxima, say  $z_{s,3} < z_{s,4}$ , with  $-\alpha^2 < z_{s,3} < z_{s,1} < z_{s,4} < -\beta^2$ .

In case  $\xi_1 < 0$  we find in the same way that the local maxima are on the second sheet, with  $-\alpha^2 < z_{s,3} < z_{s,2} < z_{s,4} < -\beta^2$ . □

The *path of steepest descent* from the saddle  $z_{s,j}$ ,  $j = 1, 2$ , is the curve  $\gamma_{sd,j}$  through  $z_{s,j}$  where the imaginary part of  $\Phi_j$  is constant and the real part decays if we move away from the saddle. Since  $\text{Re } \Phi_j$  on  $\mathcal{C}_1$  has a local minimum at  $z_{s,j}$ , the path of steepest descent meets the real line at a right angle.

Emanating from  $z_{s,j}$ ,  $j = 1, 2$ , are also curves  $\gamma_{l,j}$  and  $\gamma_{r,j}$  where the real part is constant ( $l$  stands for left, and  $r$  stands for right). The curve  $\gamma_{l,j}$  emanates from  $z_{s,j}$  at angles  $\pm 3\pi/4$ . It consists of a part in the upper half-plane and its mirror image with respect to the real line in the lower half-plane. Similarly,  $\gamma_{r,j}$  emanates from  $z_{s,j}$  at angles  $\pm \pi/4$ , and it is also symmetric in the real line. Near  $z_{s,j}$  we find that  $\gamma_{l,j}$  is to the left of the steepest descent path  $\gamma_{sd,j}$ , and  $\gamma_{r,j}$  is to the right.

Then  $\gamma_{l,j}$  and  $\gamma_{r,j}$  are parts of the boundary of the domains

$$\Omega_j^- = \{z \in \mathbb{C} \mid \operatorname{Re} \Phi_j(z) < \operatorname{Re} \Phi_j(z_{s,j})\}, \tag{6.16}$$

$$\Omega_j^+ = \{z \in \mathbb{C} \mid \operatorname{Re} \Phi_j(z) > \operatorname{Re} \Phi_j(z_{s,j})\}. \tag{6.17}$$

**Lemma 6.8.** *Suppose  $(\xi_1, \xi_2) \in \mathfrak{G}$ , and  $j = 1, 2$ . Then the following hold.*

- (a) *All three curves  $\gamma_{sd,j}$ ,  $\gamma_{l,j}$  and  $\gamma_{r,j}$  are simple closed curves enclosing the interval  $[-\beta^2, 0]$ .*
- (b) *The steepest descent path  $\gamma_{sd,j}$  intersects the positive real line at 1,  $\gamma_{l,j}$  intersects the positive real line at a value  $> 1$ , while  $\gamma_{r,j}$  intersects the positive real line at a value  $< 1$ .*
- (c)  *$\Omega_j^-$  is a bounded open set with at most three connected components. One component (which we call the main component) contains  $\gamma_{sd,j} \setminus \{z_{s,j}\}$ . There are at most two other components: a component containing  $-\beta^2$  (if  $\operatorname{Re} \Phi_j(-\beta^2) < \operatorname{Re} \Phi_j(z_{s,j})$ ) and a component containing  $-\alpha^2$  (if  $\operatorname{Re} \Phi_j(-\alpha^2) < \operatorname{Re} \Phi_j(z_{s,j})$ ). The other components (if they exist) are at positive distance from the main component.*
- (d)  *$\Omega_j^+$  is an open set with an unbounded component that contains a contour  $\Gamma_{1,j}$  that goes around  $(-\infty, -\alpha^2]$  and with a bounded component that contains a contour  $\Gamma_{2,j}$  going around  $[-\beta^2, 0]$ .*

*Proof.* The conclusion is straightforward to verify if  $\xi_1 = 0$ , since in that case

$$\operatorname{Re} \Phi_1(z) = \operatorname{Re} \Phi_2(z) = 2 \log |z - 1| - (1 + \xi_2) \log |z| \tag{6.18}$$

and  $z_{s,1} = z_{s,2} = z_c(\xi_2) = -\frac{1+\xi_2}{1-\xi_2}$ , which is in  $(-\alpha^2, -\beta^2)$  since we are in the gas phase. Then  $\gamma_{l,j}$ ,  $\gamma_{sd,j}$ , and  $\gamma_{r,j}$  are independent of  $j$  and we have a situation as in Figure 13. The curves  $\gamma_{l,j}$  and  $\gamma_{r,j}$  enclose the domain  $\Omega_j^-$  that is shaded in the figure, and  $\Omega_j^-$  has only one component in this case.

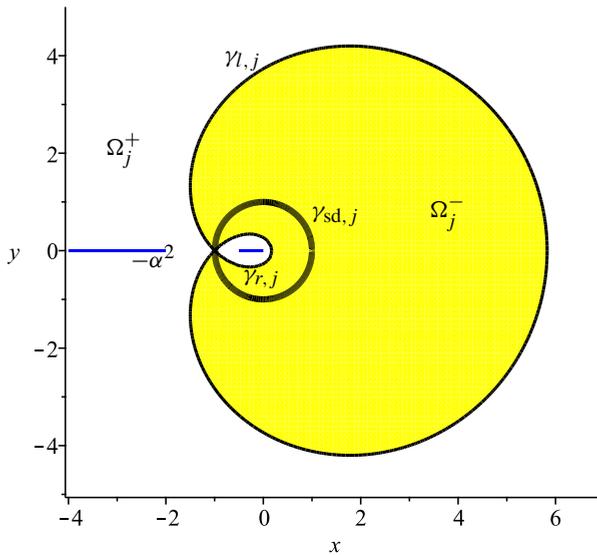
The non-shaded domain  $\Omega_j^+$  has an unbounded component with boundary  $\gamma_{l,j} \cup (-\infty, -\alpha^2]$  and a bounded component with boundary  $\gamma_{r,j} \cup [-\beta^2, 0]$ . The contours  $\Gamma_{1,j}$  and  $\Gamma_{2,j}$  can be taken in  $\Omega_j^+$  as specified in part (d) of the lemma.

From (6.18) it is easy to see that  $\operatorname{Re} \Phi_j$  is strictly decreasing on  $(-\infty, z_c(\xi_2)]$ , strictly increasing on  $(z_c(\xi_2), 0]$  with value  $+\infty$  at 0, then strictly decreasing on  $[0, 1]$  with value  $-\infty$  at 1, and finally strictly increasing on  $[1, \infty)$  with limiting value  $+\infty$  at  $\infty$ .

For  $\xi_1 \neq 0$  the behavior on the positive real axis is the same: for both  $j = 1$  and  $j = 2$ ,  $\Phi_j(x)$  is real for  $x > 0$  and it decreases from  $+\infty$  to  $-\infty$  on  $[0, 1]$  and increases from  $-\infty$  to  $+\infty$  on  $[1, \infty)$ . This is so because otherwise there would be a zero of the derivative, which would be a saddle point, but all saddle points are on the cycle  $\mathcal{C}_1$  since we are in the gas phase.

The behavior of  $\operatorname{Re} \Phi_j$  on the cuts  $(-\infty, -\alpha^2] \cup [-\beta^2, 0]$  is exactly the same as for  $\xi_1 = 0$ . Indeed, it does not depend on  $\xi_1$  at all, since  $|\lambda_{j,\pm}(x)| = 1$  for  $x$  on the cuts (see (5.10) and (6.15)). Thus

$$\begin{aligned} \operatorname{Re} \Phi_j(x) &\text{ is strictly decreasing for } x \in (-\infty, -\alpha^2], \\ \operatorname{Re} \Phi_j(x) &\text{ is strictly increasing for } x \in [-\beta^2, 0]. \end{aligned} \tag{6.19}$$



**Fig. 13.** Illustration for Lemma 6.8. The figure shows the steepest descent curve  $\gamma_{sd,j}$  (dark curve) from the saddle  $z_{s,j}$  and the level lines  $\gamma_{l,j}$  and  $\gamma_{r,j}$  of  $\text{Re } \Phi_j$  that enclose the domain  $\Omega_j^-$  for the case  $\xi_1 = 0$ . In this case  $\Omega_j^-$  has one component, but for  $\xi_1 \neq 0$ , there could be a component containing  $-\alpha^2$ , and a component containing  $-\beta^2$ .

Now let us follow the paths  $\gamma_{l,j}$  and  $\gamma_{r,j}$ , where the real part  $\text{Re } \Phi_j$  is constant, as they move away from  $z_{s,j}$  into the upper half-plane. These paths remain bounded, since  $\text{Re } \Phi_j(z) \rightarrow +\infty$  as  $|z| \rightarrow \infty$ , which follows from (6.15), Lemma 5.2(c), and the fact that  $\xi_2 < 1$ . The two paths cannot come together in the upper half-plane, since they would then enclose a domain on which  $\text{Re } \Phi_j$  is harmonic and constant on the boundary, which violates the maximum/minimum principle of harmonic functions. For the same reason they cannot meet at a point on the positive real axis.

Suppose now one of the paths comes to the cut  $(-\infty, -\alpha^2]$  at a point  $q$ . Then this path, together with its mirror image in the real line, encloses a bounded domain  $D$  and  $\text{Re } \Phi_j$  has the constant value  $\text{Re } \Phi_j(z_{s,j})$  on its boundary. The value of  $\text{Re } \Phi_j$  is smaller on  $[q, -\alpha^2]$  because of (6.19). Also  $\text{Re } \Phi_j$  is harmonic on  $D \setminus [q, -\alpha^2]$  and it follows from the maximum principle for harmonic functions that  $\text{Re } \Phi_j(z) < \text{Re } \Phi_j(z_{s,j})$  for  $z \in D$ . This is a contradiction, since  $z_{s,j}$  is a saddle at which  $\text{Re } \Phi_j$  attains a local minimum when restricted to the cycle  $\mathcal{C}_1$ .

We arrive at a similar contradiction if one of the paths  $\gamma_{l,j}$  and  $\gamma_{r,j}$  comes to the cut  $[-\beta^2, 0]$ .

Thus the two paths can only leave the upper half-plane via the positive real axis. Since  $\Phi_j$  is strictly decreasing on  $[0, 1]$  and strictly increasing on  $[1, \infty]$ , we conclude that  $\gamma_{l,j}$  comes to the unique value in  $(1, \infty)$  and  $\gamma_{r,j}$  to the unique value in  $(0, 1)$  where  $\Phi_j$  is equal to  $\text{Re } \Phi_j(z_{s,j})$ . Together with their mirror images in the real axis, they both form

simple closed curves that go around the cut  $[-\beta^2, 0]$ . They enclose a domain where  $\operatorname{Re} \Phi_j$  is smaller than  $\operatorname{Re} \Phi_j(z_{s,j})$ , i.e., it is contained in  $\Omega_j^-$  (see (6.16)).

The path of steepest descent  $\gamma_{\text{sd},j}$  lies in  $\Omega_j^-$ , and  $\operatorname{Re} \Phi_j$  decreases along  $\gamma_{\text{sd},j}$  in the upper half-plane. It will meet the positive real axis at 1, where  $\operatorname{Re} \Phi_j$  is  $-\infty$ , since if it met the positive real axis at some other point, this point would be a saddle—but there is no saddle on  $\mathcal{C}_2$ .

We have thus proved parts (a) and (b) of Lemma 6.8. We have also proved that the component of  $\Omega_j^-$  that is bounded by  $\gamma_{l,j}$  and  $\gamma_{r,j}$  contains the steepest descent curve  $\gamma_{\text{sd},j} \setminus \{z_{s,j}\}$ . We also see that  $\Omega_j^-$  is bounded since  $\operatorname{Re} \Phi_j(z) \rightarrow +\infty$  as  $|z| \rightarrow \infty$ .

Any other connected component of  $\Omega_j^-$  has to intersect the branch cuts  $(-\infty, -\alpha^2] \cup [-\beta^2, 0]$ , since otherwise we have again a contradiction with the minimum principle for harmonic functions.

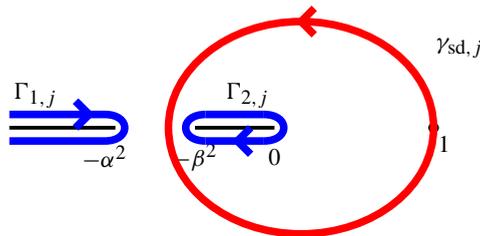
If a component of  $\Omega_j^-$  intersects  $(-\infty, -\alpha^2]$  then it will do so along an interval  $(q, -\alpha^2]$  for some  $q < -\alpha^2$ , because of (6.19). Hence there can be at most one such component, and it exists if and only if  $\operatorname{Re} \Phi_j(-\alpha^2) < \operatorname{Re} \Phi_j(z_{s,j})$ . Similarly, there is at most one component that intersects  $[-\beta^2, 0]$ , and thus in total there are at most three components.

Finally, since  $z_{s,j}$  is a simple saddle, and  $\operatorname{Re} \Phi_j$  attains a minimum at  $z_{s,j}$  when we restrict to the cycle  $\mathcal{C}_1$ , there is  $\delta > 0$  such that

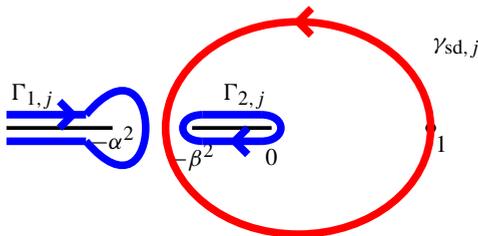
$$(z_{s,j} - \delta, z_{s,j}) \cup (z_{s,j}, z_{s,j} + \delta) \subset \Omega_j^+.$$

Thus the boundary of another component (if it exists) intersects the interval  $(-\alpha^2, -\beta^2)$  at a point different from  $z_{s,j}$ , and thus the component stays at positive distance from the main component. This proves part (c) of the lemma.

The component of  $\Omega_j^-$  that contains  $-\alpha^2$  (if it exists) is bounded and belongs to the exterior of the closed contour  $\gamma_{l,j}$ . However, by part (c), it remains at a positive distance from  $\gamma_{l,j}$ , and therefore we can find a contour  $\Gamma_{1,j}$  as in the statement of part (d) that goes around the interval  $(-\infty, -\alpha^2]$  and avoids the closure of the domain  $\Omega_j^-$  while staying to the left of  $\gamma_{l,j}$ . Similarly, we can find  $\Gamma_{2,j}$  as in part (d). See Figures 14 and 15 for plots of  $\gamma_{\text{sd},j}$  and  $\Gamma_{1,j}$  and  $\Gamma_{2,j}$  in the situations where  $\Omega_j^-$  has no component containing  $-\alpha^2$  or  $-\beta^2$  (Figure 14) and where  $\Omega_j^-$  has a component containing  $-\alpha^2$  that  $\Gamma_{1,j}$  should avoid (Figure 15).



**Fig. 14.** Steepest descent contour  $\gamma_{\text{sd},j}$  and contours  $\Gamma_{1,j}$  and  $\Gamma_{2,j}$  in case  $\Omega_j^-$  has only one component.



**Fig. 15.** Steepest descent contour  $\gamma_{sd,j}$  and contours  $\Gamma_{1,j}$  and  $\Gamma_{2,j}$  in case  $\Omega_j^-$  has a component containing  $-\alpha^2$ , but no component containing  $-\beta^2$ .

This completes the proof of Lemma 6.8. □

6.5. Gas phase: proof of Theorem 2.8

To establish Theorem 2.8 we are going to show that in the gas phase the two double integrals in (6.2) tend to 0 as  $N \rightarrow \infty$  at an exponential rate.

In the situation of Theorem 2.8 we have  $m = (1 + \xi_1)N + o(N)$ ,  $n = (1 + \xi_2)N + o(N)$ ,  $m' = (1 + \xi_1)N + o(N)$ ,  $n' = (1 + \xi_2)N + o(N)$ , so that in (6.2) we should replace  $\xi_1$  and  $\xi_1'$  by  $\xi_1 + o(1)$ , and  $\xi_2$  and  $\xi_2'$  by  $\xi_2 + o(1)$ . The  $o(1)$  terms give only a subleading contribution, and it is enough to prove that both integrals

$$\frac{1}{(2\pi i)^2} \oint_{\gamma_{0,1}} \frac{dz}{z} \int_{\Gamma} \frac{dw}{z-w} F(w)F(z)e^{N(\Phi_1(z)-\Phi_1(w))/2} \tag{6.20}$$

and

$$\frac{1}{(2\pi i)^2} \oint_{\gamma_{0,1}} \frac{dz}{z} \int_{\Gamma} \frac{dw}{z-w} F(w)(I_2 - F(z))e^{N(\Phi_2(z)-\Phi_1(w))/2} \tag{6.21}$$

are exponentially small as  $N \rightarrow \infty$ . Here we write again  $\Phi_1$  and  $\Phi_2$  instead of  $\Phi_1(\cdot; \xi_1, \xi_2)$  and  $\Phi_2(\cdot; \xi_1, \xi_2)$ .

We move the contour  $\gamma_{0,1}$  in (6.20) to the steepest descent path  $\gamma_{sd,1}$  through  $z_{s,1}$ , and we have

$$\operatorname{Re} \Phi_1(z) \leq \operatorname{Re} \Phi_1(z_{s,1}), \quad z \in \gamma_{sd,1}. \tag{6.22}$$

The contour  $\Gamma$  in (6.20) traverses the interval  $(-\alpha^2, -\beta^2)$  twice, but in opposite directions. The integrand has no branching on this interval, and therefore the two contributions cancel out. We can thus replace  $\Gamma$  by two contours  $\Gamma_{1,1}$  and  $\Gamma_{2,1}$  as in Lemma 6.8(d), where  $\Gamma_{1,1}$  goes around  $(-\infty, -\alpha^2]$ ,  $\Gamma_{2,1}$  goes around  $[-\beta^2, 0]$ , and

$$\operatorname{Re} \Phi_1(w) > \operatorname{Re} \Phi_1(z_{s,1}), \quad w \in \Gamma_{1,1} \cup \Gamma_{2,1} \tag{6.23}$$

since the contours are in  $\Omega_1^+$  (see (6.17)), and  $\operatorname{Re} \Phi_1(w) = (1 - \xi_2) \log |w| + O(1)$  as  $w \rightarrow \infty$ .

By (6.22) and (6.23) the factor  $e^{N(\Phi_1(z)-\Phi_1(w))/2}$  is  $O(e^{-cN})$  for some  $c > 0$ , uniformly for  $(z, w) \in \gamma_{sd,1} \times (\Gamma_{1,1} \cup \Gamma_{2,1})$ . Thus (6.20) tends to 0 at an exponential rate as  $N \rightarrow \infty$ .

We are going to apply a similar argument to (6.21) and deform  $\gamma_{0,1}$  to the steepest descent contour  $\gamma_{sd,2}$  passing through  $z_{s,2}$ . We again replace  $\Gamma$  by the union of two contours, one going around  $(-\infty, -\alpha^2]$  and the other around  $[-\beta^2, 0]$ , but now we are going to move these contours to the cuts, where we note that  $\lambda_{1,\pm} = \lambda_{2,\mp}$  and  $\Phi_{1,\pm} = \Phi_{2,\mp}$  plus a purely imaginary constant (depending on the precise branches of the logarithms that we choose in (2.22)). We also note that there is no pole at  $w = 0$ , and that  $I_2 - F(w)$  is the analytic continuation of  $F(w)$  to the second sheet. Then we deform the contours further to  $\Gamma_{1,2}$  and  $\Gamma_{2,2}$  as in Lemma 6.8(d), and we see that (6.21) is equal to

$$-\frac{1}{(2\pi i)^2} \oint_{\gamma_{sd,2}} \frac{dz}{z} \int_{\Gamma_{1,2} \cup \Gamma_{2,2}} \frac{dw}{z-w} (I_2 - F(w))(I_2 - F(z)) e^{N(\Phi_2(z) - \Phi_2(w))/2}. \tag{6.24}$$

The minus sign comes since we use the orientation on  $\Gamma_{1,2}$  and  $\Gamma_{2,2}$  as in Figure 15.

Analogously to (6.22) and (6.23) we now have

$$\operatorname{Re} \Phi_2(z) \leq \operatorname{Re} \Phi_2(z_{s,2}) < \operatorname{Re} \Phi_2(w), \quad z \in \gamma_{sd,2}, \quad w \in \Gamma_{1,2} \cup \Gamma_{2,2},$$

and the integral (6.24) tends to 0 at an exponential rate as  $N \rightarrow \infty$ .

This completes the proof of Theorem 2.8. □

### 6.6. Cusp points: proof of Theorem 2.10

For the proof of Theorem 2.10 we are going to use (6.2) once more.

In the scaling (2.30) of the parameters, the saddle points coalesce to a triple saddle point at  $-\alpha^2$ . We are going to deform  $\gamma_{0,1}$  so that it comes close to  $-\alpha^2$ . The main contribution to the integrals in (6.2) then comes from the triple saddle at  $-\alpha^2$ , as the integrands are exponentially small if  $w$  and/or  $z$  are outside a small neighborhood of  $-\alpha^2$ . This follows as in the proof of the gas phase.

Hence, our task is to investigate how the entries in the integrals in (6.2) behave for  $z$  and  $w$  close to  $-\alpha^2$ . We do this in the following two lemmas and their corollaries. Besides the constants  $c_1$  and  $c_2$  from (2.31) we also use

$$c_0 = \frac{\alpha + \beta}{\sqrt{2}}. \tag{6.25}$$

**Lemma 6.9.** *We have*

$$\begin{aligned} \Phi_{1,2}(z; \xi_1, \xi_2) &= 2 \log(\alpha^2 + 1) - (1 + \xi_2) \log(\alpha^2) + (1 - \xi_2 \pm \xi_1) \pi i \operatorname{sgn}(\operatorname{Im} z) \\ &\quad \pm \frac{2\sqrt{\alpha - \beta}}{\sqrt{\alpha + \beta}} \xi_1 \left(\frac{z + \alpha^2}{\alpha^2}\right)^{1/2} + (\xi_2 - \xi_2^*) \left(\frac{z + \alpha^2}{\alpha^2}\right) + \frac{1}{(\alpha + \beta)^2} \left(\frac{z + \alpha^2}{\alpha^2}\right)^2 \\ &\quad + \xi_1 O(z + \alpha^2) + (\xi_2 - \xi_2^*) O((z + \alpha^2)^2) + O((z + \alpha^2)^3). \end{aligned} \tag{6.26}$$

The sign  $\pm$  means that  $+$  applies to  $\Phi_1$  and  $-$  applies to  $\Phi_2$ .

*Proof.* We recall

$$\begin{aligned} \Phi_1(z; \xi_1, \xi_2) &= 2 \log(z - 1) - (1 + \xi_2) \log z + \xi_1 \log \lambda_1(z) \\ &= \Phi_1(z; 0, \xi_2^*) - (\xi_2 - \xi_2^*) \log z + \xi_1 \log \lambda_1(z). \end{aligned} \tag{6.27}$$

All terms are multi-valued because of the logarithms. We take principal branches of  $\log(z - 1)$  and  $\log z$ , and the analytic branch of  $\log \lambda_1(z)$  in  $\mathbb{C} \setminus ((-\infty, 0] \cup [1, \infty))$  that is real and positive on  $(0, 1)$  (recall  $\lambda_1(x) > 1$  for  $x \in (0, 1)$ ).

Since  $\Phi_1(z; 0, \xi_2^*)$  has a critical point at  $z = -\alpha^2$  with second derivative equal to  $\frac{d^2}{dz^2} \Phi_1(z; 0, \xi_2^*) = \frac{2}{\alpha^4(\alpha+\beta)^2}$ , we have, as  $z \rightarrow -\alpha^2$ ,

$$\begin{aligned} \Phi_1(z; 0, \xi_2^*) &= 2 \log(\alpha^2 + 1) - (1 + \xi_2^*) \log(\alpha^2) + (1 - \xi_2^*) \pi i \operatorname{sgn}(\operatorname{Im} z) \\ &\quad + \frac{1}{(\alpha + \beta)^2} \left( \frac{z + \alpha^2}{\alpha^2} \right)^2 + O((z + \alpha^2)^3). \end{aligned} \tag{6.28}$$

By an expansion of  $\log z$  around  $z = -\alpha^2$ ,

$$\log z = \log(\alpha^2) + \pi i \operatorname{sgn}(\operatorname{Im} z) - \frac{z + \alpha^2}{\alpha^2} + O((z + \alpha^2)^2). \tag{6.29}$$

From the formula (2.12) for  $\rho_1$ ,

$$\rho_1(z) = -\alpha^2(\alpha + \beta) - \alpha\sqrt{\alpha^2 - \beta^2} (z + \alpha^2)^{1/2} + O(z + \alpha^2),$$

and then from the formula (2.17) for  $\lambda_1$  we get

$$\lambda_1(z) = -1 - 2 \frac{\sqrt{\alpha - \beta}}{\sqrt{\alpha + \beta}} \left( \frac{z + \alpha^2}{\alpha^2} \right)^{1/2} + O(z + \alpha^2),$$

which confirms the fact that  $\lambda_1(x) < -1$  for  $x \in (-\alpha^2, -\beta^2)$  as stated in (5.12) of Lemma 5.2(e). Our choice of  $\log \lambda_1(z)$  is such that  $\log \lambda_1(z) \rightarrow \pm \pi i$  as  $z \rightarrow -\alpha^2$  with  $\pm \operatorname{Im} z > 0$ . Hence

$$\log \lambda_1(z) = \pi i \operatorname{sgn}(\operatorname{Im} z) + 2 \frac{\sqrt{\alpha - \beta}}{\sqrt{\alpha + \beta}} \left( \frac{z + \alpha^2}{\alpha^2} \right)^{1/2} + O(z + \alpha^2) \tag{6.30}$$

as  $z \rightarrow \infty$ . Using (6.28), (6.29), and (6.30) in (6.27) we find the expansion (6.26) for  $\Phi_1$ .

The expansion for  $\Phi_2$  also follows, since  $\lambda_2 = 1/\lambda_1$  by Lemma 5.2(d), which means  $\log \lambda_2 = -\log \lambda_1$ , and so  $\Phi_2$  is obtained from  $\Phi_1$  by simply changing the sign of  $\xi_1$ .  $\square$

**Corollary 6.10.** *Suppose*

$$\frac{z + \alpha^2}{\alpha^2} = c_0 s N^{-1/2}, \quad \frac{w + \alpha^2}{\alpha^2} = c_0 t N^{-1/2}, \tag{6.31}$$

with constant  $c_0$  as in (6.25). Then under the scaling assumptions (2.30) and (2.31),

$$\begin{aligned} &\exp(N(\Phi_{1,2}(z; \xi_1, \xi_2) - \Phi_1(w; \xi'_1, \xi'_2))/2) \\ &= (-1)^{(\Delta n - \Delta m)/2} (\pm 1)^m \alpha^{\Delta n} \frac{e^{\pm us^{1/2} + \frac{1}{2}vs + \frac{1}{4}s^2}}{e^{u't^{1/2} + \frac{1}{2}v't + \frac{1}{4}t^2}} (1 + O(N^{-1/4})) \end{aligned} \tag{6.32}$$

as  $N \rightarrow \infty$ .

*Proof.* From Lemma 6.9 and the scalings (2.30) and (6.31) we find after straightforward calculations

$$\begin{aligned} \Phi_{1,2}(z; \xi_1, \xi_2) &= 2 \log(\alpha^2 + 1) - (1 + \xi_2) \log(\alpha^2) + (1 - \xi_2 \pm \xi_1) \pi i \operatorname{sgn}(\operatorname{Im} z) \\ &\quad + (\pm 2us^{1/2} + vs + \frac{1}{2}s^2)N^{-1} + O(N^{-5/4}), \end{aligned} \tag{6.33}$$

and similarly for  $\Phi_1(w; \xi'_1, \xi'_2)$ . Thus for  $\operatorname{Im} z, \operatorname{Im} w > 0$ ,

$$\begin{aligned} \Phi_{1,2}(z; \xi_1, \xi_2) - \Phi_1(w; \xi'_1, \xi'_2) &= (\xi'_2 - \xi_2) \log(\alpha^2) + (-\xi'_1 \pm \xi_1 + \xi'_2 - \xi_2) \pi i \\ &\quad + (2(\pm us^{1/2} - u't^{1/2}) + vs - v't + \frac{1}{2}(s^2 - t^2))N^{-1} + O(N^{-5/4}). \end{aligned} \tag{6.34}$$

We note

$$e^{\frac{N}{2}(\xi'_2 - \xi_2) \log(\alpha^2)} = e^{\frac{1}{2}(n' - n) \log(\alpha^2)} = \alpha^{\Delta n} \tag{6.35}$$

since  $\xi_2 = n/N - 1$  and  $\xi'_2 = n'/N - 1$ , and

$$e^{\frac{N}{2}(-\xi'_1 \pm \xi_1 + \xi'_2 - \xi_2) \pi i} = e^{\frac{1}{2}(-m' \pm m + n' - n) \pi i} \tag{6.36}$$

since also  $\xi_1 = m/N - 1$  and  $\xi'_1 = m'/N - 1$ . Recall  $m + n$  and  $m' + n'$  are even, and  $-m' \pm m + n' - n$  is even as well. Thus

$$e^{\frac{N}{2}(-\xi'_1 \pm \xi_1 + \xi'_2 - \xi_2)} = (-1)^{(\Delta n - \Delta m)/2} (\pm 1)^m. \tag{6.37}$$

The expansion (6.32) is immediate from (6.34), (6.35), and (6.37).

We have derived (6.32) for  $\operatorname{Im} z, \operatorname{Im} w > 0$ . Similar calculations give (6.32) for other combinations of signs of  $\operatorname{Im} z$  and  $\operatorname{Im} w$ . □

We also need the behavior of  $F(z)$  near the saddle  $z = -\alpha^2$ .

**Lemma 6.11.** (a) As  $z \rightarrow -\alpha^2$ ,

$$F(z) = \frac{\sqrt{\alpha - \beta}}{2\sqrt{\alpha + \beta}} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \left( \frac{z + \alpha^2}{\alpha^2} \right)^{-1/2} + \frac{1}{2} I_2 + O((z + \alpha^2)^{1/2}). \tag{6.38}$$

(b) As  $w, z \rightarrow -\alpha^2$ ,

$$\begin{aligned} &F(w)F(z) \\ &= \frac{\sqrt{\alpha - \beta}}{4\sqrt{\alpha + \beta}} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \left( \left( \frac{w + \alpha^2}{\alpha^2} \right)^{-1/2} + \left( \frac{z + \alpha^2}{\alpha^2} \right)^{-1/2} \right) + O(1). \end{aligned} \tag{6.39}$$

(c) As  $w, z \rightarrow -\alpha^2$ ,

$$\begin{aligned} &F(w)(I_2 - F(z)) \\ &= \frac{\sqrt{\alpha - \beta}}{4\sqrt{\alpha + \beta}} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \left( \left( \frac{w + \alpha^2}{\alpha^2} \right)^{-1/2} - \left( \frac{z + \alpha^2}{\alpha^2} \right)^{-1/2} \right) + O(1). \end{aligned} \tag{6.40}$$

*Proof.* Part (a) follows from (2.10) where we have to take into account that  $\sqrt{z(z + \beta^2)}$  is the negative square root for  $z = -\alpha^2$ .

Parts (b) and (c) follow from (a) since  $\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$  is a nilpotent matrix. □

**Corollary 6.12.** *Under the same scaling (6.31) as in Corollary 6.10 we have*

$$F(w)F(z) = \frac{2^{1/4}\sqrt{\alpha - \beta}}{4(\alpha + \beta)} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} (s^{-1/2} + t^{-1/2})N^{1/4} + O(1) \tag{6.41}$$

and

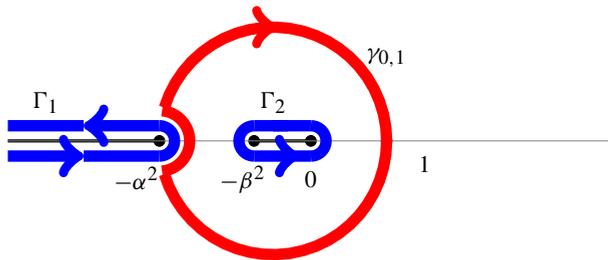
$$F(w)(I_2 - F(z)) = \frac{2^{1/4}\sqrt{\alpha - \beta}}{4(\alpha + \beta)} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} (-s^{-1/2} + t^{-1/2})N^{1/4} + O(1) \tag{6.42}$$

as  $N \rightarrow \infty$ .

*Proof.* This follows from inserting (6.31) into (6.39) and (6.40) and taking note of the value (6.25) for  $c_0$ . □

*Proof of Theorem 2.10.* We deform the contour  $\gamma_{0,1}$  in (6.2) to the one shown in Figure 16. It consists of the steepest descent contour  $\gamma_{sd,1}$  for  $\Phi_1(\cdot; 0, \xi_2^*)$  through  $-\alpha^2$ , but with a slight indentation around  $-\alpha^2$  of size  $O(N^{-1/2})$ . The steepest descent curve  $\gamma_{sd,1}$  is actually a perfect circle passing through  $-\alpha^2$  and 1, since  $\xi_1 = 0$  in the present situation. Also  $\Phi_1 = \Phi_2$ , since  $\xi_1 = 0$ .

The contour  $\Gamma$  in (6.2) is split into two components, as in the proof of Theorem 2.8. We denote them by  $\Gamma_1$  and  $\Gamma_2$  as shown in Figure 16. We reverse the orientation on both  $\Gamma_1 \cup \Gamma_2$  and  $\gamma_{0,1}$ , which does not change the values of the double integrals in (6.2).



**Fig. 16.** The contours  $\Gamma_1$  and  $\Gamma_2$  and  $\gamma_{0,1}$  with the reversed orientation.

As in the proof of Theorem 2.8, we then have

$$\operatorname{Re} \Phi_{1,2}(z) < \operatorname{Re} \Phi_1(-\alpha^2) < \operatorname{Re} \Phi_1(w)$$

whenever  $z \in \gamma_{0,1}$  and  $w \in (\Gamma_1 \setminus \{-\alpha^2\}) \cup \Gamma_2$ . From these inequalities we find that it is only a neighborhood of  $-\alpha^2$  that contributes, and in particular the component  $\Gamma_2$  that goes around  $[-\beta^2, 0]$  will not contribute to the limit. Near  $-\alpha^2$  we introduce the change of variables

$$z = -\alpha^2 + \alpha^2 c_0 s N^{-1/2}, \quad w = -\alpha^2 + \alpha^2 c_0 t N^{-1/2} \tag{6.43}$$

with  $c_0$  as in (6.25). Then

$$\frac{dzdw}{z(z-w)} = \frac{\alpha^2 c_0}{z} \frac{dsdt}{s-t} N^{-1/2} \approx \frac{\alpha + \beta}{\sqrt{2}} \frac{dsdt}{t-s} N^{-1/2} \tag{6.44}$$

where we have used  $z \sim -\alpha^2$  and the value (6.25) for  $c_0$ .

From (6.32), (6.41)–(6.42) and (6.44), we find that the leading order behavior for the sum of the two double integrals in (6.2) as  $N \rightarrow \infty$  is determined by

$$\begin{aligned} & (-1)^{(\Delta n - \Delta m)/2} \alpha^{\Delta n} \frac{\sqrt{\alpha - \beta}}{4 \cdot 2^{1/4}} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} N^{-1/4} \\ & \times \left( \frac{1}{(2\pi i)^2} \int_{s \in \Sigma_1} \int_{t \in \Sigma_2} \frac{e^{us^{1/2} + \frac{1}{2}vs + \frac{1}{4}s^2}}{e^{u't^{1/2} + \frac{1}{2}v't + \frac{1}{4}t^2}} (s^{-1/2} + t^{-1/2}) \frac{ds dt}{t-s} \right. \\ & \left. + (-1)^m \frac{1}{(2\pi i)^2} \int_{s \in \Sigma_1} \int_{t \in \Sigma_2} \frac{e^{-us^{1/2} + \frac{1}{2}vs + \frac{1}{4}s^2}}{e^{u't^{1/2} + \frac{1}{2}v't + \frac{1}{4}t^2}} (-s^{-1/2} + t^{-1/2}) \frac{ds dt}{t-s} \right) \tag{6.45} \end{aligned}$$

where  $\Sigma_1$  and  $\Sigma_2$  are as in Figure 17.  $\Sigma_2$  is the image of  $\Gamma_1$  under the mapping (6.43). It is a contour going around the negative real axis.  $\Sigma_1$  arises from applying (6.43) to the part of  $\gamma_{0,1}$  in a small disk around  $-\alpha^2$ , and deforming and extending it so that  $\Sigma_1$  is the imaginary axis with an indentation around 0.

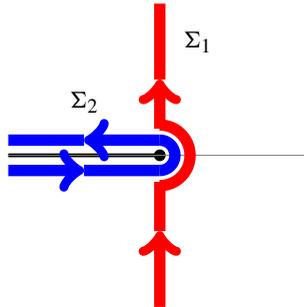


Fig. 17. The contours  $\Sigma_1$  and  $\Sigma_2$ .

We make a further change of variables by mapping  $t = (t')^2, t' \in i\mathbb{R}$ , in both integrals, and  $s = (s')^2, s' \in \Sigma$ , in the first integral, but  $s = (s')^2, s' \in (-\Sigma)$ , in the second double integral. Recall that  $\Sigma$  and  $-\Sigma$  are as in Figure 7. We find a sum of two double integrals, but the integrands are exactly the same. We pick up a Jacobian  $4s't'$  from the change of variables, and we use

$$((s')^{-1} + (t')^{-1}) \frac{s't'}{(t')^2 - (s')^2} = \frac{1}{t' - s'}$$

to simplify the expression. Then the integrand on the right-hand side of (2.32) and (2.33) follows (provided we drop the accents from  $s'$  and  $t'$ ).

This completes the proof of Theorem 2.10. □

6.7. Cusp points: proof of Theorem 2.11

The proof of Theorem 2.11 is based on the integral representation (2.29) for  $\mathbb{K}_{\text{gas}}$ .

We already noted in Remark 2.9 that  $\mathbb{K}_{\text{gas}}(m, n; m', n')$  tends to zero at an exponential rate whenever  $|\Delta m| + |\Delta n| \rightarrow \infty$ . In the situation of Theorem 2.11 we have

$$\begin{aligned} \Delta m &= (\xi'_1 - \xi_1)N = c_1(u' - u)N^{1/4} + o(N^{1/4}), \\ \Delta n &= (\xi'_2 - \xi_2)N = c_2(v' - v)N^{1/2} + o(N^{1/2}), \end{aligned}$$

as  $N \rightarrow \infty$ . Thus (2.35) holds if  $v < v'$  or if  $v = v'$  and  $u \neq u'$ .

So we assume  $v > v'$ . Then  $\mathbb{K}_{\text{gas}}(m, n; m', n')$  still decays to 0 as  $N \rightarrow \infty$ , but  $\alpha^{-\Delta n}$  increases since  $\Delta n < 0$ , and the limit in (2.35) will exist, as we shown next.

We argue similar to the proof of Theorem 2.10. We start by deforming the contour  $\gamma$  to a contour going around the interval  $(-\infty, -\alpha^2]$  as before. That is, we deform  $\gamma$  to  $\Gamma_1$  as in Figure 16 but with opposite direction. We can do this since for  $\Delta n < 0$  there is enough decay of the integrand in (2.29) as  $z \rightarrow \infty$ .

It is then helpful to rewrite (2.29) to

$$\mathbb{K}_{\text{gas}}(m, n; m', n') = \begin{cases} \frac{1}{2\pi i} \int_{\gamma} F(z) e^{N(\Phi_1(z; \xi_1, \xi_2) - \Phi_1(z; \xi'_1, \xi'_2))/2} \frac{dz}{z} & \text{if } \Delta m \geq 0, \\ \frac{1}{2\pi i} \int_{\gamma} (F(z) - I_2) e^{N(\Phi_2(z; \xi_1, \xi_2) - \Phi_2(z; \xi'_1, \xi'_2))/2} \frac{dz}{z} & \text{if } \Delta m < 0. \end{cases}$$

After a change of variables (6.43), using the expansions (6.33) and (6.38), and finally changing  $s$  to  $s^2$ , all as in the proof of Theorem 2.10, we find

$$\begin{aligned} &N^{1/4}(-1)^{(\Delta n - \Delta m)/2} \alpha^{-\Delta n} \mathbb{K}_{\text{gas}}(m, n; m', n') \\ &= \begin{cases} \frac{\sqrt{\alpha - \beta}}{2^{1/4}} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{\frac{1}{2}s^2(v-v') + s(u-u')} ds + o(1), & \Delta m \geq 0, \\ \frac{(-1)^{\Delta m} \sqrt{\alpha - \beta}}{2^{1/4}} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{\frac{1}{2}s^2(v-v') - s(u-u')} ds + o(1), & \Delta m < 0, \end{cases} \end{aligned} \tag{6.46}$$

as  $N \rightarrow \infty$ . The limit (2.35) then follows after observing that the integral is a well-known representation of the gaussian

$$\int_{-i\infty}^{i\infty} e^{\frac{1}{2}s^2(v-v') \pm s(u-u')} ds = \frac{\sqrt{2\pi} i}{\sqrt{v - v'}} e^{-\frac{(u-u')^2}{v-v'}} \quad \text{for } v > v'.$$

This concludes the proof of Theorem 2.11. □

*Acknowledgments.* We thank Sunil Chhita and Kurt Johansson for many fruitful discussions. We are also grateful to Christophe Charlier for a careful reading of a preliminary manuscript.

Most of this work was done when Arno Kuijlaars was a visiting professor at KTH in spring semester of 2017, funded by the Knut and Alice Wallenberg Foundation.

Maurice Duits is supported by the grants 2012.3128 and 2016.05450 of the Swedish Research Council and by the Göran Gustafsson foundation. Arno Kuijlaars is supported by long term structural funding—Methusalem grant of the Flemish Government, and via the support of FWO Flanders through projects G.0864.16, G.0910.20, and EOS grant 30889451.

## References

- [1] Adler, M., Cafasso, M., van Moerbeke, P.: From the Pearcey to the Airy process. *Electron. J. Probab.* **16**, 1048–1064 (2011) [Zbl 1231.60107](#) [MR 2820069](#)
- [2] Adler, M., Chhita, S., Johansson, K., van Moerbeke, P.: Tacnode GUE-minor processes and double Aztec diamonds. *Probab. Theory Related Fields* **162**, 275–325 (2015) [Zbl 1321.60103](#) [MR 3350046](#)
- [3] Adler, M., Johansson, K., van Moerbeke, P.: Double Aztec diamonds and the tacnode process. *Adv. Math.* **252**, 518–571 (2014) [Zbl 1335.60177](#) [MR 3144240](#)
- [4] Álvarez-Fernández, C., Mañas, M.: On the Christoffel–Darboux formula for generalized matrix orthogonal polynomials of multigraded-Hankel type. *J. Math. Anal. Appl.* **418**, 238–247 (2014) [Zbl 1305.33016](#) [MR 3198876](#)
- [5] Aptekarev, A., Nikishin, E.: The scattering problem for a discrete Sturm–Liouville operator, *Mat. Sb.* **121**, 327–358 (1983) (in Russian); English transl.: *Math. USSR Sb.* **49**, 325–357 (1984) [Zbl 0557.34017](#) [MR 0708000](#)
- [6] Ariznabarreta, G., Garcá-Ardila, J. C., Mañas, M., Marcellán, F.: Matrix biorthogonal polynomials on the real line: Geronimus transformations. *Bull. Math. Sci.* **9**, art. 1950007, 68 pp. (2019) [MR 3994624](#)
- [7] Baik, J., Deift, P., Suidan, T.: *Combinatorics and Random Matrix Theory*, Grad. Stud. Math. 172, Amer. Math. Soc., Providence, RI (2016) [Zbl 1342.05001](#) [MR 3468920](#)
- [8] Baik, J., Kriecherbauer, T., McLaughlin, K. T.-R., Miller, P. D.: *Discrete Orthogonal Polynomials*. Ann. of Math. Stud. 164, Princeton Univ. Press (2007) [Zbl 1119.41001](#) [MR 2283089](#)
- [9] Beffara, V., Chhita, S., Johansson, K.: Airy point process at the liquid-gas boundary. *Ann. Probab.* **46**, 2973–3013 (2018) [Zbl 1428.60063](#) [MR 3846843](#)
- [10] Bertola, M., Cafasso, M.: The transition between the gap probabilities from the Pearcey to the Airy process—a Riemann–Hilbert approach. *Int. Math. Res. Notices* **2012**, 1519–1568 [Zbl 1251.60005](#) [MR 2913183](#)
- [11] Bingham, N. H.: Multivariate prediction and matrix Szegő theory. *Probab. Surv.* **9**, 325–339 (2012) [Zbl 1285.60038](#) [MR 2956574](#)
- [12] Bleher, P. M., Kuijlaars, A. B. J.: Random matrices with external source and multiple orthogonal polynomials. *Int. Math. Res. Notices* **2004**, 109–129 [Zbl 1082.15035](#) [MR 2276453](#)
- [13] Bleher, P. M., Kuijlaars, A. B. J.: Large  $n$  limit of Gaussian random matrices with external source III: double scaling limit. *Comm. Math. Phys.* **270**, 481–517 (2007) [Zbl 1126.82010](#) [MR 2276453](#)
- [14] Borodin, A.: Determinantal point processes. In: *Oxford Handbook of Random Matrix Theory*. G. Akemann et al. (eds.), Oxford Univ. Press, 231–249 (2011) [Zbl 1238.60055](#) [MR 2932631](#)
- [15] Borodin, A., Rains, E.: Eynard–Mehta theorem, Schur process, and their Pfaffian analogs. *J. Statist. Phys.* **121**, 291–317 (2005) [Zbl 1127.82017](#) [MR 2185331](#)
- [16] Böttcher, A., Grudsky, S.: *Spectral Properties of Banded Toeplitz Matrices*. SIAM, Philadelphia, PA (2005) [Zbl 1089.47001](#) [MR 2179973](#)
- [17] Brézin, E., Hikami, S.: Level spacing of random matrices in an external source. *Phys. Rev. E* **58**, 7176–7185 (1998) [MR 1662382](#)
- [18] Bufetov, A., Knizel, A.: Asymptotics of random domino tilings of rectangular Aztec diamonds. *Ann. Inst. H. Poincaré Probab. Statist.* **54**, 1250–1290 (2018) [Zbl 1401.60174](#) [MR 3825881](#)
- [19] Cassatella-Contra, G. A., Mañas, M.: Riemann–Hilbert problems, matrix orthogonal polynomials and discrete matrix equations with singularity confinement. *Stud. Appl. Math.* **128**, 252–274 (2012) [Zbl 1246.33003](#) [MR 2899643](#)

- [20] Chhita, S., Johansson, K.: Domino statistics of the two-periodic Aztec diamond. *Adv. Math.* **294**, 37–149 (2016) [Zbl 1344.82021](#) [MR 3479561](#)
- [21] Chhita, S., Young, B.: Coupling functions for domino tilings of Aztec diamonds. *Adv. Math.* **259**, 173–251 (2014) [Zbl 1288.05038](#) [MR 3197656](#)
- [22] Cohn, H., Elkies, N., Propp, J.: Local statistics for random domino tilings of the Aztec diamond. *Duke Math. J.* **85**, 117–166 (1996) [Zbl 0866.60018](#) [MR 1412441](#)
- [23] Cohn, H., Kenyon, R., Propp, J.: A variational principle for domino tilings. *J. Amer. Math. Soc.* **14**, 297–346 (2001) [Zbl 1037.82016](#) [MR 1815214](#)
- [24] Daems, E., Kuijlaars, A. B. J.: A Christoffel–Darboux formula for multiple orthogonal polynomials. *J. Approx. Theory* **130**, 190–202 (2004) [Zbl 1063.42014](#) [MR 2100703](#)
- [25] Damanik, D., Pushnitski, A., Simon, B.: The analytic theory of matrix orthogonal polynomials, *Surv. Approx. Theory* **4**, 1–85 (2008) [Zbl 1193.42097](#) [MR 2379691](#)
- [26] Deift, P., Kriecherbauer, T., McLaughlin, K. T.-R., Venakides, S., Zhou, X.: 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**, 1335–1425 (1999) [Zbl 0944.42013](#) [MR 1702716](#)
- [27] Deift, P., Kriecherbauer, T., McLaughlin, K. T.-R., Venakides, S., Zhou, X.: Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.* **52**, 1491–1552 (1999) [Zbl 1026.42024](#) [MR 1711036](#)
- [28] Deift, P., Zhou, X.: A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the MKdV equation, *Ann. of Math.* **137**, 295–368 (1993) [Zbl 0771.35042](#) [MR 1207209](#)
- [29] Delvaux, S.: Average characteristic polynomials for multiple orthogonal polynomial ensembles. *J. Approx. Theory* **162**, 1033–1067 (2010) [Zbl 1202.42047](#) [MR 2610344](#)
- [30] Di Francesco, P., Soto-Garrido, R.: Arctic curves of the octahedron equation. *J. Phys. A.* **47**, art. 285204, 34 pp. (2014) [Zbl 1296.05121](#) [MR 3228361](#)
- [31] Duits, M., Kuijlaars, A. B. J.: Universality in the two-matrix model: a Riemann–Hilbert steepest-descent analysis. *Comm. Pure Appl. Math.* **62**, 1076–1153 (2009) [Zbl 1221.15052](#) [MR 2531553](#)
- [32] Duits, M., Kuijlaars, A. B. J.: The two periodic Aztec diamond and matrix valued orthogonal polynomials (extended version of present paper). [arXiv:1712.05636](#)
- [33] Durán, A. J., Grünbaum, F. A.: A survey on orthogonal matrix polynomials satisfying second order differential equations. *J. Comput. Appl. Math.* **178**, 169–190 (2005) [Zbl 1060.42017](#) [MR 2127878](#)
- [34] Durán, A. J., Van Assche, W.: Orthogonal matrix polynomials and higher-order recurrence relations. *Linear Algebra Appl.* **219**, 261–280 (1995) [Zbl 0827.15027](#) [MR 1327404](#)
- [35] Elkies, N., Kuperberg, G., Larsen, M., Propp, J.: Alternating sign matrices and domino tilings I, II. *J. Algebraic Combin.* **1**, 111–132, 219–234 (1992) [Zbl 0779.05009](#) [MR 1226347\(I\)](#) [MR 1194076\(II\)](#)
- [36] Eynard, B., Mehta, M. L.: Matrices coupled in a chain. I. Eigenvalue correlations. *J. Phys. A* **31**, 4449–4456 (1998) [Zbl 0938.15012](#) [MR 1628667](#)
- [37] Fokas, A., Its, A., Kitaev, A.: The isomonodromy approach to matrix models in 2D quantum gravity. *Comm. Math. Phys.* **147**, 395–430 (1992) [Zbl 0760.35051](#) [MR 1174420](#)
- [38] Gessel, I., Viennot, G.: Binomial determinants, paths, and hook length formulae. *Adv. Math.* **58**, 300–321 (1985) [Zbl 0579.05004](#) [MR 0815360](#)
- [39] Grünbaum, F. A., de la Iglesia, M. D., Martínez-Finkelshtein, A.: Properties of matrix orthogonal polynomials via their Riemann–Hilbert characterization. *SIGMA* **7**, art. 098, 31 pp. (2011) [Zbl 1242.42020](#) [MR 2861178](#)

- [40] Grünbaum, F. A., Pacharoni, I., Tirao, J.: Matrix valued spherical functions associated to the complex projective plane. *J. Funct. Anal.* **188**, 350–441 (2002) [Zbl 0996.43013](#) [MR 1883412](#)
- [41] Hachem, W., Hardy, A., Najim, J.: Large complex correlated Wishart matrices: the Pearcey kernel and expansion at the hard edge. *Electron. J. Probab.* **21**, art. 1, 36 pp. (2016) [Zbl 1336.15016](#) [MR 3485343](#)
- [42] Johansson, K.: Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. of Math.* **153**, 259–296 (2001) [Zbl 0984.15020](#) [MR 1826414](#)
- [43] Johansson, K.: Non-intersecting paths, random tilings and random matrices. *Probab. Theory Related Fields* **123**, 225–280 (2002) [Zbl 1008.60019](#) [MR 1900323](#)
- [44] Johansson, K.: The arctic circle boundary and the Airy process. *Ann. Probab.* **33**, 1–30 (2005) [Zbl 1096.60039](#) [MR 2118857](#)
- [45] Johansson, K.: Random matrices and determinantal processes. In: *Mathematical Statistical Physics*, A. Bovier et al. (eds.), Elsevier, Amsterdam, 1–55 (2006) [Zbl 1411.60144](#) [MR 2581882](#)
- [46] Johansson, K.: Edge fluctuations of limit shapes. In: *Current Developments in Mathematics 2016*, D. Jerison et al. (eds.), Int. Press, Somerville, MA, 47–110 (2018) [Zbl 1423.60167](#) [MR 2581882](#)
- [47] Kenyon, R.: Lectures on dimers. In: *Statistical Mechanics*, S. Sheffield and T. Spencer (eds.), Amer. Math. Soc., Providence, RI, 191–230 (2009) [Zbl 1180.82001](#) [MR 2523460](#)
- [48] Kenyon, R., Okounkov, A.: Limit shapes and the complex Burgers equation. *Acta Math.* **199**, 263–302 (2007) [Zbl 1156.14029](#) [MR 2358053](#)
- [49] Kenyon, R., Okounkov, A., Sheffield, S.: Dimers and amoebae. *Ann. of Math.* **163**, 1019–1056 (2006) [Zbl 1154.82007](#) [MR 2215138](#)
- [50] Koelink, E., van Pruijssen, M., Román, P.: Matrix-valued orthogonal polynomials related to  $(\mathrm{SU}(2) \times \mathrm{SU}(2), \mathrm{diag})$ . *Int. Math. Res. Notices* **2012**, 5673–5730 [Zbl 1263.33005](#) [MR 3006173](#)
- [51] Kuijlaars, A. B. J.: Multiple orthogonal polynomials in random matrix theory. In: *Proceedings ICM (Hyderabad, 2010)*, Vol. III, R. Bhatia et al. (eds.), Hindustan Book Agency, New Delhi, 1417–1432 (2010) [Zbl 1230.42034](#) [MR 2827849](#)
- [52] Kuijlaars, A. B. J., Martínez-Finkelshtein, A.: Strong asymptotics for Jacobi polynomials with varying nonstandard parameters. *J. Anal. Math.* **94**, 195–234 (2004) [Zbl 1126.33003](#) [MR 2124460](#)
- [53] Lindström, B.: On the vector representations of induced matroids. *Bull. London Math. Soc.* **5**, 85–90 (1973) [Zbl 0262.05018](#) [MR 0335313](#)
- [54] Martínez-Finkelshtein, A., Martínez-González, P., Orive, R.: Zeros of Jacobi polynomials with varying non-classical parameters. In: *Special Functions*, C. Dunkl et al. (eds.), World Sci., River Edge, NJ, 98–113 (2000) [Zbl 1221.33019](#) [MR 1805976](#)
- [55] Martínez-Finkelshtein, A., Orive, R.: Riemann–Hilbert analysis of Jacobi polynomials orthogonal on a single contour. *J. Approx. Theory* **134**, 137–170 (2005) [Zbl 1073.30031](#) [MR 2142296](#)
- [56] Nienhuis, B., Hilhorst, H. J., Blöte, H. W. J.: Triangular SOS models and cubic-crystal shapes. *J. Phys. A* **17**, 3559–3581 (1984) [MR 0772340](#)
- [57] Okounkov, A., Reshetikhin, N.: Correlation function of Schur process with application to local geometry of a random 3-dimensional Young diagram. *J. Amer. Math. Soc.* **16**, 581–603 (2003) [Zbl 1009.05134](#) [MR 1969205](#)
- [58] Okounkov, A., Reshetikhin, N.: Random skew plane partitions and the Pearcey process. *Comm. Math. Phys.* **269**, 571–609 (2007) [Zbl 1115.60011](#) [MR 2276355](#)

- 
- [59] Petrov, L.: Asymptotics of random lozenge tilings via Gelfand–Tsetlin schemes. *Probab. Theory Related Fields* **160**, 429–487 (2014) [Zbl 1315.60013](#) [MR 3278913](#)
- [60] Petrov, L.: Asymptotics of uniformly random lozenge tilings of polygons. *Gaussian free field. Ann. Probab.* **43**, 1–43 (2015) [Zbl 1315.60062](#) [MR 3298467](#)
- [61] Propp, J.: Generalized domino-shuffling. *Theoret. Computer Sci.* **303**, 267–301 (2003) [Zbl 1052.68095](#) [MR 1990768](#)
- [62] Szegő, G.: *Orthogonal Polynomials*. 4th ed., Amer. Math. Soc., Providence, RI (1975) [Zbl 0305.42011](#) [MR 0372517](#)
- [63] Tracy, C., Widom, H.: The Pearcey process. *Comm. Math. Phys.* **263**, 381–400 (2006) [Zbl 1129.82031](#) [MR 2207649](#)
- [64] Van Assche, W., Geronimo, J. S., Kuijlaars, A. B. J.: Riemann–Hilbert problems for multiple orthogonal polynomials. In: *Special Functions 2000*, J. Bustoz et al. (eds.), Kluwer, Dordrecht, 23–59 (2001) [Zbl 0997.42012](#) [MR 2006283](#)