

# Dimer models on astroidal zig-zag graphs

Tomas Berggren\*      Alexei Borodin†      Terrence George‡

May 5, 2026

## Abstract

On a finite weighted graph, the dimer model is a probability measure on its dimer covers, also known as perfect matchings, that assigns to any cover a probability proportional to the product of the weights of its edges. For planar bipartite graphs, dimer correlations are encoded by the inverse of the so-called Kasteleyn matrix, which is a version of the adjacency matrix for the graph.

The inverse Kasteleyn matrix for a large graph, typically taken as a finite domain in a periodic graph, is known explicitly only for a handful of examples. In all previously known examples, the Newton polygon – a convex lattice polygon that classifies periodic graphs up to local moves – is either a triangle or a quadrilateral.

Our main results are the following.

For any (minimal) periodic planar bipartite graph, we construct an  $(n - 3)$ -dimensional family of finite subgraphs for which we obtain an explicit inverse Kasteleyn matrix; here  $n$  is the number of sides of the Newton polygon. Their boundaries are formed by zig-zag paths and their overall shape is reminiscent of an astroid; we call them astroidal zig-zag graphs (AZ graphs). If the Newton polygon is the unit square then the corresponding AZ graph is the celebrated Aztec diamond with its size as the parameter.

Our inverse Kasteleyn matrices are given by a double contour integral on the corresponding spectral curve for any Fock weighting of the graph. This includes, in particular, all periodic weightings.

For periodic weightings, we asymptotically analyze the resulting inverse Kasteleyn matrices. We establish a phase separation in large AZ graphs into asymptotically frozen, rough (liquid), and smooth (gaseous) regions, and obtain an explicit parametrization of the ‘arctic curve’.

We also compute the deterministic limit of the height function, known as the limit shape, and prove the convergence of the local dimer correlations to the translation-invariant Gibbs measure of the slope predicted by the limit shape.

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Astroidal zig-zag graphs . . . . .	5
1.2	Formula for the inverse Kasteleyn matrix . . . . .	7
1.3	Phase regions and the arctic curve . . . . .	9
1.4	Limit shape . . . . .	11

---

\*Department of Mathematics & Statistics, University of South Florida, USA. E-mail: tberggren@usf.edu

†Department of Mathematics, Massachusetts Institute of Technology, USA. E-mail: borodin@math.mit.edu

‡Department of Mathematics, Massachusetts Institute of Technology, USA. E-mail: tegeorge@mit.edu

1.5	Local statistics . . . . .	12
1.6	Tropical limit and simulations . . . . .	12
1.7	Further directions . . . . .	13
<b>2</b>	<b>Preliminaries</b>	<b>14</b>
2.1	Conventions . . . . .	14
2.2	The dimer model . . . . .	15
2.3	M-curves . . . . .	18
2.4	Fock's dimer model . . . . .	21
<b>3</b>	<b>Astroidal zig-zag graphs</b>	<b>26</b>
3.1	Definition of astroidal zig-zag graphs . . . . .	26
3.2	Examples of AZ graphs . . . . .	29
3.3	Criterion for admissibility . . . . .	31
3.4	Characterization of AZ graphs . . . . .	32
3.5	Convex and concave vertices . . . . .	34
3.6	Chambers . . . . .	35
<b>4</b>	<b>Exact inverse Kasteleyn formula for AZ graphs</b>	<b>38</b>
4.1	Formula for $K^{-1}$ . . . . .	38
4.2	Example: Aztec diamond . . . . .	40
4.3	Example: a pentagonal Newton polygon with isoradial weights . . . . .	42
4.4	Correctness of the definition of $K^{-1}$ . . . . .	46
4.5	Proof of Theorem 4.3 . . . . .	51
<b>5</b>	<b>Scaling limits and astroidal domains</b>	<b>55</b>
5.1	Astroidal domains . . . . .	55
5.2	Cell decomposition and sign changes . . . . .	56
5.3	Characterization of astroidal domains . . . . .	57
<b>6</b>	<b>Phase regions and the arctic curve</b>	<b>59</b>
6.1	Action 1-form and action function . . . . .	59
6.2	Phase regions and boundary strata of the limit shape . . . . .	61
6.3	The rough region and the interior parametrization . . . . .	64
6.4	Boundary parametrization away from angles . . . . .	66
6.5	Turning points . . . . .	67
6.6	Global extension of $\Omega^{-1}$ and the arctic curve . . . . .	69
6.7	Geometry of the arctic curve . . . . .	70
6.8	Zeroes of $\lambda_c(x, y)$ and tangent lines to the arctic curve . . . . .	74
<b>7</b>	<b>Asymptotics: local statistics and the limit shape</b>	<b>75</b>
7.1	Local statistics . . . . .	76
7.2	The limit shape . . . . .	84
<b>A</b>	<b>Tropical limits and simulations</b>	<b>88</b>
<b>B</b>	<b>Proof of Proposition 7.10</b>	<b>92</b>

# 1 Introduction

The dimer model is a statistical mechanical model that studies random perfect matchings or dimer covers of a large graph. Via Thurston’s height function [108], random perfect matchings on planar bipartite graphs may be viewed as random discrete surfaces with prescribed boundary conditions. Classical examples include lozenge tilings and domino tilings, which have been studied extensively from the points of view of combinatorics, probability and integrable systems; see Kenyon’s lectures [81] and Gorin’s book [64] for an introduction.

A fundamental result in the subject is the variational principle, proved for domino tilings by Cohn–Kenyon–Propp [38] and generalized by Kenyon–Okounkov–Sheffield [84] and Kuchumov [86]. For a periodic planar bipartite graph with periodic edge weights, and for suitable exhaustions by finite subgraphs, the associated random height functions concentrate, after scaling, near a deterministic *limit shape* characterized as the minimizer of an explicit surface tension functional. This can also be viewed as a variant of the Wulff construction [110]; see also Nienhuis [93], Dobrushin–Kotecký–Shlosman [46], and Cerf–Kenyon [30].

The limit shape is highly sensitive to the choice of finite subgraphs, or, equivalently, on the boundary conditions. For example, the limit shape of a growing sequence of rectangles on the square lattice is constant, whereas for sequences of Aztec diamonds it develops distinct phases separated by a curve known as the *arctic curve*, as shown by Jockusch–Propp–Shor [71]. This raises a natural problem: how does one systematically choose sequences of finite graphs that exhibit nontrivial phase transitions?

The general variational theory is not accompanied by an equally general supply of concrete models with this behavior. Outside a small number of classical examples—lozenge tilings, domino tilings, and the Schur process/rail-yard graph models of Okounkov–Reshetikhin [97] and Boutillier–Bouttier–Chapuy–Cortee–Ramassamy [29]—there are no theoretical results or any simulations. The first main contribution of this paper is to identify natural finite subgraphs called *astroidal zig-zag graphs* (or *AZ graphs* for short) of an arbitrary (minimal) periodic planar bipartite graph exhibiting nontrivial phase transitions.

Once the limit shape has been identified, the next step is to study local and global fluctuations around it. In the smooth part of the limit shape, it is conjectured—and proved for the hexagonal lattice with uniform weights by Aggarwal [4] (see also Gorin [63])—that the local statistics are governed by the translation-invariant ergodic Gibbs measure of Kenyon–Okounkov–Sheffield [84] whose slope matches that of the limit shape. It is also of considerable interest to study the local fluctuations near non-smooth points, such as regular points of the arctic curve as in Johansson [73], Chhita–Dauvergne–Finn [33], Huang [66] and Aggarwal–Huang [6]; turning points as in Johansson–Nordenstam [74], Okounkov–Reshetikhin [95], Aggarwal–Gorin [5] and Berggren–Bradinoff [14]; cusps as in Tracy–Widom [109], Okounkov–Reshetikhin [96], Huang–Yang–Zhang [67]; and at other special points as discussed in Astala–Duse–Prause–Zhong [8]. Likewise, understanding the global fluctuations around the limit shape remains an important open problem; see Kenyon [80] and Kenyon–Okounkov [83]. Kasteleyn theory [75, 107, 76] provides a powerful tool for tackling these problems through the inverse Kasteleyn matrix, which, by a result of Kenyon [79], encodes all probabilistic information about the dimer model.

Explicit formulas for the inverse Kasteleyn matrix (equivalently, for the correlation kernel of the associated nonintersecting lattice path model) are, however, known only in certain special cases, including:

- domino tilings of the Aztec diamond: the uniform case by Helfgott [65] and Johansson [73];  $q^{\text{vol}}$  and other one-periodic weightings due to Chhita–Young [36] and Chhita–Johansson–Young [35]; the two-periodic and, more generally, doubly periodic models studied by Chhita–Young [36], Chhita–Johansson [34], Duits–Kuijlaars [48], Berggren–Duits [15], Berggren [10], Borodin–Duits [20], Berggren–Borodin [12], and Boutillier–de Tilière [27]; the double Aztec diamond of Adler–Chhita–Johansson–van Moerbeke [1]; and the split Aztec diamond of Shea [105].
- lozenge tilings of hexagonal and polygonal regions: the uniform hexagon by Johansson [72];  $q^{\text{vol}}$ -weighted hexagons studied by Borodin–Gorin–Rains [22]; polygonal domains studied by Petrov [100]; holey hexagons studied by Gilmore [61]; non-convex polygons studied by Adler–Johansson–van Moerbeke [2]; and periodic hexagons studied by Charlier–Duits–Kuijlaars–Lenells [31] and Kuijlaars [87].
- more general graph families related to the Schur processes of Okounkov–Reshetikhin [97], rail-yard graphs of Boutillier–Bouttier–Chapuy–Corteele–Ramassamy [29], and tower graphs by Borodin–Ferrari [21] and Nicoletti [92].

The complexity of a periodic bipartite graph is captured by an invariant called the Newton polygon. All the cases listed above correspond to Newton polygons that are triangles or quadrilaterals. *By contrast, AZ graphs exist for arbitrary Newton polygons and, to the best of our knowledge, provide the first analysis of finite subgraphs beyond triangles and quadrilaterals.*

The second main contribution of the present paper is a formula for the inverse Kasteleyn matrix for AZ graphs, generalizing the recent Aztec diamond formula of Boutillier–de Tilière [27] (which is a streamlined version of the formula obtained in Berggren–Borodin [12]). This formula allows for a detailed asymptotic analysis similar to that performed in Berggren–Borodin [12] for the Aztec diamond. In particular, we obtain an explicit parameterization of the limit shape, prove that local statistics are governed by the ergodic Gibbs measures classified in Kenyon–Okounkov–Sheffield [84], and establish geometric properties of the arctic curve analogous to Kenyon–Okounkov [83], Berggren–Borodin [12], and Bobenko–Bobenko [18]. In this sense, AZ graphs provide a general new class of dimer models whose large-scale behavior can be described with a level of detail previously available only in a few special examples.

Although we do not currently have an exact sampling algorithm for random dimer covers on general AZ graphs, we show in the appendix that certain AZ graphs can be sampled via a tropical limit of the Aztec diamond. This yields concrete simulations that can be compared with our theoretical predictions.

As we were writing this paper, we learned that Boutillier–de Tilière–Deb [28] have independent work that derives a formula for the inverse Kasteleyn matrix for Speyer’s “crosses and wrenches” graphs [106]. We expect that our class of graphs overlaps with theirs but neither class includes the other.

Let us now describe our results in more detail. Let  $G$  be a minimal  $\mathbb{Z}^2$ -periodic planar bipartite graph, and let  $G_{\mathbb{T}} := G/\mathbb{Z}^2$  be the associated torus graph. We denote the black vertices, white vertices and faces of  $G$  by  $B(G)$ ,  $W(G)$  and  $F(G)$  respectively. A *zig-zag path* of  $G$  or  $G_{\mathbb{T}}$  is a path that turns maximally right at white vertices and maximally left at black vertices. Every zig-zag path of  $G$  is a lift of a zig-zag path in  $G_{\mathbb{T}}$ . We write  $\alpha$  for the set of zig-zag paths of  $G$ .

The homology classes of the zig-zag paths of  $G_{\mathbb{T}}$  determine a convex lattice polygon  $N$ , called the *Newton polygon* of  $G$ ; see Section 2.2.3. By a theorem of Goncharov–Kenyon, minimal periodic planar bipartite graphs are classified, up to certain local moves, by their Newton polygon. These

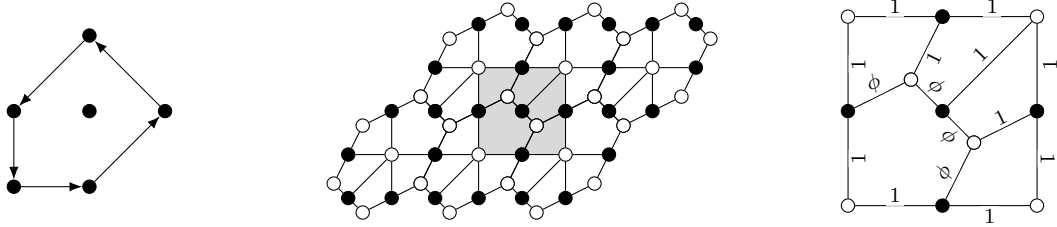


Figure 1: A pentagonal lattice polygon  $N$  (left), a minimal graph  $G$  with Newton polygon  $N$  together with a fundamental domain (middle), and a choice of isoradial weights (modulo gauge) where the angles are chosen evenly spaced around the unit circle (right). Here  $\phi = (1 + \sqrt{5})/2$  denotes the golden ratio. See Figure 23 below for an illustration of the isoradial property.

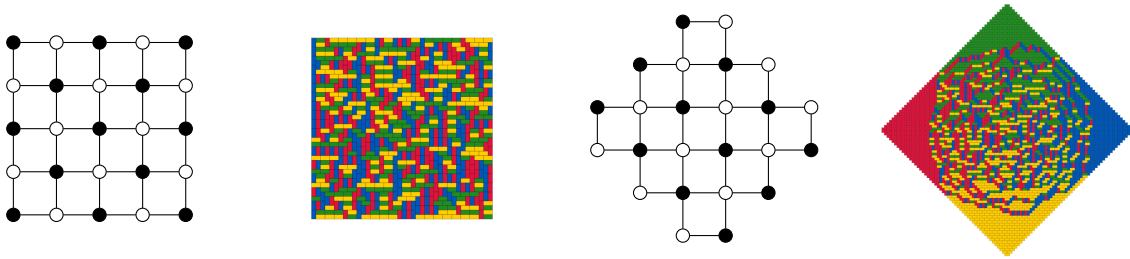


Figure 2: Two natural finite domains for the square lattice are the square and the Aztec diamond, together with their associated bipartite graphs. Only the Aztec diamond gives rise to a limit shape with a nontrivial phase transition. The images were generated using Leonid Petrov's domino tiling simulator; see <https://lpetrov.cc/simulations/>.

local moves do not affect the statistical-mechanical properties of the model. Figure 1 shows an example of a minimal graph for a convex lattice pentagon.

Let  $V(N)$  and  $E(N)$  denote the vertices and edges of  $N$ , respectively, where by an edge we mean a side of  $N$  rather than a primitive edge vector. For each  $\alpha \in \alpha$ , let  $e(\alpha) \in E(N)$  be the edge of  $N$  corresponding to the homology class of  $\alpha$ . For each edge  $e \in E(N)$ , define

$$\alpha_e := \{\alpha \in \alpha : e(\alpha) = e\}.$$

We say that the zig-zag paths in  $\alpha_e$  are *parallel to*  $e$ .

### 1.1 Astroidal zig-zag graphs

Fix a convex lattice polygon  $N$ , and let  $G$  be a graph with Newton polygon  $N$ . Our first main contribution is the definition of a natural class of finite subgraphs of  $G$  that give rise to limit shapes with nontrivial phase transitions. To motivate this definition, consider the square lattice. Two natural families of finite domains are squares and Aztec diamonds. As illustrated in Figure 2, only the Aztec diamond produces a limit shape exhibiting a phase transition. Our goal is therefore to isolate the features of the Aztec diamond responsible for this phenomenon.

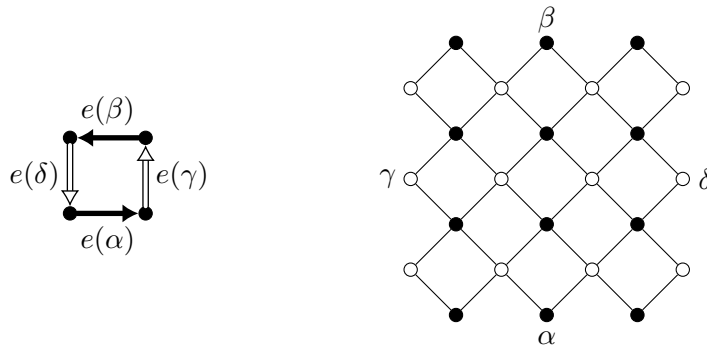


Figure 3: The square Newton polygon  $N$  (left) and the Aztec diamond (right). The cyclic order of zig-zag paths along the boundary of the Aztec diamond is opposite to the cyclic order of the corresponding edges of  $N$ . The edges of  $N$  are colored according to whether the black or white vertices are outside.

First of all, we observe that the boundary of the Aztec diamond is formed by zig-zag paths whose cyclic order along the boundary is opposite to the cyclic order of the corresponding edges of  $N$  (see Figure 3). This condition alone, however, is not sufficient: one can also construct rectangular regions satisfying it, but such regions do not admit dimer covers. This leads to a second requirement, which we call *admissibility*. These observations motivate the definition of the class of AZ graphs.

An AZ graph is constructed by choosing, for each edge  $e \in E(N)$ , a zig-zag path  $\beta_e$  parallel to  $e$ , such that these boundary zig-zag paths appear in the cyclic order opposite to that of the corresponding edges of  $N$ . We write

$$\boldsymbol{\beta} := (\beta_e)_{e \in E(N)}$$

for the resulting collection of boundary zig-zag paths. They cut out a finite region of  $\mathbf{G}$ , and the resulting subgraph  $\mathbf{G}_\beta \subset \mathbf{G}$  is the portion of  $\mathbf{G}$  enclosed by the corresponding zig-zag boundary  $\partial \mathbf{G}_\beta$ . Each boundary zig-zag path  $\beta_e$  also determines a color of the edge  $e$ , according to whether the black or white vertices lie outside  $\mathbf{G}_\beta$ . We now give an informal definition of AZ graphs; see Definition 3.2 for the precise definition.

**Definition 1.1.** The graph  $\mathbf{G}_\beta$  is called an *astroidal zig-zag graph*, or *AZ graph*, if its zig-zag boundary is simple and it is *admissible*, meaning that for any face, the sum of the distances to  $\beta_e$  over black edges  $e$  is equal to the sum of the distances to  $\beta_e$  over white edges  $e$ .

Here the distances are measured using the *discrete Abel map*

$$\mathbf{d} : W(\mathbf{G}) \sqcup F(\mathbf{G}) \sqcup B(\mathbf{G}) \rightarrow \mathbb{Z}^\alpha,$$

which assigns to each vertex or face an integer linear combination of zig-zag paths recording which zig-zag paths we cross as we move from a fixed reference face  $f_0$  to that vertex or face. In Lemma 3.13, we show that every AZ graph has exactly four convex vertices, connected by boundary arcs that may contain concave vertices. Thus, these graphs resemble the astroid curve, which explains the term “astroidal,” while “zig-zag” refers to the fact that the boundary is formed by zig-zag paths.

In Proposition 3.10 and Corollary 4.4, we show that among finite subgraphs whose boundary consists of zig-zag paths, one for each edge of the Newton polygon, in opposite cyclic order, the

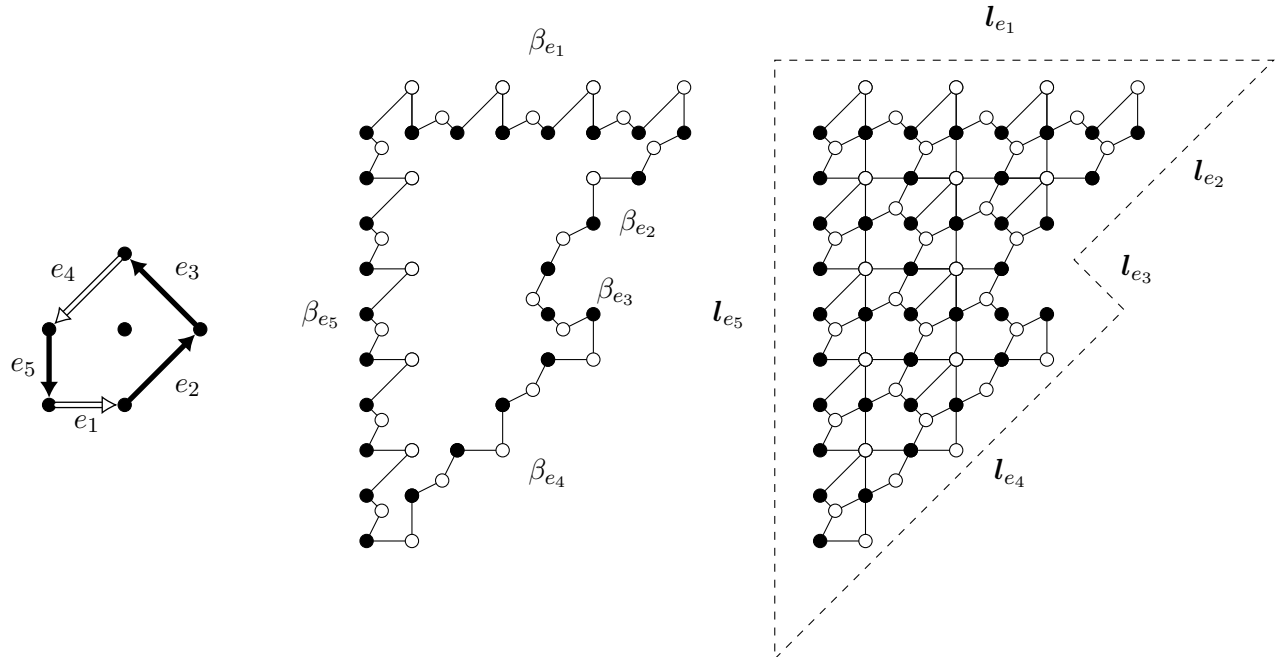


Figure 4: The construction of an AZ graph for the graph  $G$  in Figure 1. We choose zig-zag paths forming the boundary  $\partial G_\beta$  (middle), and obtain  $G_\beta$  as the enclosed subgraph (right). The dashed polygon is an astroidal domain that  $G_\beta$  approximates.

ones that admit dimer covers are precisely the AZ graphs. Thus admissibility is not an auxiliary technical condition, but exactly the condition for the graph to carry a nontrivial dimer model.

AZ graphs form a  $(\#E(N) - 3)$ -parameter family. Thus, when  $N$  is a triangle, no AZ graphs exist. This can be seen directly: any finite subgraph bounded by three zig-zag paths, one in each direction, has unequal numbers of black and white vertices. When  $N$  is the unit square, the Aztec diamond is, up to size and local moves, the unique AZ graph.

## 1.2 Formula for the inverse Kasteleyn matrix

We next turn to the dimer model itself. The periodic graph  $G$  is equipped with Fock's dimer model, built from the spectral data of an M-curve  $\Sigma$  called the *spectral curve*, a point  $t$  in the Jacobian  $\text{Jac}(\Sigma)$ , and an *angle function*  $\nu$  assigning to each zig-zag path a point in a distinguished component  $A_0$  of the real part  $\Sigma(\mathbb{R})$ . As is customary, we identify a zig-zag path with the corresponding angle in  $A_0$ . We write

$$\Sigma(\mathbb{R}) = A_0 \sqcup \left( \bigsqcup_{j=1}^g A_j \right),$$

and denote by  $\Sigma_+$  the closure of one of the two connected components of  $\Sigma \setminus \Sigma(\mathbb{R})$ . We also write  $\theta$  for the Riemann theta function on  $\Sigma$  and  $E$  for the prime form. We recall the necessary background on M-curves, the Abel map, the theta function, the prime form and Fock's dimer model in Sections 2.3–2.4. For the purposes of the introduction, it is sufficient to know that out of this geometric data, one can construct a Kasteleyn matrix  $K$  called *Fock's Kasteleyn matrix*. Moreover, Fock [52] showed

that this construction realizes all periodic weights on  $\mathbf{G}$  (see also Boutillier–Cimasoni–de Tilière [26] and George–Goncharov–Kenyon [58]).

We now describe the whole-plane inverse Kasteleyn matrices introduced in Boutillier–Cimasoni–de Tilière [26, Section 3.5], which will be used both in the definition of  $\mathbf{K}^{-1}$  for finite AZ graphs and in the description of local statistics. For a black vertex  $\mathbf{b}$  and a white vertex  $\mathbf{w}$  of  $\mathbf{G}$ , there is an associated meromorphic form  $g_{\mathbf{b},\mathbf{w}}$  on  $\Sigma$ , lying in the kernel/cokernel of  $\mathbf{K}$ . For  $\zeta \in \Sigma_+ \setminus \alpha$ , set

$$A_{\mathbf{b},\mathbf{w}}^\zeta := \frac{1}{2\pi i} \int_{C_{\mathbf{b},\mathbf{w}}^\zeta} g_{\mathbf{b},\mathbf{w}},$$

where  $C_{\mathbf{b},\mathbf{w}}^\zeta$  is a contour from  $\sigma(\zeta)$  to  $\zeta$ , with  $\sigma : \Sigma \rightarrow \Sigma$  the antiholomorphic involution of the M-curve; see Section 2.4.5. Then, for each  $\zeta$ ,  $A^\zeta$  is an inverse of Fock’s Kasteleyn matrix on the full graph  $\mathbf{G}$ .

For  $\zeta \in A_0 \setminus \alpha$ , the matrix  $A^\zeta$  depends only on the connected component of  $A_0 \setminus \alpha$  containing  $\zeta$ . These components are naturally indexed by the vertices  $u \in V(N)$ , and we denote the corresponding whole-plane inverse by  $A^u$ .

We now state the main theorem of the paper which gives an explicit formula for  $\mathbf{K}^{-1}$  for AZ graphs. Since the dimer model is determinantal, this formula gives complete access to all local correlations on  $\mathbf{G}_\beta$ . The boundary conditions determine a divisor  $\mathbf{D}_\beta$ , defined by recording, via the discrete Abel map, the signed distances from the reference face  $f_0$  to the zig-zag paths forming the boundary of  $\mathbf{G}_\beta$ . For a black vertex  $\mathbf{b}$  and a white vertex  $\mathbf{w}$  of  $\mathbf{G}_\beta$ , we define

$$\omega_{\mathbf{b},\mathbf{w}}(\zeta, \eta) := \frac{1}{(2\pi i)^2} \frac{1}{E(\zeta, \eta)} \frac{\theta(-t + \mathbf{D}_\beta + \eta - \zeta)}{\theta(-t + \mathbf{D}_\beta)} \frac{g_{f_0,\mathbf{w}}(\eta) \Psi_\beta(\eta)}{g_{f_0,\mathbf{b}}(\zeta) \Psi_\beta(\zeta)},$$

where

$$\Psi_\beta(\zeta) := \prod_{\alpha \in \alpha} E(\alpha, \zeta)^{(\mathbf{D}_\beta)_\alpha}$$

is the product of prime forms associated with the boundary divisor  $\mathbf{D}_\beta$ . The key point is that admissibility is precisely the condition under which this kernel is globally well-defined on  $\Sigma \times \Sigma$ .

**Remark 1.2.** While we use the formalism of Fock’s dimer model to state the results in full generality, our results are already new and interesting in the case of genus-zero spectral curves. From the probabilistic point of view, this case is the familiar critical/isoradial dimer model; see Kenyon [78] and Kenyon–Okounkov [82, Section 5]. The genus-zero case is especially concrete: the theta functions are identically equal to 1 and the prime form  $E(\alpha, \zeta)$  can be written as  $\zeta - \alpha$ , so the kernel reduces to an explicit rational function. For example, choosing the angles to be evenly spaced around the unit circle for the pentagonal Newton polygon in Figure 1(left) yields the isoradial weights shown in Figure 1(right).

For each variable, the zeros and poles of  $\omega_{\mathbf{b},\mathbf{w}}$  partition the boundary of  $N$  into four intervals, alternating between zeros and poles. To each of  $\mathbf{b}$  and  $\mathbf{w}$ , we assign one of the two pole intervals, denoted

$$I_{\mathbf{b}} = [\ell_{\mathbf{b}}, r_{\mathbf{b}}], \quad I_{\mathbf{w}} = [\ell_{\mathbf{w}}, r_{\mathbf{w}}].$$

We then let  $C(I_{\mathbf{b}})$  and  $C(I_{\mathbf{w}})$  be small counterclockwise-oriented contours around the corresponding angles on  $\Sigma$ .

The inverse Kasteleyn matrix is then defined by

$$K_{b,w}^{-1} = \iint_{C(I_w) \prec C(I_b)} \omega_{b,w} + \mathbf{1}_{\{\ell_b \in I_w\}} A_{b,w}^{\ell_b} - \mathbf{1}_{\{r_b \in I_w\}} A_{b,w}^{r_b}, \quad (1.1)$$

where  $C(I_w) \prec C(I_b)$  denotes the product contour  $C(I_b) \times C(I_w)$  with the convention that whenever both contours contain small circles around a common angle  $\alpha$ , the circle in  $C(I_w)$  has smaller radius than the one in  $C(I_b)$ . This formula has the same general structure as many correlation kernels appearing in the dimer and integrable-probability literature: a double contour integral with a single-integral correction term whose integrand is the residue along the diagonal  $\zeta = \eta$ .

**Theorem 1.3** (cf. Theorem 4.3). *The matrix  $K^{-1}$  defined by (1.1) is the two-sided inverse of Fock's Kasteleyn matrix  $K$  on  $\mathbf{G}_\beta$ .*

In particular, this immediately yields the following consequence which is not obvious from the definition of AZ graphs.

**Corollary 1.4** (cf. Corollary 4.4). *Every AZ graph has equally many black and white vertices and admits a dimer cover.*

We obtained the formula (1.1) experimentally, guided by the corresponding formula for the Aztec diamond in [27]. The integrand has the same structure, while the contours and the single-integral correction term are more complicated. It would, however, be very interesting to find a direct derivation explaining this formula.

We now turn to the asymptotic consequences of Theorem 1.3 which are obtained by a steepest descent analysis of the exact formula (1.1).

### 1.3 Phase regions and the arctic curve

We now consider limits of AZ graphs as the size tends to infinity. Let  $\mathbf{G}_\beta$  be an AZ graph. For each side  $e \in E(N)$ , we label the zig-zag paths in the family  $\alpha_e$  from left to right as

$$\alpha_e = \{\alpha_e^i : i \in \mathbb{Z} + \frac{1}{2}\},$$

choosing the indexing so that the reference face  $f_0$  lies between  $\alpha_e^{-1/2}$  and  $\alpha_e^{1/2}$ . We then define

$$c_\beta : E(N) \rightarrow \mathbb{Z} + \frac{1}{2}$$

by the requirement that

$$\alpha_e^{c_\beta(e)} = \beta_e.$$

In other words,  $c_\beta(e)$  measures the location of the boundary zig-zag path  $\beta_e$  among the zig-zag paths parallel to  $e$ .

Now consider a sequence of AZ graphs  $\mathbf{G}_{\beta^n}$ ,  $n = 1, 2, \dots$ , such that for every  $e \in E(N)$ ,

$$c_{\beta^n}(e) = n|e|_{\mathbb{Z}}c(e) + O(1) \quad \text{as } n \rightarrow \infty,$$

for some function  $c : E(N) \rightarrow \mathbb{R}$ , where  $|e|_{\mathbb{Z}}$  is the lattice length of  $e$ . In the limit, the admissibility condition for  $\mathbf{G}_{\beta^n}$  becomes

$$\sum_{e \in E(N)} |e|_{\mathbb{Z}}c(e) = 0. \quad (1.2)$$

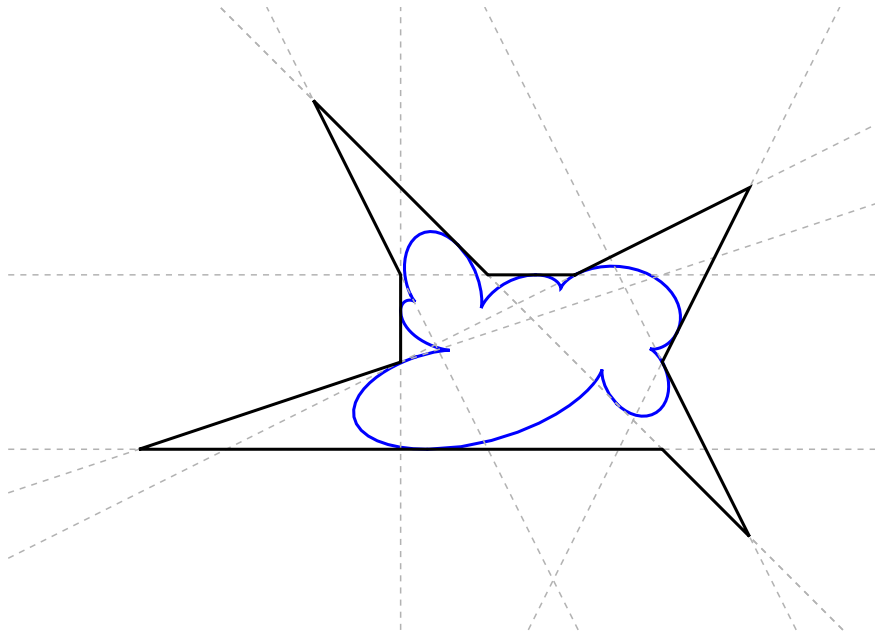


Figure 5: An astroidal domain for a decagonal Newton polygon, and an arctic curve for a choice of isoradial weights.

For each  $e \in E(N)$ , the boundary zig-zag path  $\beta_e^n$  has as its scaling limit the line

$$\mathbf{l}_e := \{(x, y) \in \mathbb{R}^2 : -b_e x + a_e y + c(e) = 0\}, \quad e \in E(N),$$

where

$$(a_e, b_e) := \frac{\vec{e}}{|e|_{\mathbb{Z}}} \in \mathbb{Z}^2$$

is the primitive integer vector in the direction of  $e$  viewed as a vector in the plane. The resulting collection of lines  $(\mathbf{l}_e)_{e \in E(N)}$  determines a domain

$$\mathcal{D}_c \subset \mathbb{R}^2,$$

called an *astroidal domain*.<sup>1</sup> This domain is the scaling limit of the AZ graphs  $\mathbf{G}_{\beta^n}$ ; see Figure 4(right).

In Section 5.3, we show that if a simple polygonal domain carries a continuous piecewise linear boundary height function whose slopes are prescribed by the Newton polygon, then it is an astroidal domain. In this sense, astroidal domains are exactly the domains compatible with the desired boundary behavior of the limit shape.

For the asymptotic results below, we assume that the angle function is periodic. This assumption is slightly more general than periodicity of the edge weights; see [26, Section 4.2]. The asymptotics of the dimer model are governed by a meromorphic 1-form

$$\lambda_c(x, y)$$

---

<sup>1</sup>Rote–Santos–Streinu [103] called similar-looking polygons *pseudo-quadrilaterals*, in parallel to the well-established term of *pseudo-triangles*. They did not involve the continuous version of the admissibility condition (1.2), however. We are grateful to Richard Kenyon and Pavel Galashin for pointing out the connection with pseudo-quadrilaterals, and to Pavel Galashin for also drawing our attention to [103].

on  $\Sigma$ , called the *action 1-form*, which is characterized as the unique imaginary-normalized meromorphic 1-form satisfying

$$\operatorname{res}_\alpha(\lambda_c(x, y)) = -b_{e(\alpha)}x + a_{e(\alpha)}y + c(e(\alpha)), \quad \text{for all } \alpha \in \alpha.$$

The signs of the residues determine the locations of all but two zeroes of  $\lambda_c(x, y)$ . The two remaining zeroes are either both real or form a  $\sigma$ -conjugate pair. This leads to a phase decomposition of  $\mathcal{D}_c^\circ$ : the *rough* (or *liquid*) region  $\mathcal{L}$ , where the remaining zeroes form a  $\sigma$ -conjugate pair; the *frozen* regions  $\mathcal{F}_{(\alpha, \beta)}$ , where the remaining zeroes lie in an open interval  $(\alpha, \beta) \subset A_0$  between consecutive angles with  $e(\alpha) \neq e(\beta)$ ; the *quasi-frozen* regions  $\mathcal{Q}_{(\alpha, \beta)}$ , where the remaining zeroes lie in an interval  $(\alpha, \beta) \subset A_0$  with  $e(\alpha) = e(\beta)$ ; and the *smooth* (or *gaseous*) regions  $\mathcal{G}_j$ ,  $1 \leq j \leq g$ , where the remaining zeroes lie on the oval  $A_j$ . The boundary between the rough and non-rough phases is the *arctic curve*, denoted by  $\mathcal{A}$ .

In the rough region there is a unique zero  $\zeta \in \Sigma_+^\circ$ , and this defines a map

$$\Omega : \mathcal{L} \rightarrow \Sigma_+^\circ, \quad \Omega(x, y) = \zeta.$$

**Proposition 1.5** (cf. Proposition 6.17). *The map  $\Omega : \mathcal{L} \rightarrow \Sigma_+^\circ$  is a diffeomorphism. Its inverse extends smoothly to a map*

$$\Omega^{-1} : \Sigma_+ \rightarrow \overline{\mathcal{L}}$$

which restricts to a bijection

$$\Omega^{-1} : \Sigma(\mathbb{R}) \rightarrow \mathcal{A}.$$

Thus, the real locus  $\Sigma(\mathbb{R})$  gives a global parametrization of the arctic curve. Figure 5 shows an example of the arctic curve plotted using this global parameterization for a decagonal Newton polygon and a choice of isoradial weights.

Several methods have been developed to compute arctic curves. One prominent approach is the tangent method of Colomo–Sportiello [39] and its variants; see Di Francesco–Lapa [43], Di Francesco–Guitter [41, 42], Aggarwal [3], and Ruelle [104]. Another approach uses analytic combinatorics in several variables, particularly for dimer models arising from the octahedron recurrence and  $T$ -systems; see Petersen–Speyer [99], Baryshnikov–Pemantle [9], Pemantle–Wilson–Melczer [98], Di Francesco–Soto-Garrido [44], Di Francesco–Vu [45], and references therein. In contrast, the inverse Kasteleyn approach gives access not only to the arctic curve, but also to the local statistics throughout the limit shape.

## 1.4 Limit shape

Let  $h_M$  be the height function on  $\mathbf{G}_{\beta^n}$ , normalized with respect to the extremal matching corresponding to a fixed vertex  $u_0 \in V(N)$ . If  $\mathbf{f}_n$  is a face whose location is  $n(x, y) + O(1)$ , define

$$\mathfrak{h}(x, y) := \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{n} h_M(\mathbf{f}_n) \right].$$

For periodic dimer models the existence of this limit is guaranteed by the variational principle. In our setting of AZ graphs, we derive an explicit parameterization for  $\mathfrak{h}(x, y)$ .

Define  $A_{(x, y)}$  by

$$A_{(x, y)} := \begin{cases} (\alpha, \beta), & \text{if } (x, y) \text{ lies in a frozen or quasi-frozen region corresponding to } (\alpha, \beta) \subset A_0, \\ A_j, & \text{if } (x, y) \in \mathcal{G}_j. \end{cases}$$

**Theorem 1.6** (cf. Theorem 7.9). *For  $(x, y) \in \mathcal{D}_c^\circ$ , let  $\zeta_* = \Omega(x, y)$  if  $(x, y) \in \mathcal{L}$ , and otherwise let  $\zeta_*$  be any point of  $A_{(x,y)}$ . Then*

$$\mathfrak{h}(x, y) = -\frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_*}} \lambda_c(x, y),$$

where  $C_{u_0}^{\zeta_*}$  is a contour from  $\sigma(\zeta_*)$  to  $\zeta_*$  intersecting the component of  $A_0 \setminus \alpha$  corresponding to  $u_0$ .

## 1.5 Local statistics

Fix a finite graph  $C \subset G$  and translate it to a macroscopic point  $(x, y) \in \mathcal{D}_c^\circ$  by setting

$$C_n = (j_n, k_n) + C, \quad (j_n, k_n) = n(x, y) + O(1) \in \mathbb{Z}^2.$$

**Theorem 1.7** (cf. Theorem 7.2). *Let  $(x, y) \in \mathcal{D}_c^\circ \setminus \mathcal{A}$ . In the rough region, let  $\zeta_* = \Omega(x, y)$ , and in a frozen, quasi-frozen or smooth region, let  $\zeta_*$  be any point in the corresponding component  $A_{(x,y)} \subset \Sigma(\mathbb{R}) \setminus \alpha$ . If  $(x, y) \in \mathcal{F}_{(\alpha, \beta)}$ , assume in addition that*

$$(x, y) \notin \mathbf{l}_{e(\alpha)} \cup \mathbf{l}_{e(\beta)}.$$

For a black vertex  $\mathbf{b}$  and a white vertex  $\mathbf{w}$  of  $C$ , let

$$\mathbf{b}_n := \mathbf{b} + (j_n, k_n), \quad \mathbf{w}_n := \mathbf{w} + (j_n, k_n)$$

denote their translates in  $C_n$ . Then

$$K_{\mathbf{b}_n, \mathbf{w}_n}^{-1} \rightarrow A_{\mathbf{b}, \mathbf{w}}^{\zeta_*} \quad \text{as } n \rightarrow \infty,$$

uniformly over  $\mathbf{b}$  and  $\mathbf{w}$ .

**Corollary 1.8** (cf. Corollary 7.12). *The local correlations in a neighborhood of  $(x, y)$  of the dimer model on  $G_{\beta^n}$  defined by the Kasteleyn matrix (2.4) converge to those of the Gibbs measure defined by the corresponding whole-plane Kasteleyn matrix with slope  $\nabla \mathfrak{h}(x, y)$ .*

**Remark 1.9.** If the edge weights of the dimer model are periodic, the Gibbs measures obtained in Corollary 1.8 are the translation-invariant Gibbs measures classified by Kenyon–Okounkov–Sheffield [84].

## 1.6 Tropical limit and simulations

In Appendix A, we outline, via an example, how AZ graphs can be indirectly simulated inside the Aztec diamond by choosing a suitable tropicalization in which the tropical spectral curve is not smooth. This generalizes Berggren–Borodin [11], which treats smooth tropical curves that do not leave any dimer randomness. Figure 6 shows one such simulation, containing an astroidal domain for the pentagonal Newton polygon in Figure 1 together with the arctic curve predicted by our theory. Understanding this phenomenon was one of the starting points for the present work, as it naturally led us to the notion of astroidal domains. Trying to understand their discrete counterparts in turn led us to AZ graphs.

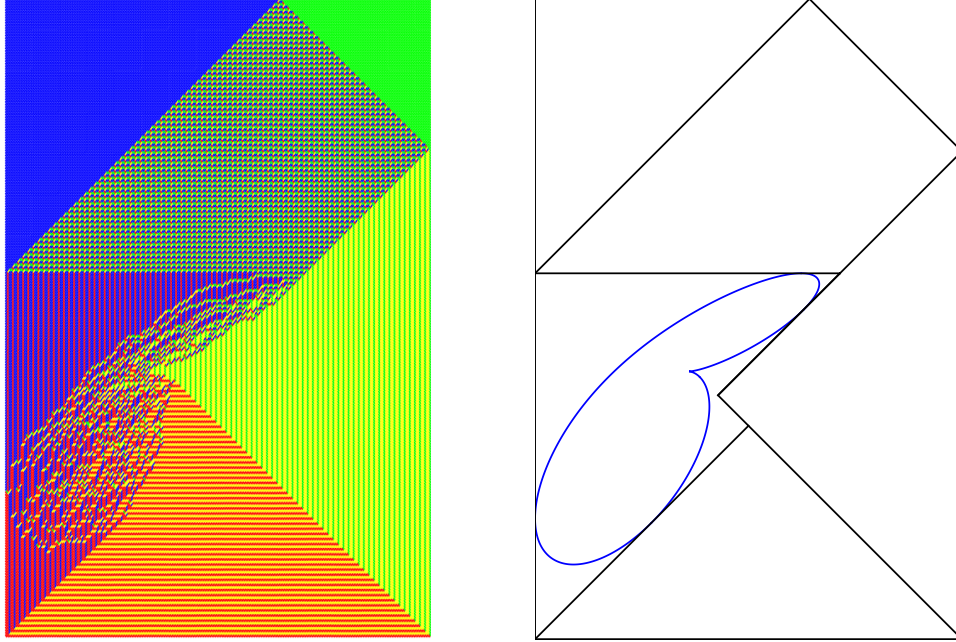


Figure 6: A simulation of the arctic curve in the tropical limit of the Aztec diamond (left) and the analytically computed arctic curve (right). The Aztec diamond has been vertically rescaled by a factor of  $3/2$  to match  $\mathcal{D}_c$ .

### 1.7 Further directions

We conclude the introduction with a few directions that we believe are promising for further study.

- It might be possible to give a direct proof that the limit shape of Theorem 1.6 solves the variational problem, in the spirit of Bobenko–Bobenko [18] for the Aztec diamond.
- By construction, each edge of the Newton polygon contributes exactly one zig-zag path to the boundary of an AZ graph. However, there are natural domains that do not fit into this framework, for example lozenge tilings of a hexagon, or, more generally, the polygonal domains of Kenyon–Okounkov [83], where a given parallel class of zig-zag paths may appear several times around the boundary. It would be very interesting to extend our constructions to this setting.
- If we drop the assumption that the zig-zag boundary of an AZ graph is simple, we obtain graphs with branching as in Lai–Musiker [89, Figures 46–49]. As the number of edges of  $N$  grows, the subset of such graphs among admissible graphs grows, and we would hope that our results extend to this setting.
- AZ graphs form a  $(\#E(N) - 3)$ -parameter family. On the other hand, the cluster modular group of generalized shufflings also gives rise to a  $(\#E(N) - 3)$ -parameter family, as shown by George–Inchiesta [59] (see also Fock–Marshakov [53] and George–Ramassamy [60]). It is therefore natural to expect that AZ graphs can be studied via shuffling, generalizing the relationship between domino shuffling and the Aztec diamond.

- Is it possible to extend the results of Berggren–Borodin [11] from the case of smooth tropical curves to the non-smooth tropical curves in Appendix A?
- It would be very nice to characterize which AZ graphs can be simulated inside the Aztec diamond as in Figure 6 and Appendix A. For instance, this procedure never produces a graph with self-intersecting boundary.

## Organization of the paper

The paper is organized as follows. Section 2 collects preliminaries on the dimer model. Section 3 introduces AZ graphs and develops their basic combinatorics. Section 4 proves the exact inverse Kasteleyn formula for AZ graphs. Section 5 turns to scaling limits and introduces astroidal domains. Section 6 analyzes the geometry of the phase regions and arctic curve. Section 7 uses steepest descent analysis to derive the local statistics and the explicit formula for the limit shape. Appendix A explains how certain astroidal domains arise from a tropical limit of the Aztec diamond and provides simulations illustrating the theory.

## Acknowledgements

We are grateful to Cédric Boutillier, Béatrice de Tilière, and Bishal Deb for communicating to us their ongoing work [28]. We also thank Cédric Boutillier, Béatrice de Tilière, and Marianna Russkikh for sharing with us the proof of Proposition 7.10, and for allowing us to reproduce it here. We thank Gregg Musiker for drawing our attention to further examples of AZ graphs in the literature. A. B. was partially supported by the NSF grant DMS-2450323. A. B. and T. G. were also partially supported by the Simons Investigator program.

## 2 Preliminaries

### 2.1 Conventions

Since we will simultaneously refer to vertices and edges of both the graph and the Newton polygon, we use the usual font for objects associated with the polygon and the sans-serif font for those associated with the graph. Whenever it makes sense, we write

$$V(\cdot), \quad E(\cdot), \quad F(\cdot)$$

for the sets of vertices, edges, and faces of the object under consideration (for example, for graphs, polygons, strands, etc).

Given a lattice polygon  $N \subset \mathbb{R}^2$ , we orient its boundary  $\partial N$  counterclockwise. This induces a cyclic order on  $\partial N$ : for  $u, v, w \in V(N)$ , we write  $u < v < w$  if they appear in counterclockwise order around  $\partial N$ . By an *edge* of  $N$  we mean a side of the polygon (possibly containing interior lattice points), while a *vertex* always means a corner. For  $a, b \in \partial N$ , we write

$$I = [a, b]$$

for the closed cyclic interval in  $\partial N$  from  $a$  to  $b$ , oriented counterclockwise.

We use the subscript  $-$  (resp.  $+$ ) to indicate the object immediately clockwise (resp. counterclockwise) from a given one. For example, when we write

$$v = e_- \cap e_+,$$

the edge  $e_-$  (resp.  $e_+$ ) is the edge immediately clockwise (resp. counterclockwise) from  $v$ . Similarly, for an oriented edge we write

$$e = [v_-, v_+],$$

where  $v_-$  and  $v_+$  are its clockwise and counterclockwise endpoints. We freely identify edges of  $N$  with the corresponding closed intervals in  $\partial N$  and with their displacement vectors in  $\mathbb{R}^2$ . Thus we may write an edge either as  $e = [v_-, v_+]$  or as the vector  $\vec{e}$ .

For a side  $e \in E(N)$ , let  $|e|_{\mathbb{Z}}$  denote the *integral length* of the vector  $e$ , *i.e.* the number of primitive lattice segments contained in  $e$ , or equivalently  $|e \cap \mathbb{Z}^2| - 1$ . Let

$$(a_e, b_e) := \frac{\vec{e}}{|e|_{\mathbb{Z}}} \in \mathbb{Z}^2$$

denote the primitive vector along  $\vec{e}$ .

## 2.2 The dimer model

Let  $\mathbf{G}$  be a planar bipartite graph. We denote the black vertices, white vertices, edges and faces of  $\mathbf{G}$  by  $B(\mathbf{G}), W(\mathbf{G}), E(\mathbf{G})$  and  $F(\mathbf{G})$  respectively. If an edge  $\mathbf{e} \in E(\mathbf{G})$  is incident to a black vertex  $\mathbf{b} \in B(\mathbf{G})$  and a white vertex  $\mathbf{w} \in W(\mathbf{G})$ , we write  $\mathbf{e} = \mathbf{bw}$ . A *dimer cover* on  $\mathbf{G}$  is a subset of  $E(\mathbf{G})$  such that every vertex is used exactly once.

An *edge weight* on  $\mathbf{G}$  is a function  $\mathbf{wt} : E(\mathbf{G}) \rightarrow \mathbb{R}_{>0}$ . Two edge weights  $\mathbf{wt}_1$  and  $\mathbf{wt}_2$  are said to be *gauge equivalent* if there is a function  $\lambda : B(\mathbf{G}) \sqcup W(\mathbf{G}) \rightarrow \mathbb{R}_{>0}$  such that for every edge  $\mathbf{e} = \mathbf{bw}$ , we have  $\mathbf{wt}_2(\mathbf{e}) = \lambda(\mathbf{b})\mathbf{wt}_1(\mathbf{e})\lambda(\mathbf{w})$ .

When  $\mathbf{G}$  is finite, the *Boltzmann measure* associated to an edge weight  $\mathbf{wt}$  assigns to each dimer cover  $M$  the probability

$$\mathbb{P}(M) := \frac{\mathbf{wt}(M)}{Z},$$

where  $\mathbf{wt}(M) := \prod_{\mathbf{e} \in M} \mathbf{wt}(\mathbf{e})$  and  $Z := \sum_{\text{dimer covers } M} \mathbf{wt}(M)$  is the *partition function*. Gauge equivalent edge weights define the same Boltzmann measure.

Motivated by this, we define a *dimer model* to be a pair  $(\mathbf{G}, [\mathbf{wt}])$ , where  $\mathbf{G}$  is a (possibly infinite) planar bipartite graph and  $[\mathbf{wt}]$  is a gauge equivalence class of edge weights.

### 2.2.1 The height function

Let  $\mathbf{G}$  be a planar bipartite graph. Each dimer cover  $M$  can be viewed as a white-to-black flow  $\omega_M$  on  $\mathbf{G}$  that sends 1 along each edge of  $M$  and 0 along all other edges. Fix a reference face  $\mathbf{f}_0$  and a reference dimer cover  $M_0$ . Following Thurston [108], we define a function

$$h_M : F(\mathbf{G}) \rightarrow \mathbb{Z},$$

called the *height function* of  $M$ , as follows. For any face  $\mathbf{f}$ , let  $\gamma^*$  be a path in the dual graph  $\mathbf{G}^*$  of  $\mathbf{G}$  from  $\mathbf{f}_0$  to  $\mathbf{f}$  and let

$$h_M(\mathbf{f}) := (M - M_0) \wedge \gamma^* = \sum_{\mathbf{e}=\mathbf{wb}} (\mathbf{e} \wedge \gamma^*) (\mathbf{1}_{\{\mathbf{e} \in M\}} - \mathbf{1}_{\{\mathbf{e} \in M_0\}}),$$

where  $\wedge$  denotes the intersection pairing in the plane with the convention that

$$e \wedge \gamma^* = \begin{cases} -1, & \text{if } \gamma^* \text{ crosses } e \text{ from left to right,} \\ 1, & \text{if } \gamma^* \text{ crosses } e \text{ from right to left.} \end{cases} \quad (2.1)$$

The height function allows us to interpret a random dimer cover of  $G$  as a random discrete surface, which is sometimes a more suitable object for studying scaling limits.

### 2.2.2 The Kasteleyn matrix

The fundamental tool in the study of the dimer model is the Kasteleyn matrix which we now introduce. A *Kasteleyn sign* is a function  $\kappa : E(G) \rightarrow \mathbb{C}$  such that  $|\kappa(e)| = 1$  for all  $e \in E(G)$  and the alternating product of  $\kappa$  around the boundary  $\partial f$  of each face  $f \in F(G)$  is  $(-1)^{\frac{|\partial f|}{2}+1}$ . The *Kasteleyn matrix* is the  $W(G) \times B(G)$  weighted signed bipartite adjacency matrix defined by

$$K_{w,b} := \sum_{e=bw} \text{wt}(e) \kappa(e).$$

**Theorem 2.1** ([75, 107]). *If  $G$  is finite, the partition function satisfies*

$$Z = |\det K|.$$

A fundamental consequence is that the inverse Kasteleyn matrix encodes all local statistics of the dimer model.

**Theorem 2.2** ([79]). *Let  $G$  be finite and let  $\{e_i = b_i w_i\}_{i=1}^k \subset E(G)$  be a fixed collection of its edges. Then the probability that all  $e_i$  occur in a random dimer configuration  $M$  is*

$$\mathbb{P}(e_1, \dots, e_k \in M) = \prod_{i=1}^k K_{w_i, b_i} \det(K_{b_i, w_j}^{-1})_{1 \leq i, j \leq k}.$$

### 2.2.3 Minimal periodic bipartite graphs

Let  $G \subset \mathbb{R}^2$  be a  $\mathbb{Z}^2$ -periodic graph embedded in the plane such that every face is a topological disk. Let  $\mathbb{T} := \mathbb{R}^2/\mathbb{Z}^2$  be the torus,  $p : \mathbb{R}^2 \rightarrow \mathbb{T}$  be the quotient map and let  $G_{\mathbb{T}} := G/\mathbb{Z}^2$  be the corresponding graph in  $\mathbb{T}$ .

A *zig-zag path* in  $G$  is an oriented path that turns maximally right at white vertices and maximally left at black vertices. Equivalently, zig-zag paths can be viewed as certain paths in the medial graph. Recall that the (directed) medial graph  $G^\times$  has a vertex at the midpoint of each edge of  $G$ , and an edge whenever two edges of  $G$  are consecutive around a face. Thus we have a bijection

$$V(G^\times) \cong E(G),$$

and we will use this identification hereafter without comment. We orient the edges of  $G^\times$  so that they go clockwise around black vertices and counterclockwise around white vertices. Every vertex of  $G^\times$  has degree four, and the cyclic order of the incident edges determines two pairs of opposite edges. A (*medial strand*) is a directed path in  $G^\times$  with the property that whenever the path enters a vertex, it exits along the opposite edge (*i.e.*, it goes straight through the vertex). The bijection

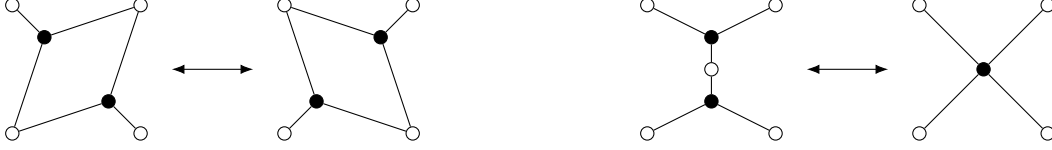


Figure 7: Local moves: spider move (left) and contraction-uncontraction move (right).

with zig-zag paths is obtained as follows. Given a zig-zag path in  $G$ , record the sequence of edges it traverses; the corresponding strand is the path in  $G^\times$  visiting the midpoints of these edges in order. The maximal-turn rule at black and white vertices is equivalent to the condition that the strand goes straight at each medial vertex. Zig-zag paths and strands in  $G_{\mathbb{T}}$  are defined in the same way. We say that a vertex *lies on* a strand if it lies on the corresponding zig-zag path.

We say that  $G$  is *minimal* if:

1. no strand is a closed loop,
2. no strand has a self-intersection,
3. there is no pair of strands  $\alpha, \beta$  (or, rather, zig-zag paths) that have two edges  $e_1 \neq e_2 \in E(G)$  in common such that both of them are oriented from  $e_1$  to  $e_2$ .

Minimal graphs have no multiple edges, leaves or isolated vertices. Informally, the strands in a minimal graph “behave like straight lines”.

Every strand  $\alpha$  in  $G_{\mathbb{T}}$  is a cycle and therefore, has a homology class  $[\alpha] \in H_1(\mathbb{T}, \mathbb{Z}) \cong \mathbb{Z}^2$ . The unique convex lattice polygon  $N \subset \mathbb{R}^2$ , defined modulo translation by  $\mathbb{Z}^2$ , whose sides consist of the vectors  $[\alpha] \in \mathbb{Z}^2$  for all strands  $\alpha$  in  $G_{\mathbb{T}}$  in counterclockwise cyclic order is called the *Newton polygon* of  $G$ . This construction produces a closed polygon because the homology classes of all strands in  $G_{\mathbb{T}}$  sum to zero: a strand has the same homology class as the corresponding zig-zag path and every edge of the graph is traversed by two zig-zag paths with opposite orientations.

**Theorem 2.3** ([62, Theorem 2.5]). *For any convex lattice polygon  $N \subset \mathbb{R}^2$ , there is a nonempty family of minimal graphs with Newton polygon  $N$ . Moreover, any two graphs in the same family are related by sequences of the two types of local moves shown in Figure 7.*

These local moves do not affect the statistical-mechanical properties of the model like the limit shape or local limits.

We denote by  $\alpha$  (resp.  $\alpha_{\mathbb{T}}$ ) the set of strands of  $G$  (resp.  $G_{\mathbb{T}}$ ). Each strand  $\alpha \in \alpha$  projects to a strand  $p(\alpha) \in \alpha_{\mathbb{T}}$ . To each strand  $\alpha \in \alpha$ , we associate the unique edge  $e(\alpha) \in E(N)$  containing the homology class  $[p(\alpha)]$ . We say that  $\alpha$  is *parallel to*  $e(\alpha)$ . For each edge  $e \in E(N)$ , let

$$\alpha_e := \{\alpha \in \alpha : e(\alpha) = e\}$$

be the set of strands parallel to  $e$ . The counterclockwise orientation of  $\partial N$  induces a partial cyclic order on  $\alpha$  with respect to which two members of  $\alpha_e$  for any  $e$  are incomparable.

**Lemma 2.4** ([24, Lemma 8]; see also [62, Lemma 3.4]). *Let  $G$  be a minimal graph, and let  $x \in B(G) \sqcup W(G)$  be a vertex. Then the strands that  $x$  lies on are parallel to distinct edges of  $N$ . Moreover, the cyclic order of these strands (or the corresponding zig-zag paths) around  $x$  agrees with the cyclic order of the corresponding edges along  $\partial N$  if  $x$  is white, and with the opposite cyclic order if  $x$  is black.*

## 2.3 M-curves

We briefly recall the basic material on M-curves that we will use, following the notation and conventions of [26].

An *M-curve* is a compact Riemann surface  $\Sigma$  of genus  $g$  together with an anti-holomorphic involution  $\sigma : \Sigma \rightarrow \Sigma$  such that the fixed-point set

$$\Sigma(\mathbb{R}) := \{\zeta \in \Sigma : \sigma(\zeta) = \zeta\},$$

called the *real locus*, consists of  $g + 1$  connected components called *ovals*, each homeomorphic to a circle. We denote these components by  $A_0, \dots, A_g$ . The real locus separates  $\Sigma$  into two connected components

$$\Sigma \setminus \Sigma(\mathbb{R}) = \Sigma_+^\circ \sqcup \Sigma_-^\circ.$$

Let  $\Sigma_\pm := \overline{\Sigma_\pm^\circ}$  denote their closures. We fix the orientations of the ovals  $A_j$  so that

$$\partial\Sigma_+ = A_0 - \sum_{j=1}^g A_j.$$

Let  $B_1, \dots, B_g$  be dual cycles such that  $\sigma_* B_j = -B_j$  and

$$A_j \wedge B_k = \delta_{jk}$$

where  $\wedge : H_1(\Sigma, \mathbb{Z}) \times H_1(\Sigma, \mathbb{Z}) \rightarrow \mathbb{Z}$  is the intersection pairing. Let  $\omega_1, \dots, \omega_g$  be the dual basis of holomorphic 1-forms on  $\Sigma$  such that

$$\int_{A_j} \omega_k = \delta_{jk}.$$

It is known as the basis of normalized holomorphic 1-forms.

### 2.3.1 Abel map and Jacobian

The  $g \times g$ -matrix  $\Omega$  defined by  $\Omega_{jk} = \int_{B_j} \omega_k$  is called the *period matrix*. For an M-curve, the matrix  $\Omega$  is purely imaginary. The complex torus

$$\text{Jac}(\Sigma) := \mathbb{C}^g / (\mathbb{Z}^g \oplus \Omega\mathbb{Z}^g)$$

is called the *Jacobian variety* of  $\Sigma$ .

The *Abel map* with basepoint  $\zeta_0 \in A_0$  is the map

$$\mu : \Sigma \rightarrow \text{Jac}(\Sigma), \quad \mu(\zeta) := \left( \int_{\zeta_0}^{\zeta} \omega_j \right)_{j=1}^g.$$

The Abel map is independent of the path from  $\zeta_0$  to  $\zeta$  because the integrals of  $\omega_j$  along cycles are contained in  $\mathbb{Z}^g \oplus \Omega\mathbb{Z}^g$ .

A *divisor* on  $\Sigma$  is a finite formal  $\mathbb{Z}$ -linear combination of points of  $\Sigma$ ,

$$D = \sum_{j=1}^n a_j \zeta_j.$$

Its *degree* is

$$\deg(D) := \sum_{j=1}^n a_j.$$

If  $s$  is a meromorphic section of a holomorphic line bundle on  $\Sigma$ , we define its divisor

$$\operatorname{div}(s) := \sum_{\zeta \in \Sigma} \operatorname{ord}_{\zeta}(s)\zeta,$$

where  $\operatorname{ord}_{\zeta}$  denotes the order of vanishing at  $\zeta$ .

We extend the Abel map to divisors by linearity,

$$\mu(D) := \sum_{j=1}^n a_j \mu(\zeta_j).$$

By abuse of notation, we will henceforth suppress the Abel map from the notation, writing  $\zeta$  for  $\mu(\zeta)$  and similarly  $D$  for its image  $\mu(D) \in \operatorname{Jac}(\Sigma)$ .

### 2.3.2 Theta function and prime form

The *Riemann theta function* is the multi-valued function on  $\operatorname{Jac}(\Sigma)$  defined by

$$\theta(\zeta) := \sum_{n \in \mathbb{Z}^g} e^{i\pi \langle n, \Omega n \rangle + 2\pi i \langle n, \zeta \rangle}.$$

Pulling back via the Abel map, we obtain a multi-valued function on  $\Sigma$ . By the same abuse of notation as in Section 2.3.1, we will usually suppress the Abel map and write  $\theta(\zeta)$  for  $\theta(\mu(\zeta))$ .

The *prime form*  $E(\zeta, \eta)$  is a holomorphic  $(-\frac{1}{2}, -\frac{1}{2})$ -form on  $\Sigma \times \Sigma$  that is antisymmetric, vanishes simply along the diagonal with the expansion

$$E(\zeta, \eta) = \frac{\eta - \zeta}{\sqrt{d\zeta} \sqrt{d\eta}} (1 + O((\eta - \zeta)^2)),$$

and has no other zeroes and poles.

### 2.3.3 Meromorphic differentials of the third kind

We follow [51, Chapter I]. A *meromorphic differential of the third kind* on  $\Sigma$  is a meromorphic 1-form all of whose poles are simple. Although Fay [51] states the construction for integral divisors, all formulas are linear in the coefficients, so they extend immediately to degree-zero divisors with real coefficients. If

$$D = \sum_{j=1}^n a_j \zeta_j$$

is a degree-zero real divisor, meaning that  $a_j \in \mathbb{R}$  and

$$\sum_{j=1}^n a_j = 0,$$

then, after fixing a basepoint  $\zeta_0 \in \Sigma$ , there exists a unique meromorphic 1-form  $\omega_D$  with vanishing  $A$ -periods,

$$\int_{A_k} \omega_D = 0, \quad 1 \leq k \leq g,$$

whose only singularities are simple poles at the points  $\zeta_j$ , with residues

$$\operatorname{res}_{\zeta_j}(\omega_D) = a_j.$$

It is given by

$$\omega_D(\zeta) := d \sum_{j=1}^n a_j \log \left( \frac{E(\zeta_j, \zeta)}{E(\zeta_0, \zeta)} \right); \quad (2.2)$$

here  $d$  is the differential. By the *Riemann bilinear relations*, for  $1 \leq k \leq g$ ,

$$\int_{B_k} \omega_D = 2\pi i \sum_{j=1}^n a_j \int_{\zeta_0}^{\zeta_j} \omega_k. \quad (2.3)$$

### 2.3.4 Imaginary-normalized differentials

Instead of normalizing a meromorphic differential of the third kind to have vanishing  $A$ -periods, we may normalize it by requiring all of its periods to be purely imaginary. We say that a meromorphic 1-form is *imaginary-normalized* if

$$\Re \int_{\gamma} \omega = 0$$

for every cycle  $\gamma$  in  $\Sigma$ .

**Lemma 2.5** ([85, Lemma 2.1]). *Let  $D = \sum_{j=1}^n a_j \zeta_j$  be a degree-0 real divisor on  $\Sigma$ . Then there exists a unique imaginary-normalized meromorphic 1-form on  $\Sigma$  whose only singularities are simple poles at the points  $\zeta_j$  with residues  $a_j$ .*

Recall the meromorphic 1-forms  $\omega_D$  associated to degree-0 divisors  $D$  with real coefficients defined in Section 2.3.3. We now show that, when the support of  $D$  lies on  $A_0$ , this differential is also imaginary-normalized.

**Lemma 2.6.** *For any degree-0 real divisor  $D = \sum_{j=1}^n a_j \zeta_j$  such that  $\zeta_j \in A_0$ , the 1-form  $\omega_D$  is imaginary-normalized and satisfies  $\overline{\sigma^* \omega_D} = \omega_D$ ; in particular,  $\omega_D$  is real-valued on  $\Sigma(\mathbb{R})$ .*

*Proof.* Since the  $A$ -periods of  $\omega_D$  vanish by definition, it remains only to show that the  $B$ -periods are purely imaginary. By the Riemann bilinear relations (2.3),

$$\int_{B_k} \omega_D = 2\pi i \sum_{j=1}^n a_j \int_{\zeta_0}^{\zeta_j} \omega_k,$$

which is purely imaginary since  $\int_{\zeta_0}^{\zeta_j} \omega_k \in \mathbb{R}$  for any  $\zeta_j \in A_0$  by [26, Lemma 14] as  $\zeta_0 \in A_0$ .

The 1-form  $\overline{\sigma^* \omega_D}$  is imaginary-normalized since

$$\Re \int_{\gamma} \overline{\sigma^* \omega_D} = \Re \int_{\sigma_*(\gamma)} \omega_D = 0.$$

Since  $A_0$  is fixed by  $\sigma$ ,  $\overline{\sigma^*\omega_D}$  has the same poles as  $\omega_D$ . To see that the residues are also the same, let  $C$  be a small counterclockwise-oriented circle around  $\alpha$ . Then  $\sigma_*C = -C$ , so

$$\operatorname{res}_\alpha(\overline{\sigma^*\omega_D}) = \frac{1}{2\pi i} \int_C \overline{\sigma^*\omega_D} = \frac{1}{2\pi i} \int_{-C} \omega_D = \operatorname{res}_\alpha(\omega_D).$$

By Lemma 2.5,  $\overline{\sigma^*\omega_D} = \omega_D$ . □

Throughout the paper, we will count zeroes with multiplicity.

**Lemma 2.7.** *Let  $D = \sum_{j=1}^n a_j \zeta_j$  be a degree-0 real divisor such that  $a_j \neq 0$ ,  $\zeta_j \in A_0$  for all  $j$ , and index the points  $\zeta_1, \dots, \zeta_n$  cyclically along  $A_0$ . Then the differential  $\omega_D$  satisfies:*

1. *for each  $1 \leq k \leq g$ , the number of zeroes of  $\omega_D$  on  $A_k$  is even and at least 2,*
2. *for each interval  $(\zeta_j, \zeta_{j+1}) \subset A_0$ , the number of zeroes of  $\omega_D$  in  $(\zeta_j, \zeta_{j+1})$  is odd if  $\operatorname{res}_{\zeta_j}(\omega_D)$  and  $\operatorname{res}_{\zeta_{j+1}}(\omega_D)$  have the same sign, and even if they have opposite signs. In particular,  $\omega_D$  has at least one zero in  $(\zeta_j, \zeta_{j+1})$  whenever the residues at  $\zeta_j$  and  $\zeta_{j+1}$  have the same sign.*

*Proof.* Both assertions are consequences of the real-valuedness of  $\omega_D$  on  $\Sigma(\mathbb{R})$  (Lemma 2.6).

Let  $1 \leq k \leq g$ . Since  $\omega_D$  is imaginary-normalized, we have

$$\int_{A_k} \omega_D = 0.$$

Since  $\omega_D$  is not identically zero on  $A_k$ , it cannot have constant sign on  $A_k$ . Thus, it must change sign, which forces the existence of at least two zeroes on  $A_k$ . Since sign changes correspond to zeroes of odd multiplicity and the total number of sign changes is even, the total number of zeroes counted with multiplicity is even.

Now consider an interval  $(\zeta_j, \zeta_{j+1}) \subset A_0$ . Since  $\omega_D$  is real-valued on this interval, the parity of the number of zeroes is determined by its sign near the endpoints. Near  $\zeta_j$ , the sign of  $\omega_D$  on the interval is the sign of  $\operatorname{res}_{\zeta_j}(\omega_D)$ , while near  $\zeta_{j+1}$ , it is the opposite of the sign of  $\operatorname{res}_{\zeta_{j+1}}(\omega_D)$ . Therefore, these endpoint signs are opposite precisely when  $\operatorname{res}_{\zeta_j}(\omega_D)$  and  $\operatorname{res}_{\zeta_{j+1}}(\omega_D)$  have the same sign. In that case the number of zeroes in  $(\zeta_j, \zeta_{j+1})$  is odd, hence, at least one. If the residues have opposite signs, then the endpoint signs agree, so the number of zeroes is even. □

## 2.4 Fock's dimer model

In this subsection we recall Fock's construction of a dimer model from data associated to an M-curve following [52, 26].

### 2.4.1 Angle functions

Let  $\Sigma$  be an M-curve. We call a function  $\nu : \alpha \rightarrow A_0$  an *angle* function if it is compatible with the partial cyclic order on  $\alpha$  and the total cyclic order on  $A_0$ , and it maps non-parallel strands to distinct points. By a customary abuse of notation, we identify the angle  $\nu(\alpha)$  with the strand  $\alpha$ .

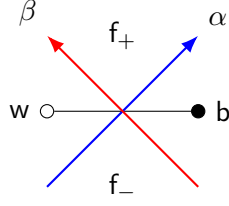


Figure 8: Local configuration near an edge in the definition of Fock's Kasteleyn matrix  $K$ .

### 2.4.2 The discrete Abel map

Let  $\Sigma$  be an M-curve and let  $\nu$  be an angle function. Via  $\nu$ , we henceforth identify  $\mathbb{Z}^\alpha$  with the group of divisors supported on  $\nu(\alpha) \subset A_0$ . In particular, for  $\mathbf{D} \in \mathbb{Z}^\alpha$ , we define its *degree* to be the degree of the corresponding divisor.

For  $\mathbf{D} \in \mathbb{Z}^\alpha$  and  $e \in E(N)$ , let

$$\mathbf{D}_{\alpha_e} \in \mathbb{Z}^{\alpha_e}$$

denote the restriction of  $\mathbf{D}$  to  $\alpha_e$ , so that

$$\mathbf{D} = \sum_{e \in E(N)} \mathbf{D}_{\alpha_e}.$$

For  $\alpha \in \alpha$ , let

$$\mathbf{D}_\alpha \in \mathbb{Z}^\alpha \cong \mathbb{Z}$$

denote the restriction of  $\mathbf{D}$  to  $\alpha$ ; we identify this with the coefficient of  $\alpha$  in  $\mathbf{D}$ .

The *discrete Abel map*

$$\mathbf{d} : W(\mathbb{G}) \sqcup F(\mathbb{G}) \sqcup B(\mathbb{G}) \rightarrow \mathbb{Z}^\alpha$$

is defined as follows:

1. Choose a reference face  $f_0$  and set  $\mathbf{d}(f_0) = 0$ .
2. For a white vertex  $w$  incident to a face  $f$  such that they are separated by a strand  $\alpha$ , we have  $\mathbf{d}(w) = \mathbf{d}(f) - \alpha$ .
3. For a black vertex  $b$  incident to a face  $f$  such that they are separated by a strand  $\alpha$ , we have  $\mathbf{d}(b) = \mathbf{d}(f) + \alpha$ .

Equivalently,

$$\mathbf{d}(y) - \mathbf{d}(x) = \sum_{\alpha \in \alpha} (\gamma \wedge \alpha) \alpha,$$

where  $\gamma$  is any path from  $x$  to  $y$  that intersects all strands transversely and  $\wedge$  denotes the intersection pairing in the plane with the same convention as in (2.1). By definition,  $\deg(\mathbf{d}(x))$  is  $-1$  if  $x$  is a white vertex,  $0$  if  $x$  is a face and  $1$  if  $x$  is a black vertex.

### 2.4.3 Fock's Kasteleyn matrix

Given an M-curve  $\Sigma$ , a point  $t \in \mathbb{R}^g/\mathbb{Z}^g \subset \text{Jac}(\Sigma)$  and an angle function  $\nu$ , *Fock's Kasteleyn matrix*  $K$  is defined by

$$K_{w,b} := \begin{cases} \frac{E(\alpha, \beta)}{\theta(t + \mathbf{d}(f_-))\theta(t + \mathbf{d}(f_+))}, & \text{if } \mathbf{e} = \mathbf{bw} \text{ is an edge,} \\ 0, & \text{otherwise,} \end{cases} \quad (2.4)$$

where  $\alpha, \beta$  are the two strands containing  $\mathbf{e}$ , and  $f_-, f_+$  are the two faces incident to  $\mathbf{e}$  (Figure 8).

**Proposition 2.8** ([26, Proposition 3]). *Fock's Kasteleyn matrix is the Kasteleyn matrix of a dimer model  $(G, [\mathbf{wt}])$ .*

### 2.4.4 Forms in the kernel of $K$

Following [52] (see also [82]), we define some differential forms on  $\Sigma$  that will be part of the integrand in the integral formula for  $K^{-1}$ . When  $\mathbf{x} \in W(G) \sqcup F(G) \sqcup B(G)$ , define  $g_{\mathbf{x}}$  as follows:

1. Set  $g_{f_0}(\zeta) \equiv 1$ .
2. For a white vertex  $\mathbf{w}$  incident to a face  $\mathbf{f}$  such that they are separated by a strand  $\alpha$ , we have

$$g_{\mathbf{w}}(\zeta) = \frac{\theta(t + \zeta + \mathbf{d}(\mathbf{w}))}{E(\alpha, \zeta)} g_{\mathbf{f}}(\zeta).$$

3. For a black vertex  $\mathbf{b}$  incident to a face  $\mathbf{f}$  such that they are separated by a strand  $\alpha$ , we have

$$g_{\mathbf{b}}(\zeta) = \frac{E(\alpha, \zeta)}{\theta(-t + \zeta - \mathbf{d}(\mathbf{b}))} g_{\mathbf{f}}(\zeta).$$

More generally, for  $\mathbf{x}, \mathbf{y} \in W(G) \sqcup F(G) \sqcup B(G)$ , define

$$g_{\mathbf{x}, \mathbf{y}} := \frac{g_{\mathbf{y}}}{g_{\mathbf{x}}}. \quad (2.5)$$

In particular,  $g_{\mathbf{x}} = g_{f_0, \mathbf{x}}$ .

**Lemma 2.9** ([26, Lemma 13]; see also [82]). *For any white vertex  $\mathbf{w} \in W(G)$ , we have*

$$\sum_{\mathbf{b} \sim \mathbf{w}} K_{\mathbf{w}, \mathbf{b}} g_{\mathbf{b}, \mathbf{x}} = 0$$

for any  $\mathbf{x} \neq \mathbf{w} \in W(G) \sqcup F(G) \sqcup B(G)$ .

**Lemma 2.10** ([51, Proposition 2.10] and [26, Equation (28)]). *Let  $\mathbf{w} \in W(G)$  and  $\mathbf{b} \in B(G)$  be adjacent white and black vertices. Then*

$$K_{\mathbf{w}, \mathbf{b}} g_{\mathbf{b}, \mathbf{w}} = \omega_{\beta - \alpha} + \sum_{k=1}^g \left( \frac{\partial \log \theta}{\partial z_k}(t + \mathbf{d}(f_+)) - \frac{\partial \log \theta}{\partial z_k}(t + \mathbf{d}(f_-)) \right) \omega_k,$$

where  $\alpha$  and  $\beta$  are the strands containing the edge  $\mathbf{bw}$ , and  $f_+$  and  $f_-$  are the adjacent faces as in Figure 8.

We now describe the structure of the zeroes and poles of  $g_{x,y}$  along  $A_0$ .

**Lemma 2.11** ([26, Section 3.4]). *Let  $x, y \in W(\mathbf{G}) \sqcup F(\mathbf{G}) \sqcup B(\mathbf{G})$ . Then*

$$\operatorname{div}(g_{x,y})|_{A_0} = \mathbf{d}(y) - \mathbf{d}(x).$$

**Lemma 2.12.** *Let  $x, y \in W(\mathbf{G}) \sqcup F(\mathbf{G}) \sqcup B(\mathbf{G})$ . Then  $g_{x,y}$  cannot have both zeroes and poles at angles in  $\alpha_e$  for any  $e \in E(N)$ .*

*Proof.* By Lemma 2.11, we just need to check that for any  $e \in E(N)$ , there is a path  $\gamma$  from  $x$  to  $y$  whose intersections with  $\alpha \in \alpha_e$  all have the same orientation. By minimality, the strands in  $\alpha_e$  are pairwise disjoint and hence, they divide the plane into an ordered sequence of strips between consecutive strands. Among all paths from  $x$  to  $y$ , choose  $\gamma$  transverse to those strands and minimizing the number of intersections with  $\alpha_e$ . Then  $\gamma$  cannot go back and forth along strips, so it crosses every strand  $\alpha \in \alpha_e$  between  $x$  and  $y$  with the same orientation exactly once.  $\square$

**Lemma 2.13** ([25, Lemma 20]; see also [26, Lemma 34]). *Let  $x, y \in W(\mathbf{G}) \sqcup F(\mathbf{G}) \sqcup B(\mathbf{G})$ . As we move cyclically along  $\partial N$ , the edges corresponding to angles at which  $g_{x,y}$  has zeros and the edges corresponding to angles at which it has poles do not interlace.*

**Definition 2.14.** Let  $\mathbf{b} \in B(\mathbf{G})$  and  $\mathbf{w} \in W(\mathbf{G})$ . We define the *null sector*

$$\mathfrak{N}_{\mathbf{b},\mathbf{w}} \subset \partial N$$

as follows.

1. Suppose  $g_{\mathbf{b},\mathbf{w}}$  has at least one zero in  $A_0$ . By Lemma 2.13, the zeros and poles of  $g_{\mathbf{b},\mathbf{w}}$  along  $\partial N$  do not interlace. We then define  $\mathfrak{N}_{\mathbf{b},\mathbf{w}}$  to be the maximal closed connected interval of  $\partial N$  which contains every edge  $e$  such that  $g_{\mathbf{b},\mathbf{w}}$  has a zero at some angle  $\alpha \in \alpha_e$ , and contains no edge  $e$  such that  $g_{\mathbf{b},\mathbf{w}}$  has a pole at some angle  $\alpha \in \alpha_e$ .
2. Suppose  $g_{\mathbf{b},\mathbf{w}}$  has no zeros in  $A_0$ . Then

$$\operatorname{div}(g_{\mathbf{b},\mathbf{w}})|_{A_0} = -\alpha - \beta$$

for some angles  $\alpha, \beta \in \alpha$ . Then there are two complementary cyclic intervals in  $\partial N \setminus \{e(\alpha), e(\beta)\}$ . In this case, we define  $\mathfrak{N}_{\mathbf{b},\mathbf{w}}$  by “continuity in  $(\mathbf{b}, \mathbf{w})$ .” Precisely, by contraction moves as in Figure 7(right), we may assume that at least one of  $\mathbf{b}$  and  $\mathbf{w}$  has degree at least three. We then extend a path from  $\mathbf{b}$  to  $\mathbf{w}$  into an adjacent face of one of these vertices so as to cross a strand  $\gamma \neq \alpha, \beta$ . We define  $\mathfrak{N}_{\mathbf{b},\mathbf{w}}$  to be the interval in  $\partial N \setminus \{e(\alpha), e(\beta)\}$  containing  $e(\gamma)$ .

#### 2.4.5 Whole-plane inverse Kasteleyn matrices

We recall the construction of a two-parameter family of inverse Kasteleyn matrices for graph  $\mathbf{G}$  from [26]. These whole-plane inverse Kasteleyn matrices will be a basic ingredient in the definition of inverse Kasteleyn matrices of finite subgraphs as well as in the description of local statistics.

Since the angle function is compatible with the partial cyclic order on  $\alpha$ , there is an open interval in  $A_0 \setminus \alpha$  corresponding to each  $v = e_- \cap e_+ \in V(N)$ , namely the interval between the angles in  $\alpha_{e_-}$  and those in  $\alpha_{e_+}$ .

**Definition 2.15.** For  $v \in V(N)$  and  $\zeta \in \Sigma_+^\circ$ , let  $C_v^\zeta$  be a curve from  $\sigma(\zeta)$  to  $\zeta$  intersecting  $A_0$  exactly once in the component of  $A_0 \setminus \alpha$  corresponding to  $v$ , and such that  $\sigma_*(C_v^\zeta) = -C_v^\zeta$ . If  $\zeta \in \Sigma(\mathbb{R}) \setminus \alpha$ , let  $C_v^\zeta$  denote the natural limit as  $\zeta$  approaches  $\Sigma(\mathbb{R})$ .

Given  $\mathbf{b} \in B(\mathbb{G})$ ,  $\mathbf{w} \in W(\mathbb{G})$  and  $\zeta \in \Sigma_+ \setminus \alpha$ , define

$$C_{\mathbf{b},\mathbf{w}}^\zeta := C_v^\zeta,$$

where  $v$  is any vertex in the null sector  $\mathfrak{N}_{\mathbf{b},\mathbf{w}}$ .

**Theorem 2.16** ([26, Theorem 40]). *For any  $\zeta \in \Sigma_+ \setminus \alpha$ , the matrix*

$$A_{\mathbf{b},\mathbf{w}}^\zeta := \frac{1}{2\pi i} \int_{C_{\mathbf{b},\mathbf{w}}^\zeta} g_{\mathbf{b},\mathbf{w}}$$

*is an inverse of  $K$  in the sense that  $KA^\zeta = A^\zeta K = I$ .*

We now introduce some additional notation for the case  $\zeta \in \Sigma(\mathbb{R}) \setminus \alpha$  which will be used in Section 4.

**Definition 2.17.** Given an oriented interval  $I \subset \partial N$ , let  $C(I) \subset \Sigma$  be a curve which is the union of small counterclockwise-oriented circles that enclose all the angles  $\alpha_e$  for  $e \subset I$ .

If  $\zeta \in \Sigma(\mathbb{R}) \setminus \alpha$ , then  $A_{\mathbf{b},\mathbf{w}}^\zeta$  depends only on the connected component of  $\Sigma(\mathbb{R}) \setminus \alpha$  containing  $\zeta$ . If this component corresponds to  $u \in V(N)$ , then we can take

$$C_{\mathbf{b},\mathbf{w}}^\zeta = C(I_{\mathbf{b},\mathbf{w}}^u), \quad I_{\mathbf{b},\mathbf{w}}^u := [u, v],$$

where  $v$  is any vertex in  $\mathfrak{N}_{\mathbf{b},\mathbf{w}}$ , and we write  $A_{\mathbf{b},\mathbf{w}}^u := A_{\mathbf{b},\mathbf{w}}^\zeta$ .

#### 2.4.6 Gibbs measures

In this section we assume that the angle function  $\nu : \alpha \rightarrow A_0$  is *periodic*, i.e.

$$\nu(\alpha + (i, j)) = \nu(\alpha) \quad \text{for all } \alpha \in \alpha \text{ and } (i, j) \in \mathbb{Z}^2.$$

Equivalently, we have a well-defined angle function

$$\nu : \alpha_{\mathbb{T}} \rightarrow A_0$$

on the quotient. As explained in [26, Section 4.2], this setting is strictly more general than assuming that the weights are periodic; see [26, Proposition 43] for a characterization of the subset of periodic weights.

**Theorem 2.18** ([26, Theorem 62]). *For any  $\zeta \in \Sigma_+ \setminus \alpha$ , the operator  $A^\zeta$  defines a Gibbs measure  $\mathbb{P}^\zeta$  on dimer covers of  $\mathbb{G}$ . For any fixed collection of edges  $\{\mathbf{e}_i = \mathbf{b}_i \mathbf{w}_i\}_{i=1}^k \subset E(\mathbb{G})$ , the probability that all of the edges  $\mathbf{e}_1, \dots, \mathbf{e}_k$  are present is*

$$\mathbb{P}^\zeta(\mathbf{e}_1, \dots, \mathbf{e}_k \in M) = \prod_{i=1}^k K_{\mathbf{w}_i, \mathbf{b}_i} \det(A_{\mathbf{b}_i, \mathbf{w}_j}^\zeta)_{1 \leq i, j \leq k}.$$

The *slope*  $(s^\zeta, t^\zeta)$  of the Gibbs measure  $\mathbb{P}^\zeta$  is defined as the expected horizontal and vertical height change:

$$s^\zeta := \mathbb{E}^\zeta [h_M(\mathbf{f}_0 + (1, 0)) - h_M(\mathbf{f}_0)], \quad t^\zeta := \mathbb{E}^\zeta [h_M(\mathbf{f}_0 + (0, 1)) - h_M(\mathbf{f}_0)],$$

where  $\mathbb{E}^\zeta$  denotes expectation with respect to the probability measure  $\mathbb{P}^\zeta$ .

For  $u \in V(N)$ , let  $\zeta \in A_0 \setminus \alpha$  be any point in the connected component corresponding to  $u$ . Since  $\mathbb{P}^\zeta$  depends only on this component, we set

$$\mathbb{P}^u := \mathbb{P}^\zeta, \quad \mathbb{E}^u := \mathbb{E}^\zeta, \quad (s^u, t^u) := (s^\zeta, t^\zeta).$$

By [26, Remark 53], the measure  $\mathbb{P}^u$  is deterministic: more precisely,

$$\mathbb{P}^u(\mathbf{e} \in M) = \mathbf{1}_{\mathbf{e} \in M_u},$$

where  $M_u$  is the dimer cover consisting of all edges  $\mathbf{e}$  such that  $e(\alpha) < u < e(\beta)$  along  $\partial N$ , with  $\alpha, \beta$  the two strands containing  $\mathbf{e}$  as in Figure 8; we call  $M_u$  the *extremal dimer cover* corresponding to  $u$ .

Recall from Section 2.1 that  $|e|_{\mathbb{Z}}$  is the lattice length of  $e$  and  $(a_e, b_e) = \vec{e}/|e|_{\mathbb{Z}} \in \mathbb{Z}^2$ . Let  $u_0 \in V(N)$  and let  $M_0 := M_{u_0}$  be the reference dimer cover used to define the height function. Define

$$z(\zeta) := \prod_{\alpha \in \alpha_{\mathbb{T}}} E(\alpha, \zeta)^{-b_{e(\alpha)}}, \quad w(\zeta) := \prod_{\alpha \in \alpha_{\mathbb{T}}} E(\alpha, \zeta)^{a_{e(\alpha)}}. \quad (2.6)$$

Then

$$(s^\zeta, t^\zeta) = \left( -\frac{1}{2\pi i} \int_{C_{u_0}^\zeta} d \log z, -\frac{1}{2\pi i} \int_{C_{u_0}^\zeta} d \log w \right); \quad (2.7)$$

see [26, Equation 32]. The set of possible slopes is the Newton polygon  $N$  rotated by  $-\pi/2$ ; see [26, Corollary 56] and [84, Section 3.1.4].

### 3 Astroidal zig-zag graphs

In this section, we introduce a class of finite graphs that we will work with and study their basic properties.

#### 3.1 Definition of astroidal zig-zag graphs

Let  $\beta = (\beta_e)_{e \in E(N)}$  be a collection of strands such that  $\beta_e \in \alpha_e$ . For each vertex  $v = e_- \cap e_+ \in V(N)$ , choose

$$\mathbf{e}_v \in \beta_{e_-} \cap \beta_{e_+} \in V(\mathbf{G}^\times) \cong E(\mathbf{G}),$$

where  $e_-$  (resp.  $e_+$ ) is the edge immediately clockwise (resp. counterclockwise) from  $v$ .

We first define the boundary cycle in the medial graph (see Figure 9). Let  $\partial \mathbf{G}_\beta^\times$  be the counterclockwise-oriented cycle in  $\mathbf{G}^\times$  obtained as follows: it visits the medial vertices  $\mathbf{e}_v$  in the clockwise cyclic order of  $v \in V(N)$ , and for each edge  $e = [v_-, v_+] \in E(N)$  it follows the strand  $\beta_e$  from  $\mathbf{e}_{v_+}$  to  $\mathbf{e}_{v_-}$ . We assume throughout that  $\partial \mathbf{G}_\beta^\times$  is a simple cycle.

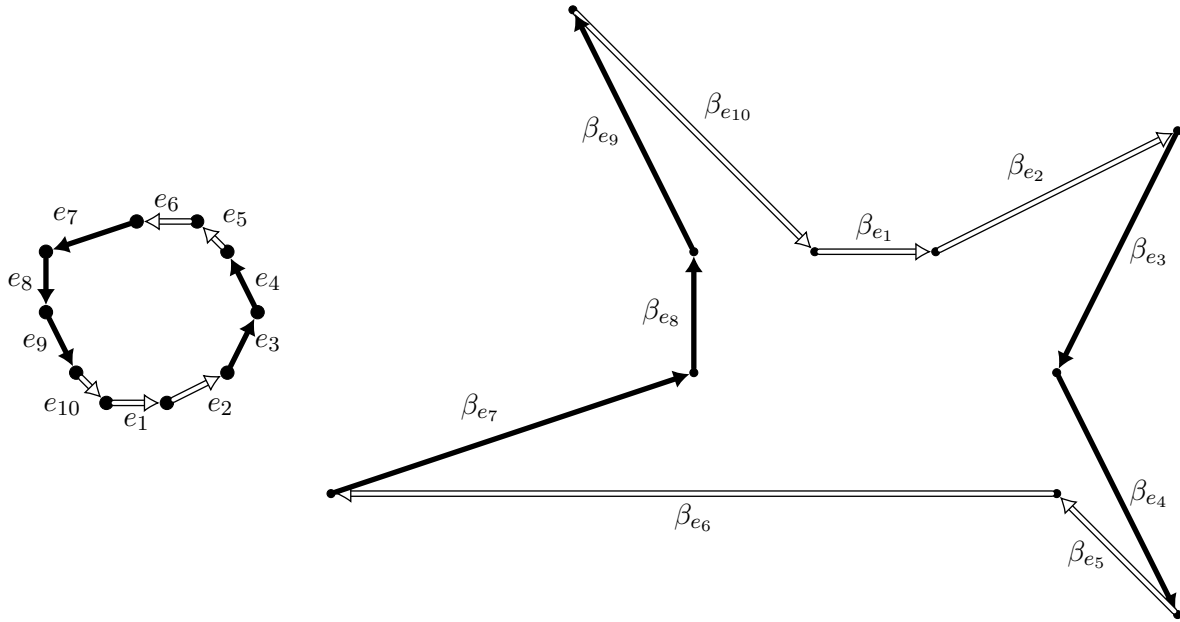


Figure 9: A Newton polygon  $N$  and an illustration of  $\partial G_{\beta}^{\times}$ , along with the edge coloring. Note that the actual  $\beta_{e_j}$ 's are (pieces of) strands, but they have the same asymptotic direction as the vectors pictured on the right.



Figure 10: The left panel shows an example illustrating the necessity of Condition 2 in Definition 3.2. The white arrows are a portion of  $\partial G_{\beta}^{\times}$ , the thick edges are the corresponding portion of  $\partial G_{\beta}$  and the shaded region consists of faces of  $G_{\beta}$ . The white vertex  $w$  on the left-hand side is visited twice by non-consecutive zig-zag paths so  $\partial G_{\beta}$  has a self-intersection. The right panel shows the neighborhood of an outer white vertex that is allowed.

From this we obtain a subgraph  $\partial\mathbf{G}_\beta$  of  $\mathbf{G}$  by defining

$$\partial\mathbf{G}_\beta := \bigcup_{e \in V(\partial\mathbf{G}_\beta^\times)} e \subset \mathbf{G},$$

*i.e.* the union of the corresponding zig-zag path segments.

Even if the medial cycle  $\partial\mathbf{G}_\beta^\times$  is simple, the resulting subgraph  $\partial\mathbf{G}_\beta$  need not be: a vertex of  $\mathbf{G}$  may lie on several strands  $\beta_e$ . This is harmless provided such occurrences are consecutive along  $\partial\mathbf{G}_\beta^\times$ . The problematic situation is when two visits to a vertex are separated by other parts of the boundary, which would produce a genuine self-intersection in  $\mathbf{G}$  (see Figure 10(left)). Throughout, we assume this does not happen; *i.e.* we assume that whenever a vertex of  $\mathbf{G}$  is visited more than once by  $\partial\mathbf{G}_\beta$ , the corresponding strands appear consecutively in the cyclic order of  $\partial\mathbf{G}_\beta^\times$ .

We then define the graph  $\mathbf{G}_\beta$  to be the finite portion of  $\mathbf{G}$  contained weakly inside  $\partial\mathbf{G}_\beta$ . The reader may find it helpful to look ahead at the examples in Section 3.2 before proceeding with the definitions below.

We next assign a color to each edge  $e \in E(N)$ . We say that  $e$  is *black* if its induced orientation agrees with the counterclockwise orientation of  $\partial\mathbf{G}_\beta^\times$ , and *white* otherwise (see Figure 9). We also define a sign function

$$\text{sign}_\beta : E(N) \rightarrow \{+, -\}, \quad \text{sign}_\beta(e) := \begin{cases} +1, & \text{if } e \text{ is black,} \\ -1, & \text{if } e \text{ is white.} \end{cases} \quad (3.1)$$

We will use the terminology “black/white” and “sign  $\pm 1$ ” interchangeably.

For each edge  $e \in E(N)$ , the portion of the strand  $\beta_e$  between  $\mathbf{e}_{v_-}$  and  $\mathbf{e}_{v_+}$  separates the vertices of the corresponding zig-zag path into two sides. We call the vertices on the exterior of  $\partial\mathbf{G}_\beta^\times$   $\beta_e$ -*outer* vertices. A vertex of  $\mathbf{G}_\beta$  that is not  $\beta_e$ -outer for any  $e \in E(N)$  is called an *inner* vertex. Note that a vertex can be  $\beta_e$ -outer for several  $e \in E(N)$ . By Lemma 2.4, such edges all have the same color and form a cyclic interval in  $\partial N$  (see Figure 10(right)).

Informally, we think of  $\partial\mathbf{G}_\beta$  as defining a polygonal region whose “vertices” are the edges  $\mathbf{e}_v$  and whose “sides” run along the zig-zag paths corresponding to  $\beta_e$ . An edge  $e$  has the color of the corresponding  $\beta_e$ -outer vertices.

**Definition 3.1.** Given such a  $\beta$ , define an element  $\mathbf{D}_\beta \in \mathbb{Z}^\alpha$  by

$$(\mathbf{D}_\beta)_{\alpha_e} := -\mathbf{d}(\mathbf{x})_{\alpha_e} \quad (3.2)$$

for all  $e \in E(N)$ , where  $\mathbf{x}$  is any  $\beta_e$ -outer vertex. We say that  $\mathbf{G}_\beta$  is *admissible* if

$$\deg(\mathbf{D}_\beta) = 0. \quad (3.3)$$

Since

$$\deg(\mathbf{D}_\beta) = \sum_{e \in E(N)} \deg((\mathbf{D}_\beta)_{\alpha_e}),$$

the admissibility condition (3.3) says that the total number of strands between the reference face and  $\beta_e$ , summed over white edges  $e$ , is equal to the corresponding total summed over black edges.

The role of admissibility will be clarified in Lemma 4.1.

**Definition 3.2.** We call  $\mathbf{G}_\beta$  an *astroidal zig-zag graph* (AZ graph) if:

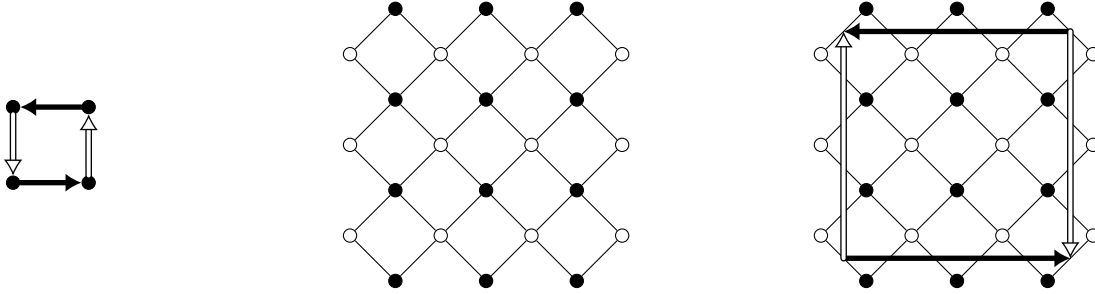


Figure 11: A square Newton polygon  $N$  (left), the Aztec diamond  $G_\beta$  (middle) and its medial cycle  $\partial G_\beta^\times$  (right).

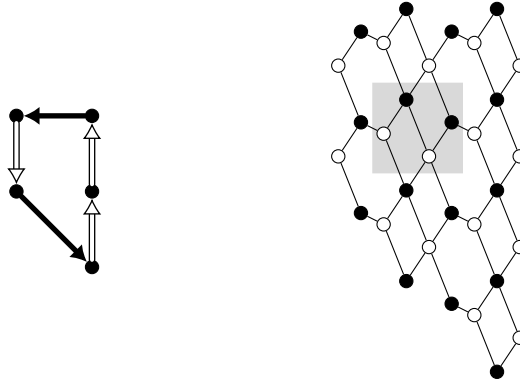


Figure 12: A quadrilateral Newton polygon and an example of a corresponding AZ graph.

1. the medial cycle  $\partial G_\beta^\times$  is simple,
2. whenever a vertex of  $G$  is visited more than once as we go around  $\partial G_\beta^\times$ , all such visits occur consecutively along  $\partial G_\beta^\times$ ,
3.  $G_\beta$  is admissible.

The reason for the name is that the boundary  $\partial G_\beta$  resembles an astroid curve (see Lemma 3.13) and is made up of zig-zag paths. Since  $G$  is  $\mathbb{Z}^2$ -periodic and we have one constraint (3.3), AZ graphs form a  $(\#E(N) - 3)$ -parameter family. In concrete examples, AZ graphs are easy to construct directly. A general construction showing that the class is nonempty for every Newton polygon with at least four sides will be given in Remark 5.2.

### 3.2 Examples of AZ graphs

**Example 3.3.** When  $N$  is any triangle, the  $G$  is the hexagonal lattice [68, Proposition 11.3], we have a 0-parameter family, and there are no AZ graphs.

**Example 3.4.** When  $N$  is a convex lattice quadrilateral, there is a one-parameter family of AZ graphs. If  $N$  is the unit square, one choice of minimal graph  $G$  is the square lattice. A general

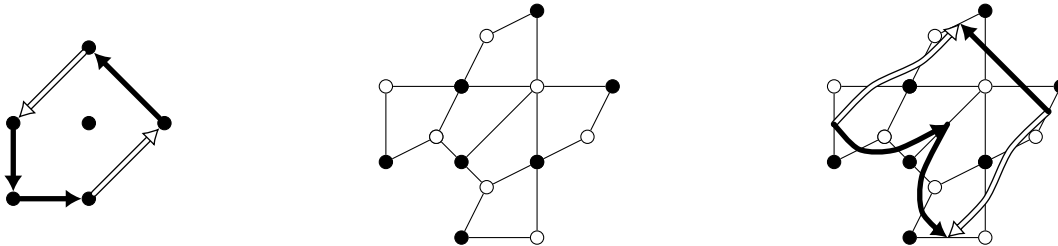


Figure 13: An example of an AZ graph for the pentagonal Newton polygon in Figure 1 which we will use as our running example.

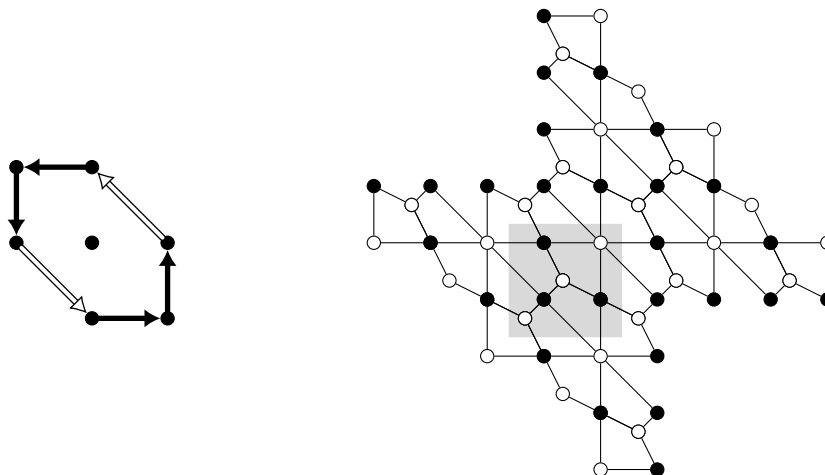


Figure 14: A hexagonal Newton polygon and an example of an AZ graph.

choice of  $\beta$  produces a rectangular subgraph, and admissibility forces this subgraph to be the Aztec diamond (Figure 11). Any other choice of minimal graph is related to the square lattice by local moves and has the same large-scale statistical-mechanical behavior. Thus, up to local moves, the Aztec diamonds are the only AZ graphs associated with the unit square Newton polygon. Figure 12 shows another quadrilateral Newton polygon together with an associated AZ graph, called a *tower graph*, which is the unique AZ graph up to size, and it has been studied in [21, 17, 92] and in the physics literature as the brane tiling/dimer model for the suspended pinch point singularity; see for example, [54, 55, 77, 49]. Another example is given by the *pinecone graphs* associated with Gale–Robinson sequences studied in [23, 70, 32].

**Example 3.5.** Figures 4 and 13 show two examples of AZ graphs for the pentagonal Newton polygon in Figure 1. For this choice of  $G$ , the space of AZ graphs forms a two-parameter family, which was studied from the point of view of cluster algebras in [57].

**Example 3.6.** Let  $N$  be the hexagon shown in Figure 14(left). Figure 14(right) shows a choice of fundamental domain for  $G_{\mathbb{T}}$  and an example of an AZ graph. The class of AZ graphs for this  $G$  has been studied extensively from the cluster algebra point of view in [40, 111, 90, 88, 37].

**Example 3.7.** Figure 15 shows an example of an AZ graph for an octagonal Newton polygon. The corresponding torus graph was studied in [7].

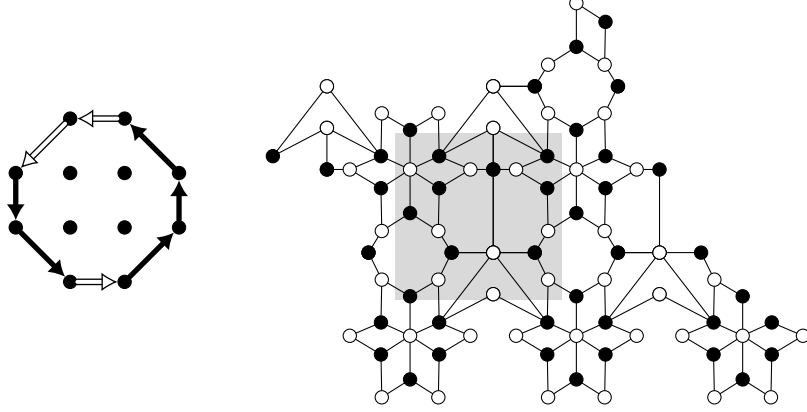


Figure 15: An octagonal Newton polygon  $N$  and an AZ graph.

### 3.3 Criterion for admissibility

For any  $e \in E(N)$ , we label the strands  $\alpha_e = \{\alpha_e^i : i \in \mathbb{Z} + \frac{1}{2}\}$  parallel to  $e$  from left to right so that the reference face  $f_0$  lies in between  $\alpha_e^{-\frac{1}{2}}$  and  $\alpha_e^{\frac{1}{2}}$ . Define

$$c_\beta : E(N) \rightarrow \mathbb{Z} + \frac{1}{2}$$

by  $\alpha_e^{c_\beta(e)} = \beta_e$ .

**Lemma 3.8.** *For each  $e \in E(N)$ ,*

$$\deg((\mathbf{D}_\beta)_{\alpha_e}) = -c_\beta(e) + \begin{cases} -\frac{1}{2}, & \text{if } e \text{ is black,} \\ \frac{1}{2}, & \text{if } e \text{ is white.} \end{cases}$$

*In particular,  $\mathcal{G}_\beta$  is admissible if and only if*

$$\sum_{e \in E(N)} c_\beta(e) = \frac{1}{2} \left( \#\{e \in E(N) : e \text{ is white}\} - \#\{e \in E(N) : e \text{ is black}\} \right). \quad (3.4)$$

*Proof.* We have, for a  $\beta_e$ -outer vertex  $\mathbf{b}$ ,

$$\mathbf{d}(\mathbf{b})_{\alpha_e} = \begin{cases} -\alpha_e^{c_\beta(e)+1} - \dots - \alpha_e^{-\frac{1}{2}}, & \text{if } c_\beta(e) < 0, \\ \alpha_e^{\frac{1}{2}} + \dots + \alpha_e^{c_\beta(e)}, & \text{if } c_\beta(e) > 0, \end{cases}$$

if  $e$  is black, so  $\deg((\mathbf{D}_\beta)_{\alpha_e}) = -c_\beta(e) - \frac{1}{2}$ . Similarly, for a  $\beta_e$ -outer vertex  $\mathbf{w}$ ,

$$\mathbf{d}(\mathbf{w})_{\alpha_e} = \begin{cases} -\alpha_e^{c_\beta(e)} - \dots - \alpha_e^{-\frac{1}{2}}, & \text{if } c_\beta(e) < 0, \\ \alpha_e^{\frac{1}{2}} + \dots + \alpha_e^{c_\beta(e)-1}, & \text{if } c_\beta(e) > 0, \end{cases}$$

if  $e$  is white, so  $\deg((\mathbf{D}_\beta)_{\alpha_e}) = -c_\beta(e) + \frac{1}{2}$ . Summing over all edges gives (3.4).  $\square$

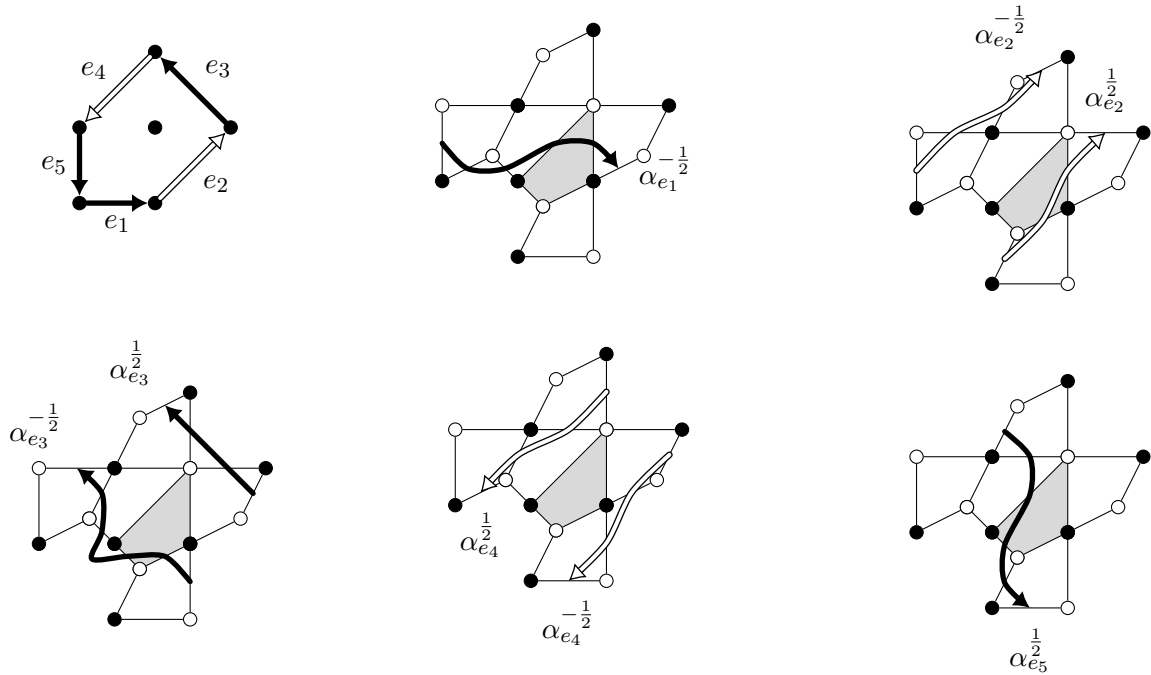


Figure 16: Classes of parallel strands for the graph in Figure 13. The shaded face is the reference face  $f_0$ .

**Example 3.9.** Let us verify that the graph in Figure 13 is admissible. The strands are shown in Figure 16, from which we read

$$c_\beta(e_1) = -\frac{1}{2}, \quad c_\beta(e_2) = -\frac{1}{2}, \quad c_\beta(e_3) = \frac{1}{2}, \quad c_\beta(e_4) = -\frac{1}{2}, \quad c_\beta(e_5) = \frac{1}{2}.$$

Summing gives

$$\sum_{i=1}^5 c_\beta(e_i) = -\frac{1}{2} = \frac{1}{2}(2-3),$$

so  $G_\beta$  is admissible by Lemma 3.8. Alternatively, from Figure 13 and Figure 16 we see that

$$D_\beta = \alpha_{e_2}^{-\frac{1}{2}} - \alpha_{e_3}^{\frac{1}{2}} + \alpha_{e_4}^{-\frac{1}{2}} - \alpha_{e_5}^{\frac{1}{2}}, \quad (3.5)$$

hence  $\deg(D_\beta) = 0$ , which again shows admissibility.

### 3.4 Characterization of AZ graphs

In this section we prove that, among graphs whose boundary consists of zig-zag paths in the prescribed order, the only ones that admit dimer covers are the AZ graphs. The converse, namely that every AZ graph admits a dimer cover, is proved later in Corollary 4.4.

Let  $\partial G_\beta^\times$ ,  $\partial G_\beta$ , and  $G_\beta$  be as in Section 3.1. We say that a face  $f \in F(G) \setminus F(G_\beta)$  is *outer* if it is adjacent to a vertex of  $G_\beta$ , and  $\beta_e$ -*outer* for  $e \in E(N)$  if it is adjacent to a vertex in the zig-zag path corresponding to the strand  $\beta_e$ .

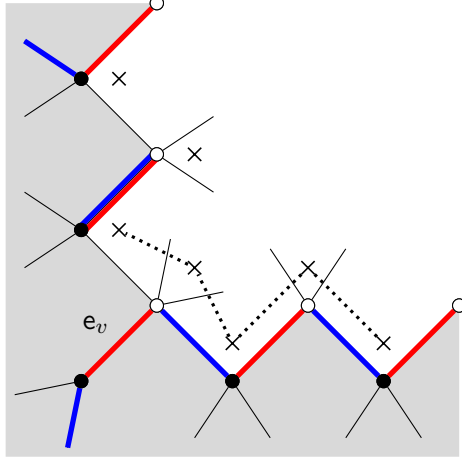


Figure 17: The faces marked by  $\times$  are the  $\beta_{e_-}$ - and  $\beta_{e_+}$ -outer faces associated with the vertex  $v = e_- \cap e_+ \in V(N)$ . The shaded region indicates  $\mathbf{G}_\beta$ . The blue edges form a dimer cover  $M$  of  $\mathbf{G}_\beta$ , while the red edges form the extremal dimer cover  $M_v$ . Any dual path between two marked outer faces, such as the dotted path shown, crosses no edges of either  $M$  or  $M_v$ . Hence, the difference  $h_{M_v}(\mathbf{f}) - \partial h(\mathbf{f})$  in (3.6) is constant on these outer faces.

**Proposition 3.10.** *Assume that Items 1 and 2 in Definition 3.2 hold. If  $\mathbf{G}_\beta$  admits a dimer cover, then it is an AZ graph.*

*Proof.* We prove that if  $\mathbf{G}_\beta$  admits a dimer cover, then Item 3 in Definition 3.2, *i.e.* the admissibility condition, is satisfied.

Fix a vertex  $u_0 \in V(N)$  and let  $M_0 := M_{u_0}$  denote the extremal dimer cover from Section 2.4.6. Let  $M$  be a dimer cover of  $\mathbf{G}_\beta$  and let  $h_M$  be the associated height function on  $F(\mathbf{G}_\beta)$  with reference dimer cover  $M_0$  and reference face  $\mathbf{f}_0$ . The union of two dimer covers  $M$  and  $M'$  of  $\mathbf{G}_\beta$  forms a collection of loops. The nontrivial loops can be interpreted as level lines of the height difference  $h_M - h_{M'}$ . In particular, this difference vanishes on the outer faces, so  $h_M$  and  $h_{M'}$  induce the same height function on the outer faces, *i.e.* the boundary values of  $h_M$  are independent of  $M$ . We denote this common boundary height function by  $\partial h$ .

Recall that for  $v \in V(N)$ , the extremal dimer cover  $M_v$  is, by definition, the dimer cover consisting of all edges  $e$  such that  $e(\alpha) < v < e(\beta)$  along  $\partial N$ , where  $\alpha$  and  $\beta$  are the two strands passing through  $e$  as in Figure 8. In particular, if  $e \in E(N)$  is an edge containing  $v$ , then every second edge of  $\beta_e$  is contained in  $M_v$ . Hence,

$$h_{M_v}(\mathbf{f}) - \partial h(\mathbf{f}) \tag{3.6}$$

is constant for all  $\beta_{e_-}$ - and  $\beta_{e_+}$ -outer faces  $\mathbf{f}$ , where  $v = e_- \cap e_+$ ; see Figure 17. We denote this constant value by  $d_v$ .

Now let  $\mathbf{f}$  be any  $\beta_e$ -outer face for  $e = [v_-, v_+]$ . Then

$$d_{v_+} - d_{v_-} = \sum_{e \in E(\mathbf{G})} (e \wedge \gamma^*) \mathbb{E}[\mathbf{1}_{e \in M_{v_+}} - \mathbf{1}_{e \in M_{v_-}}], \tag{3.7}$$

where  $\gamma^*$  is any dual path from the reference face  $\mathbf{f}_0$  to  $\mathbf{f}$ . The right-hand side of (3.7) can be

expressed in terms of the whole-plane inverse Kasteleyn matrices from Section 2.4.5 as

$$\begin{aligned} \sum_{\mathbf{e} \in E(\mathbf{G})} (\mathbf{e} \wedge \gamma^*) \mathbb{E}[\mathbf{1}_{\mathbf{e} \in M_{v_+}} - \mathbf{1}_{\mathbf{e} \in M_{v_-}}] &= \sum_{\mathbf{e} = \mathbf{bw} \in E(\mathbf{G})} (\mathbf{e} \wedge \gamma^*) \left( A_{\mathbf{b},\mathbf{w}}^{v_+} K_{\mathbf{w},\mathbf{b}} - A_{\mathbf{b},\mathbf{w}}^{v_-} K_{\mathbf{w},\mathbf{b}} \right) \\ &= - \sum_{\mathbf{e} \in E(\mathbf{G})} (\mathbf{e} \wedge \gamma^*) \frac{1}{2\pi i} \int_{C([v_-, v_+])} g_{\mathbf{b},\mathbf{w}} \cdot K_{\mathbf{w},\mathbf{b}}. \end{aligned}$$

If  $\mathbf{e} = \mathbf{bw} \in E(\mathbf{G})$ , and  $\mathbf{f}_+, \mathbf{f}_- \in F(\mathbf{G})$  are adjacent to  $\mathbf{e}$  as in Figure 8, then by Lemma 2.10,

$$\frac{1}{2\pi i} \int_{C([v_+, v_-])} g_{\mathbf{b},\mathbf{w}} \cdot K_{\mathbf{w},\mathbf{b}} = \frac{1}{2\pi i} \int_{C([v_+, v_-])} \omega_{\mathbf{d}(\mathbf{f}_+) - \mathbf{d}(\mathbf{f}_-)} = \deg((\mathbf{d}(\mathbf{f}_+) - \mathbf{d}(\mathbf{f}_-))_{\alpha_e}).$$

Summing over the edges crossing  $\gamma^*$  yields

$$d_{v_+} - d_{v_-} = -\deg(\mathbf{d}(\mathbf{f})_{\alpha_e}). \quad (3.8)$$

Finally, summing (3.8) over all  $e = [v_-, v_+] \in E(N)$ , and using the definition of  $\mathbf{D}_\beta$  in (3.2), we obtain

$$0 = \sum_{e = [v_-, v_+] \in E(N)} (d_{v_+} - d_{v_-}) = \deg(\mathbf{D}_\beta),$$

which is the admissibility condition.  $\square$

**Remark 3.11.** The proof of Proposition 3.10 shows more generally that, if Items 1 and 2 of Definition 3.2 hold, then  $\mathbf{G}_\beta$  is an AZ graph whenever the boundary height function is single-valued on the  $\beta_e$ -outer faces; equivalently, the height increments along the boundary close up.

### 3.5 Convex and concave vertices

**Definition 3.12.** We call a vertex  $v = e_- \cap e_+ \in V(N)$  *convex* if  $e_-$  and  $e_+$  have opposite colors (equivalently,  $\text{sign}_\beta(e_-) \neq \text{sign}_\beta(e_+)$ ), and *concave* otherwise.

**Lemma 3.13.** *There are exactly four convex vertices in  $V(N)$ .*

*Proof.* Consider the ‘‘Gauss map’’

$$\gamma : V(\partial \mathbf{G}_\beta^\times) \setminus \{\mathbf{e}_v\}_{v \in V(N)} \rightarrow \mathbb{R}/2\pi\mathbb{Z}, \quad \gamma(\mathbf{e}) := \arg(\text{sign}_\beta(e)\vec{e}) \quad \text{if } \mathbf{e} \in V(\beta_e),$$

sending a medial vertex  $\mathbf{e} \in V(\partial \mathbf{G}_\beta^\times)$  to the oriented ‘‘tangent direction’’ when  $\partial \mathbf{G}_\beta^\times$  is oriented counterclockwise. At  $v = e_- \cap e_+$ , the map  $\gamma$  changes by (see Figure 18)

$$\pi - (\arg(\vec{e}_+) - \arg(\vec{e}_-)) \quad \text{if } v \text{ is convex}, \quad -(\arg(\vec{e}_+) - \arg(\vec{e}_-)) \quad \text{if } v \text{ is concave}.$$

Since  $\partial \mathbf{G}_\beta^\times$  is simple, the total change of  $\gamma$  along  $\partial \mathbf{G}_\beta^\times$  is  $2\pi$ . Therefore,

$$\pi \cdot \#\{\text{convex vertices}\} - \sum_{v = e_- \cap e_+ \in V(N)} (\arg(\vec{e}_+) - \arg(\vec{e}_-)) = 2\pi.$$

On the other hand, since  $\partial N$  is a simple closed curve,

$$\sum_{v = e_- \cap e_+ \in V(N)} (\arg(\vec{e}_+) - \arg(\vec{e}_-)) = 2\pi.$$

Substituting gives  $\pi \cdot \#\{\text{convex vertices}\} = 4\pi$ , hence, there are exactly four convex vertices.  $\square$

Thus, the convex vertices partition  $\partial N$  and  $\partial \mathbf{G}_\beta^\times$  into four intervals that alternate in color (see Figure 9).

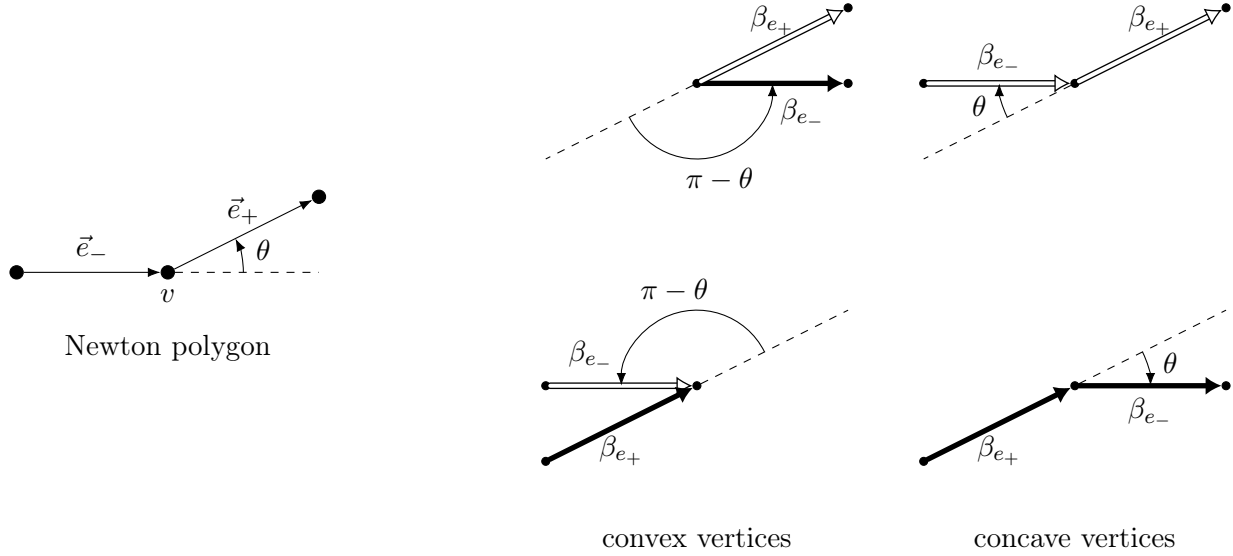


Figure 18: Change of Gauss map  $\gamma$  at a vertex  $v = e_- \cap e_+$ , where  $\theta = \arg(\vec{e}_+) - \arg(\vec{e}_-)$ .

### 3.6 Chambers

Unlike the case of the Aztec diamond, the relative position of vertices with respect to strands need not be uniform. Thus, we introduce “local sign functions” in addition to the global sign function  $\text{sign}_\beta$  from (3.1).

The collection of strands  $\beta$  partitions the plane into several (open) *chambers*. Precisely, the chambers are the connected components of

$$\mathbb{R}^2 \setminus \bigcup_{e \in E(N)} \beta_e.$$

We call the chambers contained inside (resp. outside)  $\partial G_\beta^\times$  *inner* (resp. *outer*) chambers. Each chamber  $R$  corresponds to a *sign function*

$$\text{sign}_R : E(N) \rightarrow \{+, -\}, \quad \text{sign}_R(e) := \begin{cases} +, & \text{if } R \text{ lies to the left of } \beta_e, \\ -, & \text{if } R \text{ lies to the right of } \beta_e. \end{cases} \quad (3.9)$$

Recall that we identify signs with colors, and therefore,  $\text{sign}_R$  equivalently defines a coloring of  $E(N)$  for each chamber.

**Definition 3.14.** Let  $v = e_- \cap e_+ \in V(N)$  be a vertex of the Newton polygon. We say that a chamber  $R$  has a *sign (or color) change* at  $v$  if

$$\text{sign}_R(e_-) \neq \text{sign}_R(e_+).$$

We write  $\text{SC}(R) \subset V(N)$  for the set of sign changes of  $R$ .

**Lemma 3.15.** *An inner chamber has exactly four sign changes, while an outer chamber has exactly two.*

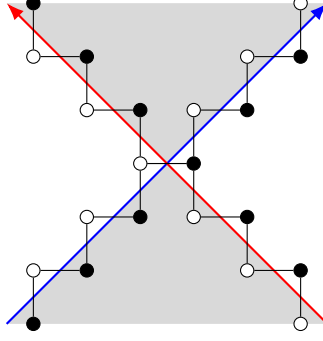


Figure 19: The four conical regions. The shaded regions are the inconsistently oriented cones.

*Proof.* The argument is identical to that of Lemma 3.13, with the Gauss map replaced by

$$\mathbf{e} \mapsto \arg(\text{sign}_{\mathbf{R}}(e)\vec{e}) \quad \text{if } \mathbf{e} \in V(\beta_e),$$

*i.e.* the tangent direction as seen from  $\mathbf{R}$  and using that the total winding is  $2\pi$  for an inner chamber and 0 for an outer chamber.  $\square$

We now describe the sign changes geometrically. Let  $v = e_- \cap e_+ \in V(N)$ . The strands  $\beta_{e_-}$  and  $\beta_{e_+}$  subdivide  $\mathbb{R}^2$  into four conical regions (Figure 19). Two of these cones have boundaries that are consistently oriented and two have boundaries that are inconsistently oriented. By definition, a chamber  $\mathbf{R}$  has a sign change at  $v$  if and only if  $\mathbf{R}$  is contained in one of the two inconsistently oriented cones determined by  $\beta_{e_-}$  and  $\beta_{e_+}$  at  $v$ .

Consider the adjacency graph  $\mathcal{R}$  of inner chambers: its vertices are the inner chambers, and two vertices are joined by an edge if and only if the boundaries of the corresponding chambers share a nontrivial segment of a strand  $\beta_e$ . In particular, chambers meeting only at a crossing, such as the two gray regions in Figure 19, are not considered adjacent.

**Lemma 3.16.** *Across each edge of  $\mathcal{R}$ , exactly one of the four sign changes moves one step clockwise or counterclockwise along  $\partial N$ .*

*Proof.* Let  $\mathbf{R}$  and  $\mathbf{R}'$  be adjacent chambers, separated by a segment of a strand  $\beta_e$ . Then  $\text{sign}_{\mathbf{R}'}(e) = -\text{sign}_{\mathbf{R}}(e)$ , while  $\text{sign}_{\mathbf{R}'}(e') = \text{sign}_{\mathbf{R}}(e')$  for all  $e' \neq e$ .

If  $e$  is not adjacent to the boundary of one of the four sign intervals for  $\text{sign}_{\mathbf{R}}$  (equivalently, if  $e$  is not adjacent along  $\partial N$  to a vertex in  $\text{SC}(\mathbf{R})$ ), then changing only the sign at  $e$  would create two additional sign changes (one at each endpoint of  $e$  along  $\partial N$ ), while the previous four sign changes persist. This would give  $\#\text{SC}(\mathbf{R}') \geq 6$ , contradicting Lemma 3.15. Hence,  $e$  must lie at the boundary of exactly one sign interval, so exactly one sign change shifts by one step along  $\partial N$ , and the other three remain fixed.  $\square$

In particular, for every edge  $\mathbf{R}\mathbf{R}'$  of  $\mathcal{R}$  there is a canonical bijection

$$\text{SC}(\mathbf{R}) \cong \text{SC}(\mathbf{R}') \tag{3.10}$$

which is the identity on three of the four elements. Consider the discrete bundle over  $\mathcal{R}$  whose fiber over an inner chamber  $\mathbf{R}$  is the four-element set  $\text{SC}(\mathbf{R})$ , equipped with the connection determined by the bijections (3.10) along edges.

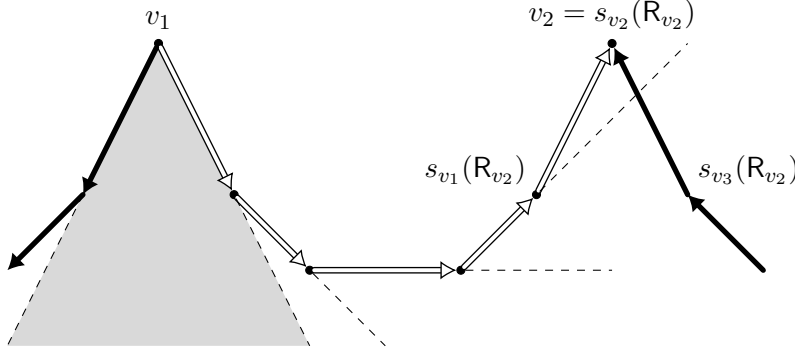


Figure 20: The chosen inconsistently oriented cones in  $[v_1, v_2)$  whose boundary orientations match the boundary orientation of the reference cone at  $v_1$  (shaded).

**Lemma 3.17.** *The connection thus defined is flat, i.e. the parallel transport around any closed loop in  $\mathcal{R}$  is the identity.*

*Proof.* Since  $\mathcal{R}$  is planar, it suffices to check triviality around each face. Each face is a 4-cycle of inner chambers surrounding an intersection point of two strands  $\beta_{e_1}$  and  $\beta_{e_2}$  in  $\mathcal{G}$ . Label the chambers cyclically by

$$\mathcal{R} \xrightarrow{\beta_{e_1}} \mathcal{R}_1 \xrightarrow{\beta_{e_2}} \mathcal{R}_{12} \xrightarrow{\beta_{e_1}} \mathcal{R}_2 \xrightarrow{\beta_{e_2}} \mathcal{R}.$$

Crossing  $\beta_{e_i}$  flips the sign at  $e_i$  and, by Lemma 3.16, moves exactly one sign change by one step along  $\partial N$ , fixing the other three. Since the intersection  $\beta_{e_1} \cap \beta_{e_2}$  is not on  $\partial \mathcal{G}_\beta$ , the edges  $e_1$  and  $e_2$  are not adjacent along  $\partial N$ , hence, the moved sign change for  $\beta_{e_1}$  is different from the moved sign change for  $\beta_{e_2}$ . Therefore, going once around the 4-cycle moves the  $e_1$ -sign change one step and then moves it back, and similarly for the  $e_2$ -sign change; all other sign changes remain fixed throughout. Thus, the composed bijection around the face is the identity, and the connection is flat.  $\square$

The discrete bundle of sign changes has four canonical sections which we now define. Each convex vertex  $v \in V(N)$  determines a section

$$s_v : \{\text{inner chambers}\} \rightarrow V(N), \quad s_v(\mathcal{R}) \in \text{SC}(\mathcal{R}),$$

as follows. Let  $\mathcal{R}_v$  be the inner chamber adjacent to  $v$ . Then  $v \in \text{SC}(\mathcal{R}_v)$ , and we set  $s_v(\mathcal{R}_v) = v$ . For any other inner chamber  $\mathcal{R}$ , define  $s_v(\mathcal{R})$  by parallel transporting  $v$  along any path in  $\mathcal{R}$  from  $\mathcal{R}_v$  to  $\mathcal{R}$ . By Lemma 3.17, this is well-defined.

**Lemma 3.18.** *For every inner chamber  $\mathcal{R}$ , the four sections associated to the four convex vertices are pairwise distinct at  $\mathcal{R}$ , together exhaust  $\text{SC}(\mathcal{R})$ , and appear in the same cyclic order along  $\partial N$  as the convex vertices.*

*Proof.* Let  $v_1, v_2, v_3, v_4$  be the convex vertices in cyclic order along  $\partial N$ . Consider any edge  $\mathcal{R}\mathcal{R}'$  of  $\mathcal{R}$ . By Lemma 3.16, under the connection bijection  $\text{SC}(\mathcal{R}) \cong \text{SC}(\mathcal{R}')$  exactly one element of  $\text{SC}(\mathcal{R})$  moves by one step along  $\partial N$ , while the other three are fixed. In particular, no two transported elements can collide or cross: a collision would force  $\#\text{SC}(\mathcal{R}') < 4$ , contradicting Lemma 3.15, and a crossing would require two elements to pass each other even though at most one moves.

Consequently, if  $s_{v_i}(\mathbf{R}) \neq s_{v_j}(\mathbf{R})$  for  $i \neq j$  at some chamber  $\mathbf{R}$ , then  $s_{v_i}$  and  $s_{v_j}$  remain distinct over all chambers, and their cyclic order along  $\partial N$  is preserved throughout  $\mathcal{R}$ .

It therefore suffices to verify the distinctness and the cyclic ordering in a single chamber. Without loss of generality, we check the cyclic order at the inner chamber  $\mathbf{R}_{v_2}$  adjacent to  $v_2$ . Consider a path in  $\mathcal{R}$  that moves clockwise along the boundary  $\partial G_\beta$  from the inner chamber  $\mathbf{R}_{v_1}$  adjacent to  $v_1$  to  $\mathbf{R}_{v_2}$ . Initially, we have  $s_{v_1}(\mathbf{R}_{v_1}) = v_1$ . Along this path, the section  $s_{v_1}$  evolves as follows. For each vertex  $u \in [v_1, v_2)$  there are two inconsistently oriented cones. Among them, choose the one whose boundary orientation agrees with that of the reference cone at  $v_1$  containing  $G_\beta$  (*i.e.*, in both cones the strands point either toward or away from the apex). Then for every chamber  $\mathbf{R}$  encountered along the path, we set  $s_{v_1}(\mathbf{R}) = u$ , where  $u$  is the unique vertex in  $[v_1, v_2)$  such that  $\mathbf{R}$  lies in this chosen cone at  $u$  (see Figure 20). In particular, when the path reaches  $\mathbf{R}_{v_2}$ , the section  $s_{v_1}$  has advanced monotonically to the vertex immediately counterclockwise of  $v_2$ . Applying the same argument with  $v_3$  in place of  $v_1$ , we get that  $s_{v_3}(\mathbf{R}_{v_2})$  is the vertex immediately clockwise of  $v_2$ . Hence, the sections  $s_{v_1}, s_{v_2}, s_{v_3}$  are distinct and appear in the correct cyclic order at  $\mathbf{R}_{v_2}$ .

Finally, since  $\#\text{SC}(\mathbf{R}) = 4$  for every chamber  $\mathbf{R}$ , the four sections exhaust  $\text{SC}(\mathbf{R})$  for all  $\mathbf{R}$ .  $\square$

## 4 Exact inverse Kasteleyn formula for AZ graphs

In this section, we define  $K^{-1}$  by a double integral plus single-integral correction term.

### 4.1 Formula for $K^{-1}$

Let  $\mathbf{b} \in B(G_\beta)$  and  $\mathbf{w} \in W(G_\beta)$ . Define the kernel

$$\omega_{\mathbf{b},\mathbf{w}}(\zeta, \eta) := \frac{1}{(2\pi i)^2} \frac{1}{E(\zeta, \eta)} \frac{\theta(-t + \mathbf{D}_\beta + \eta - \zeta)}{\theta(-t + \mathbf{D}_\beta)} \frac{g_{\mathbf{w}}(\eta) \Psi_\beta(\eta)}{g_{\mathbf{b}}(\zeta) \Psi_\beta(\zeta)}, \quad (4.1)$$

where

$$\Psi_\beta(\zeta) := \prod_{\alpha \in \alpha} E(\alpha, \zeta)^{(\mathbf{D}_\beta)_\alpha}$$

encodes the dependence on the boundary data. As in Section 2.3.1, we identify  $\mathbf{D}_\beta + \eta - \zeta$  with its image  $\mu(\mathbf{D}_\beta + \eta - \zeta)$ . The parameter  $t$  here is as in Section 2.4.3.

**Lemma 4.1.** *The kernel  $\omega_{\mathbf{b},\mathbf{w}}(\zeta, \eta)$  is a meromorphic  $(1, 1)$ -form on  $\Sigma \times \Sigma$ .*

*Proof.* This is standard; see for instance [26, Lemma 32]. Fix  $\eta$  and view  $\omega_{\mathbf{b},\mathbf{w}}(\zeta, \eta)$  as a function of  $\zeta$ . To say it is a meromorphic 1-form in  $\zeta$  is equivalent to saying that its divisor is linearly equivalent to the canonical divisor:

$$\text{div}_\zeta(\omega_{\mathbf{b},\mathbf{w}}) \sim K_\Sigma.$$

By Abel's theorem, this holds provided

$$\deg(\text{div}_\zeta(\omega_{\mathbf{b},\mathbf{w}})) = \deg(K_\Sigma) = 2g - 2, \quad \mu(\text{div}_\zeta(\omega_{\mathbf{b},\mathbf{w}})) = K_\Sigma \in \text{Jac}(\Sigma).$$

By Riemann's theorem [51, Theorem 1.1], for any  $e \in \mathbb{C}^g$  such that  $\theta(e + \zeta)$  is not identically 0,

$$\deg(\text{div}_\zeta(\theta(e + \zeta))) = g, \quad \mu(\text{div}_\zeta(\theta(e + \zeta))) = -e + \Delta,$$

where  $\Delta$  is the vector of Riemann constants, satisfying  $2\Delta = K_\Sigma$ .

As a function of  $\zeta$ ,

- $E(\zeta, \eta)^{-1}$  has a simple pole at  $\zeta = \eta$ , hence degree  $-1$ , and  $\mu(\operatorname{div}_\zeta(E(\zeta, \eta))) = \eta$  (recall that we often suppress the Abel map; the  $\eta$  on the right-hand side stands for  $\mu(\eta)$ ).
- $\theta(-t + \mathbf{D}_\beta + \eta - \zeta)$  contributes degree  $g$ , and  $\mu(\operatorname{div}_\zeta(\theta(-t + \mathbf{D}_\beta + \eta - \zeta))) = -t + \mathbf{D}_\beta + \eta + \Delta$ .
- $(g_b \Psi_\beta)^{-1}$  contributes degree  $g - 1 + \deg(\operatorname{div}(\Psi_\beta))$ , and

$$\mu(\operatorname{div}(g_b)) = -t - \Delta, \quad \mu(\operatorname{div}(\Psi_\beta)) = \mathbf{D}_\beta.$$

By admissibility,  $\deg(\operatorname{div}(\Psi_\beta)) = \deg(\mathbf{D}_\beta) = 0$ . Thus,

$$\deg(\operatorname{div}_\zeta(\omega_{b,w})) = -1 + g + (g - 1 + \deg(\operatorname{div}(\Psi_\beta))) = 2g - 2.$$

Moreover,

$$\begin{aligned} \mu(\operatorname{div}_\zeta(\omega_{b,w})) &= -\mu(\operatorname{div}_\zeta(E(\zeta, \eta))) + \mu(\operatorname{div}_\zeta(\theta(-t + \mathbf{D}_\beta + \eta - \zeta))) - \mu(\operatorname{div}_\zeta(g_b(\zeta))) - \mu(\operatorname{div}_\zeta(\Psi_\beta(\zeta))) \\ &= -\eta + (-t + \mathbf{D}_\beta + \eta + \Delta) + (t + \Delta) - \mathbf{D}_\beta \\ &= 2\Delta = K_\Sigma. \end{aligned}$$

The same argument in the  $\eta$ -variable shows that  $\omega_{b,w}(\zeta, \eta)$  is a meromorphic  $(1, 1)$ -form on  $\Sigma \times \Sigma$ .  $\square$

The residue of  $\omega_{b,w}$ , viewed as a meromorphic 1-form in  $\zeta$ , along the diagonal  $\zeta = \eta$ , is

$$\operatorname{res}_{\zeta=\eta}(\omega_{b,w}) = -\frac{1}{2\pi i} g_{b,w}.$$

For an interval  $I = [\ell, r] \subset \partial N$ , define its boundary to be the 0-chain

$$\partial I := r - \ell.$$

Recall from Section 2.4.5 the definitions of the contour  $C(I_{b,w}^u)$  and the whole-plane inverse  $A_{b,w}^u$ . By integration, we obtain a linear functional on 0-chains defined by

$$u \mapsto \langle u, A_{b,w} \rangle := A_{b,w}^u = \frac{1}{2\pi i} \int_{C(I_{b,w}^u)} g_{b,w}, \quad (4.2)$$

and extended to arbitrary 0-chains by linearity.

Let  $v$  be a convex vertex and let  $\mathbf{b} \in B(\mathbf{G}_\beta)$  and  $\mathbf{w} \in W(\mathbf{G}_\beta)$ . For  $\{x, y\} = \{\mathbf{b}, \mathbf{w}\}$ , define

$$K_{b,w}^{-1} := \operatorname{sign}(v) \left( \iint_{C(I_y) \prec C(I_x)} \omega_{b,w} - \varepsilon(x) \langle \partial I_x \cap I_y, A_{b,w} \rangle \right) \quad (4.3)$$

$$= \operatorname{sign}(v) \left( \iint_{C(I_y) \prec C(I_x)} \omega_{b,w} - \varepsilon(x) \left( \mathbf{1}_{\{r_x \in I_y\}} A_{b,w}^{r_x} - \mathbf{1}_{\{\ell_x \in I_y\}} A_{b,w}^{\ell_x} \right) \right), \quad (4.4)$$

where:

1. We set  $\operatorname{sign}(v) = +1$  if the color changes from black to white at the section  $s_v$ , and  $\operatorname{sign}(v) = -1$  otherwise.

2.  $\varepsilon(x) = +1$  (resp.  $-1$ ) if  $x$  is black (resp. white) in consistency with our sign/color convention (3.1).
3. For each  $x \in B(\mathbb{G}_\beta) \sqcup W(\mathbb{G}_\beta)$ , let  $R_x$  denote the chamber containing  $x$ . We associate to  $x$  an inner chamber  $R_x^{\text{in}}$  by setting

$$R_x^{\text{in}} := \begin{cases} R_x, & \text{if } x \text{ is inner,} \\ \text{the inner chamber adjacent to } x \text{ across } \beta_e, & \text{if } x \text{ is } \beta_e\text{-outer.} \end{cases}$$

If  $x$  is  $\beta_e$ -outer for more than one  $e \in E(N)$ , we choose one such  $e$  arbitrarily. The resulting constructions are independent of this choice; see Proposition 4.11.

The interval  $I_x = [\ell_x, r_x] \subset \partial N$  is defined as the sign interval of  $\text{sign}_{R_x^{\text{in}}}$  whose color is opposite to that of  $x$ , and which is adjacent to  $s_v(R_x^{\text{in}})$ ; the interval  $I_y$  is defined analogously. The contours  $C(I_x)$  and  $C(I_y)$  are then defined by Definition 2.17.

Equivalently, instead of specifying the convex vertex  $v$ , we can choose  $I_x$  to be either of the sign intervals for  $\text{sign}_{R_x^{\text{in}}}$  with opposite color as  $x$ , and similarly for  $I_y$ . The convex vertex is then uniquely determined as the common endpoint of both intervals when they are parallel transported to the same chamber.

4. Throughout,  $\zeta$  is integrated over  $C(I_b)$  and  $\eta$  is integrated over  $C(I_w)$ . The notation

$$C(I_y) \prec C(I_x)$$

denotes the product  $C(I_b) \times C(I_w)$  with the convention that if both contours  $C(I_x)$  and  $C(I_y)$  contain small circles around the same angle  $\alpha$ , we choose them so that the circle in  $C(I_y)$  has smaller radius than the one in  $C(I_x)$ .

**Remark 4.2.** Since there are four choices of convex vertices and two choices of  $(x, y) = (b, w)$  or  $(w, b)$ , the definition (4.3) gives eight different expressions. We will prove in Proposition 4.11 that the resulting quantities for all of these choices are equal. If we take  $x = b$  and choose the intervals  $I_b$  and  $I_w$  so that  $v$  is a convex vertex at which  $\text{sign}_\beta$  changes from black to white, then the signs  $\text{sign}(v)$  and  $\varepsilon(x)$  in (4.3)–(4.4) are  $+1$ . However, the freedom to vary these choices will be important in our proofs.

Our main result of this section is:

**Theorem 4.3.** *The matrix  $K^{-1}$  defined by (4.3) is the two-sided inverse of Fock's Kasteleyn matrix  $K$  on  $\mathbb{G}_\beta$ .*

**Corollary 4.4.** *The graph  $\mathbb{G}_\beta$  has an equal number of black and white vertices, and admits a dimer cover.*

## 4.2 Example: Aztec diamond

Consider the Aztec diamond with reference face  $f_0$  and strands labeled as shown in Figure 21(right). We switch to the notation in [27] and denote the four families of parallel strands by  $\alpha, \beta, \gamma$  and  $\delta$

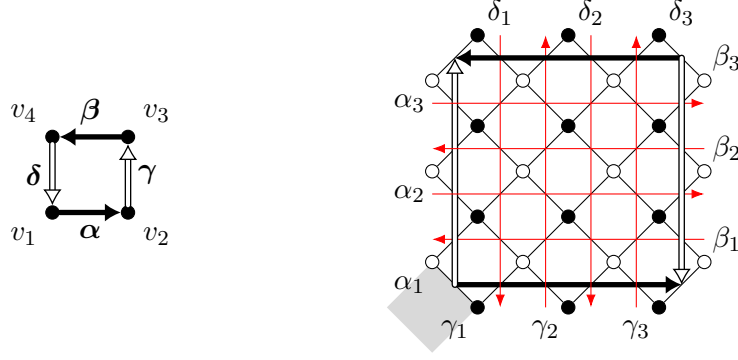


Figure 21: The sign intervals (left) and strands (right) in the Aztec diamond. The shaded face is the reference face  $f_0$ .

as shown in Figure 21(left). We check that (4.3) recovers the formula

$$K_{\mathbf{b},\mathbf{w}}^{-1} = \frac{1}{(2\pi i)^2} \int_{C_w} \int_{C_b} \frac{1}{E(\zeta, \eta)} \frac{\theta(-t + \mathbf{D}_\beta + \eta - \zeta)}{\theta(-t + \mathbf{D}_\beta)} \frac{g_w(\eta)}{g_b(\zeta)} \prod_{i=1}^n \frac{E(\delta_i, \eta)}{E(\beta_i, \eta)} \frac{E(\beta_i, \zeta)}{E(\delta_i, \zeta)} \quad (4.5)$$

$$- \mathbf{1}_{\{\mathbf{b} \text{ right of } \mathbf{w}\}} \frac{1}{2\pi i} \int_{C_w} g_{\mathbf{b},\mathbf{w}}$$

in [27, (17)], where  $C_b$  (resp.  $C_w$ ) is a counterclockwise oriented curve that encloses only the angles  $\gamma$  (resp.  $\alpha$ ).

In this example,

$$\mathbf{D}_\beta = \sum_{i=1}^n \delta_i - \sum_{i=1}^n \beta_i, \quad \Psi_\beta(\zeta) = \prod_{i=1}^n \frac{E(\delta_i, \zeta)}{E(\beta_i, \zeta)},$$

so the integrand inside the double integral is the same as in (4.3). There is only one inner chamber with sign intervals as shown in Figure 21(left). Thus, we can take

$$I_b = [v_2, v_3], \quad I_w = [v_1, v_2], \quad v = v_2, \quad \text{sign}(v) = 1,$$

so that  $C_b = C(I_b)$  and  $C_w = C(I_w)$  have no common angles. Therefore, the double integral terms are identical.

Our single integral term is

$$-\langle \partial I_b \cap I_w, \mathbf{A}_{\mathbf{b},\mathbf{w}} \rangle = A_{\mathbf{b},\mathbf{w}}^{v_2}.$$

If  $\mathbf{b}$  is on the left of (resp. above)  $\mathbf{w}$ ,  $g_{\mathbf{b},\mathbf{w}}$  has no poles at angles in  $\gamma$  (resp.  $\alpha$ ), so

$$-\mathbf{1}_{\{\mathbf{b} \text{ right of } \mathbf{w}\}} \frac{1}{2\pi i} \int_{C_w} g_{\mathbf{b},\mathbf{w}} = 0.$$

On the other hand, in both cases,  $v_2 \in \mathfrak{N}_{\mathbf{b},\mathbf{w}}$  so  $A_{\mathbf{b},\mathbf{w}}^{v_2} = 0$  as well. If  $\mathbf{b}$  is below and to the right of  $\mathbf{w}$ , then  $g_{\mathbf{b},\mathbf{w}}$  has no poles at angles in  $\beta$  and  $\delta$ , so  $v_1 \in \mathfrak{N}_{\mathbf{b},\mathbf{w}}$ , so we can take  $I_{\mathbf{b},\mathbf{w}}^{v_2} = [v_2, v_1]$ . Since the total sum of residues of  $g_{\mathbf{b},\mathbf{w}}$  is 0,

$$\frac{1}{2\pi i} \int_{C([v_1, v_2])} g_{\mathbf{b},\mathbf{w}} + \frac{1}{2\pi i} \int_{C([v_2, v_1])} g_{\mathbf{b},\mathbf{w}} = 0.$$

Hence,

$$A_{\mathbf{b},\mathbf{w}}^{v_2} = -\frac{1}{2\pi i} \int_{C_{\mathbf{w}}} g_{\mathbf{b},\mathbf{w}}.$$

Therefore, the single integral terms agree in all cases, and the two formulas for  $K_{\mathbf{b},\mathbf{w}}^{-1}$  coincide.

**Remark 4.5.** In this example, the eight *a priori* different choices reduce to four distinct formulas: when we interchange the roles of  $\mathbf{b}$  and  $\mathbf{w}$ , the double integral is unchanged since  $C(I_{\mathbf{b}})$  and  $C(I_{\mathbf{w}})$  do not enclose any common angles, and the single integral is unchanged since  $I_{\mathbf{b}} \cap I_{\mathbf{w}} = \{v\}$  and so changing  $\varepsilon(\mathbf{b})$  to  $\varepsilon(\mathbf{w})$  cancels the sign coming from changing  $\partial I_{\mathbf{b}}$  to  $\partial I_{\mathbf{w}}$ . The resulting expressions, indexed by the choice of convex vertex  $v$ , are

$$v_1 : \quad - \int_{C([v_1, v_2])} \int_{C([v_4, v_1])} \omega_{\mathbf{b},\mathbf{w}} + A_{\mathbf{b},\mathbf{w}}^{v_1}, \quad (4.6)$$

$$v_2 : \quad \int_{C([v_1, v_2])} \int_{C([v_2, v_3])} \omega_{\mathbf{b},\mathbf{w}} + A_{\mathbf{b},\mathbf{w}}^{v_2}, \quad (4.7)$$

$$v_3 : \quad - \int_{C([v_3, v_4])} \int_{C([v_2, v_3])} \omega_{\mathbf{b},\mathbf{w}} + A_{\mathbf{b},\mathbf{w}}^{v_3},$$

$$v_4 : \quad \int_{C([v_3, v_4])} \int_{C([v_4, v_1])} \omega_{\mathbf{b},\mathbf{w}} + A_{\mathbf{b},\mathbf{w}}^{v_4}.$$

Let us check that the expressions (4.6) and (4.7), corresponding to  $v_1$  and  $v_2$ , agree. Keep  $C(I_{\mathbf{w}}) = C([v_1, v_2])$  fixed and deform  $C(I_{\mathbf{b}})$  from  $C([v_4, v_1])$  to  $-C([v_2, v_3])$ . During this deformation, the only pole of  $\omega_{\mathbf{b},\mathbf{w}}$  crossed is along the diagonal  $\zeta = \eta$ , so the change in the double integral is the residue contribution

$$-\frac{1}{2\pi i} \int_{C([v_1, v_2])} g_{\mathbf{b},\mathbf{w}} = A_{\mathbf{b},\mathbf{w}}^{v_2} - A_{\mathbf{b},\mathbf{w}}^{v_1}.$$

Finally, replacing the inner contour  $-C([v_2, v_3])$  by  $C([v_2, v_3])$  reverses the sign of the double integral. Therefore, the formulas (4.6) and (4.7) coincide. The general case, proved in Proposition 4.11, uses the same underlying idea.

**Remark 4.6.** It is worth emphasizing that the indicators that appear in our formula (4.4) are of a different nature than the indicator in (4.5); indeed, as we have seen above, the latter indicator term is simply the whole-plane inverse  $A_{\mathbf{b},\mathbf{w}}^{v_2}$ .

### 4.3 Example: a pentagonal Newton polygon with isoradial weights

For simplicity, we consider the setting of *isoradial weights* from [78]. Namely, we take  $\Sigma = \mathbb{CP}^1$  and let  $A_0 = \{\zeta \in \mathbb{CP}^1 : |\zeta| = 1\}$  be the unit circle. In this case, the theta function is identically equal to 1, and the prime form is

$$E(\zeta, \eta) = \frac{\eta - \zeta}{\sqrt{d\zeta} \sqrt{d\eta}}.$$

After performing a gauge transformation that removes the factor  $\frac{1}{\sqrt{d\zeta} \sqrt{d\eta}}$ , Fock's Kasteleyn matrix (2.4) simplifies to

$$K_{\mathbf{w},\mathbf{b}} = \beta - \alpha. \quad (4.8)$$

These weights have the following geometric interpretation. We associate to  $\mathbf{G}$  its *quad graph*  $\mathbf{G}^\diamond$ , whose vertices are the vertices and faces of  $\mathbf{G}$ , and whose edges are the pairs  $(x, f)$  with  $x \in B(\mathbf{G}) \sqcup$



$W(\mathbf{G}), f \in F(\mathbf{G})$ , such that  $\times$  lies on the boundary of  $f$ . Thus, every face of  $\mathbf{G}^\diamond$  is a quadrilateral, and each edge of  $\mathbf{G}$  appears as one of its diagonals. The choice of angles determines an embedding of  $\mathbf{G}^\diamond$  in which every quadrilateral is a rhombus; see Figure 23(left). Precisely, if an edge  $e = \mathbf{b}w$  is crossed by the strands  $\alpha, \beta$ , then the corresponding rhombus has opposite sides in the directions given by the unit complex numbers  $\alpha$  and  $\beta$ . In this embedding, the Kasteleyn matrix entry (4.8) is the complex vector representing the dual diagonal of the rhombus. In particular, the isoradial weight is its modulus, *i.e.* the Euclidean length of the dual edge.

For small graphs, the formula (4.3) can be evaluated explicitly using residues.

Consider the graph shown in Figure 22. For each edge  $e \in E(N)$ , we assume that all angles in  $\alpha_e$  are equal. To simplify notation, we denote these angles by  $\alpha, \beta, \gamma, \delta, \epsilon$ , and we label each edge of  $N$  by its associated angle. Then, Fock's Kasteleyn matrix is

$$\mathbf{K} = \begin{array}{c} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \mathbf{w}_3 \\ \mathbf{w}_4 \\ \mathbf{w}_5 \\ \mathbf{w}_6 \\ \mathbf{w}_7 \end{array} \begin{array}{c} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \\ \mathbf{b}_4 \\ \mathbf{b}_5 \\ \mathbf{b}_6 \\ \mathbf{b}_7 \end{array} \begin{bmatrix} \delta - \gamma & \epsilon - \delta & 0 & \alpha - \epsilon & 0 & \beta - \alpha & \gamma - \beta \\ \gamma - \beta & \beta - \epsilon & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma - \beta & \beta - \alpha & 0 & 0 & 0 & 0 \\ 0 & \delta - \gamma & \alpha - \delta & \gamma - \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \epsilon - \gamma & \beta - \epsilon & \gamma - \beta & 0 \\ 0 & 0 & 0 & 0 & \epsilon - \delta & \delta - \gamma & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha - \delta & \delta - \gamma \end{bmatrix}.$$

Directly computing the inverse, we find, for instance,

$$\mathbf{K}_{\mathbf{b}_1, \mathbf{w}_1}^{-1} = -\frac{(\alpha - \beta)(\beta - \epsilon)(\gamma - \delta)}{(\alpha - \gamma)(\beta - \delta)^2(\gamma - \epsilon)}, \quad \mathbf{K}_{\mathbf{b}_5, \mathbf{w}_1}^{-1} = \frac{(\beta - \gamma)(\gamma - \delta)^2}{(\alpha - \gamma)(\beta - \delta)^2(\gamma - \epsilon)}. \quad (4.9)$$

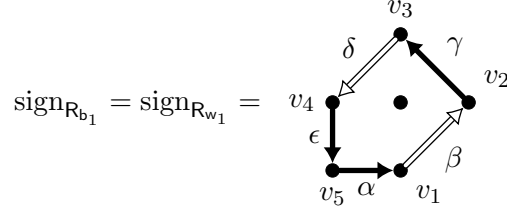
We now verify that these expressions agree with the evaluation of (4.3). From Figure 16, we compute the discrete Abel map:

	$\mathbf{d}$		$\mathbf{d}$
$\mathbf{w}_1$	$-\alpha$	$\mathbf{b}_1$	$-\alpha + \gamma + \delta$
$\mathbf{w}_2$	$-\alpha - \beta + \delta$	$\mathbf{b}_2$	$-\alpha + \delta + \epsilon$
$\mathbf{w}_3$	$-\alpha - \beta - \gamma + \delta + \epsilon$	$\mathbf{b}_3$	$-\gamma + \delta + \epsilon$
$\mathbf{w}_4$	$-\alpha + \gamma + \epsilon$	$\mathbf{b}_4$	$\epsilon$
$\mathbf{w}_5$	$-\gamma$	$\mathbf{b}_5$	$\beta - \gamma + \epsilon$
$\mathbf{w}_6$	$\beta - \gamma - \delta$	$\mathbf{b}_6$	$\beta$
$\mathbf{w}_7$	$-\alpha + \beta - \delta$	$\mathbf{b}_7$	$-\alpha + \beta + \gamma$

With our notation, (3.5) is  $\mathbf{D}_\beta = \beta - \gamma + \delta - \epsilon$ , so the boundary term

$$\Psi_\beta(\zeta) = \frac{(\zeta - \beta)(\zeta - \delta)}{(\zeta - \gamma)(\zeta - \epsilon)}.$$

We first consider  $K_{\mathbf{b}_1, \mathbf{w}_1}^{-1}$ . We have



so we can take

$$I_{\mathbf{b}_1} = [v_1, v_2], \quad I_{\mathbf{w}_1} = [v_4, v_1], \quad v = v_1, \quad \text{sign}(v) = 1.$$

The double integral is

$$\begin{aligned} & \frac{1}{(2\pi i)^2} \int_{C(I_{\mathbf{w}_1})} \int_{C(I_{\mathbf{b}_1})} \frac{1}{(\eta - \zeta)} \frac{(\zeta - \alpha)(\zeta - \epsilon)}{(\zeta - \beta)(\zeta - \delta)^2} \frac{(\eta - \beta)(\eta - \delta)}{(\eta - \alpha)(\eta - \gamma)(\eta - \epsilon)} d\zeta d\eta \\ &= -\frac{(\alpha - \beta)(\beta - \epsilon)}{(\beta - \delta)^2} \frac{1}{2\pi i} \int_{C(I_{\mathbf{w}_1})} \frac{(\eta - \delta)}{(\eta - \alpha)(\eta - \gamma)(\eta - \epsilon)} d\eta \\ &= -\frac{(\alpha - \beta)(\beta - \epsilon)}{(\beta - \delta)^2} \left( \frac{(\alpha - \delta)}{(\alpha - \gamma)(\alpha - \epsilon)} - \frac{(\delta - \epsilon)}{(\alpha - \epsilon)(\gamma - \epsilon)} \right) \\ &= -\frac{(\alpha - \beta)(\beta - \epsilon)(\gamma - \delta)}{(\alpha - \gamma)(\beta - \delta)^2(\gamma - \epsilon)}. \end{aligned}$$

Since  $\mathbf{d}(\mathbf{w}_1) - \mathbf{d}(\mathbf{b}_1) = -\gamma - \delta$ , the null sector  $\mathfrak{N}_{\mathbf{b}_1, \mathbf{w}_1} = [v_4, v_2]$ . Since

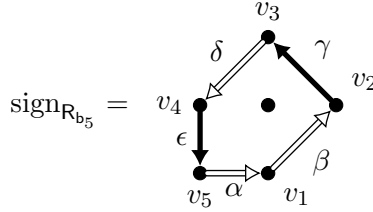
$$\partial I_{\mathbf{b}} \cap I_{\mathbf{w}} = (v_2 - v_1) \cap [v_4, v_1] = -v_1,$$

the single integral term is

$$-\langle \partial I_{\mathbf{b}} \cap I_{\mathbf{w}}, \mathbf{A}_{\mathbf{b}, \mathbf{w}} \rangle = \mathbf{A}_{\mathbf{b}, \mathbf{w}}^{v_1} = 0$$

since  $v_1 \in \mathfrak{N}_{\mathbf{b}_1, \mathbf{w}_1}$ . Thus, the total agrees with (4.9).

We now consider  $K_{\mathbf{b}_5, \mathbf{w}_1}^{-1}$ . We have



so we can take

$$I_{\mathbf{b}_5} = [v_3, v_4], \quad I_{\mathbf{w}_1} = [v_2, v_3], \quad v = v_3, \quad \text{sign}(v) = 1.$$

The double integral is

$$\begin{aligned} & \frac{1}{(2\pi i)^2} \int_{C(I_{\mathbf{w}_1})} \int_{C(I_{\mathbf{b}_5})} \frac{1}{(\eta - \zeta)} \frac{(\zeta - \gamma)^2}{(\zeta - \beta)^2(\zeta - \delta)} \frac{(\eta - \beta)(\eta - \delta)}{(\eta - \alpha)(\eta - \gamma)(\eta - \epsilon)} d\zeta d\eta \\ &= \frac{(\gamma - \delta)^2}{(\beta - \delta)^2} \frac{1}{2\pi i} \int_{C(I_{\mathbf{w}_1})} \frac{(\eta - \beta)}{(\eta - \alpha)(\eta - \gamma)(\eta - \epsilon)} d\eta \\ &= \frac{(\beta - \gamma)(\gamma - \delta)^2}{(\alpha - \gamma)(\beta - \delta)^2(\gamma - \epsilon)}, \end{aligned}$$

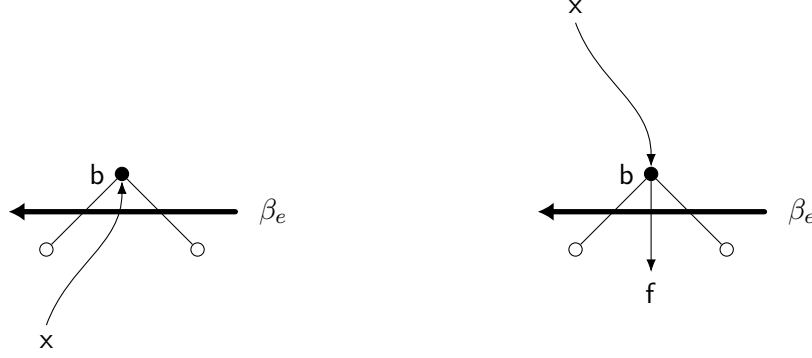


Figure 24: The two cases in the proof of Lemma 4.7:  $\text{sign}_{\mathbf{R}_x}(e) = +$  (left) and  $\text{sign}_{\mathbf{R}_x}(e) = -$  (right).

Since  $\mathbf{d}(w_1) - \mathbf{d}(b_5) = -\alpha - \beta + \gamma - \epsilon$ , the null sector  $\mathfrak{N}_{b_5, w_1} = [v_2, v_4]$ . Since

$$\partial I_b \cap I_w = (v_4 - v_3) \cap [v_2, v_3] = -v_3,$$

the single integral term is

$$-\langle \partial I_b \cap I_w, \mathbf{A}_{b, w} \rangle = \mathbf{A}_{b, w}^{v_3} = 0$$

since  $v_3 \in \mathfrak{N}_{b_5, w_1}$ . Thus, the total again agrees with (4.9).

#### 4.4 Correctness of the definition of $\mathbf{K}^{-1}$

**Lemma 4.7.** *Let  $e \in E(N)$ . If  $x \in V(G)$  is a vertex and  $\text{sign}_{\mathbf{R}_x}(e) = +$  (resp.  $-$ ), then*

$$\text{ord}_\alpha \left( \frac{1}{g_x \Psi_\beta} \right) \geq 0 \quad (\text{resp. } \leq 0)$$

for all  $\alpha \in \alpha_e$ .

*Proof.* By (3.2) and Lemma 2.11, we have

$$\text{ord}_\alpha \left( \frac{1}{g_x \Psi_\beta} \right) = (\mathbf{d}(y) - \mathbf{d}(x))_\alpha \tag{4.10}$$

where  $y$  is any  $\beta_e$ -outer vertex. Assume first that  $e$  is black so that  $y = \mathbf{b}$  is a black vertex. We have two cases (see Figure 24):

1.  $\text{sign}_{\mathbf{R}_x}(e) = +$ : Any path from  $x$  to  $\mathbf{b}$  crosses  $\beta_e$  positively, *i.e.*  $(\mathbf{d}(\mathbf{b}) - \mathbf{d}(x))_{\beta_e} = 1$ . By Lemma 2.12,  $(\mathbf{d}(\mathbf{b}) - \mathbf{d}(x))_\alpha \geq 0$  for all  $\alpha \in \alpha_e$ .
2.  $\text{sign}_{\mathbf{R}_x}(e) = -$ : Let  $f$  be the face separated from  $\mathbf{b}$  by  $\beta_e$ . Any path from  $x$  to  $f$  crosses  $\beta_e$  negatively, *i.e.*  $(\mathbf{d}(f) - \mathbf{d}(x))_{\beta_e} = -1$ . By Lemma 2.12,  $(\mathbf{d}(f) - \mathbf{d}(x))_\alpha \leq 0$  for all  $\alpha \in \alpha_e$ . Since  $\mathbf{d}(\mathbf{b}) - \mathbf{d}(x) = \mathbf{d}(f) - \mathbf{d}(x) + \beta_e$ , we get  $(\mathbf{d}(\mathbf{b}) - \mathbf{d}(x))_\alpha \leq 0$  for all  $\alpha \in \alpha_e$ .

The argument for white  $e$  is identical, with  $\mathbf{b}$  replaced by  $w$ . □

Recall Definition 2.14 of the null sector  $\mathfrak{N}_{b, w}$ .

**Lemma 4.8.** *Let  $e \in E(N)$ ,  $\mathbf{b} \in B(\mathbf{G}_\beta)$  and  $\mathbf{w} \in W(\mathbf{G}_\beta)$ . Assume that one of the following holds:*

1.  $\text{sign}_{\mathbf{R}_\mathbf{b}}(e) = +$  and  $\text{sign}_{\mathbf{R}_\mathbf{w}}(e) = -$ .
2.  $\text{sign}_{\mathbf{R}_\mathbf{w}}(e) = -$  and  $\mathbf{b}$  lies on  $\beta_e$ .
3.  $\text{sign}_{\mathbf{R}_\mathbf{b}}(e) = +$  and  $\mathbf{w}$  lies on  $\beta_e$ .

Then  $e \in \mathfrak{N}_{\mathbf{b},\mathbf{w}}$ .

*Proof.* We consider each of the three cases:

1. Suppose  $\text{sign}_{\mathbf{R}_\mathbf{b}}(e) = +$  and  $\text{sign}_{\mathbf{R}_\mathbf{w}}(e) = -$ . Then

$$(\mathbf{d}(\mathbf{w}) - \mathbf{d}(\mathbf{b}))_{\beta_e} = 1,$$

so Lemma 2.11 shows that  $g_{\mathbf{b},\mathbf{w}}$  has a zero at  $\beta_e$ . The conclusion therefore follows from the definition of  $\mathfrak{N}_{\mathbf{b},\mathbf{w}}$ .

2. Suppose  $\text{sign}_{\mathbf{R}_\mathbf{w}}(e) = -$  and  $\mathbf{b}$  lies on  $\beta_e$ . Let  $\mathbf{f}$  be the face on the other side of  $\beta_e$  from  $\mathbf{b}$ , as in Figure 24(right). Then

$$\mathbf{d}(\mathbf{w}) - \mathbf{d}(\mathbf{f}) = \mathbf{d}(\mathbf{w}) - \mathbf{d}(\mathbf{b}) + \beta_e.$$

Thus,  $g_{\mathbf{f},\mathbf{w}}$  has the same zeros and poles in  $A_0$  as  $g_{\mathbf{b},\mathbf{w}}$ , except for one additional zero at  $\beta_e$ .

We now distinguish the two cases in Definition 2.14 of  $\mathfrak{N}_{\mathbf{b},\mathbf{w}}$ .

- (a) If  $g_{\mathbf{b},\mathbf{w}}$  has at least one zero, then  $e \notin \mathfrak{N}_{\mathbf{b},\mathbf{w}}$  would imply that the edges carrying zeros and poles of  $g_{\mathbf{f},\mathbf{w}}$  interlace along  $\partial N$ , contradicting Lemma 2.13.
- (b) Suppose  $g_{\mathbf{b},\mathbf{w}}$  has no zeros in  $A_0$  and

$$\text{div}(g_{\mathbf{b},\mathbf{w}})|_{A_0} = -\alpha - \beta,$$

where  $\alpha, \beta \in \boldsymbol{\alpha}$  are as in Definition 2.14. Since  $\text{sign}_{\mathbf{R}_\mathbf{w}}(e) = -$ , the vertex  $\mathbf{w}$  does not lie on the other side of  $\beta_e$  from  $\mathbf{b}$ , so  $\alpha$  and  $\beta$  are distinct from  $\beta_e$ . Choosing  $\gamma = \beta_e$  in Case 2 of Definition 2.14, we get  $e \in \mathfrak{N}_{\mathbf{b},\mathbf{w}}$ .

3. This is symmetric to Case 2.

□

We say that a cyclic interval  $[a, b] \subset \partial N$  is *nontrivial* if  $a \neq b$ ; equivalently,  $[a, b]$  is neither a single point nor all of  $\partial N$ .

**Lemma 4.9.** *Let  $J = [a, b]$  and  $K = [c, d]$  be nontrivial cyclic closed intervals in  $\partial N$  with endpoints in  $V(N)$ . Then, as 0-chains,*

$$\partial(J \cap K) = \mathbf{1}_{\{b \in K\}}b - \mathbf{1}_{\{a \in K\}}a + \mathbf{1}_{\{d \in J\}}d - \mathbf{1}_{\{c \in J\}}c + \mathbf{1}_{\{a=c\}}a - \mathbf{1}_{\{b=d\}}b. \quad (4.11)$$

*Proof.* If  $J$  and  $K$  have no endpoints in common, then the formula is immediate: the boundary of  $J \cap K$  consists of those endpoints of  $J$  that lie in  $K$ , together with those endpoints of  $K$  that lie in  $J$ , with the usual signs. This gives

$$\partial(J \cap K) = \mathbf{1}_{\{b \in K\}}b - \mathbf{1}_{\{a \in K\}}a + \mathbf{1}_{\{d \in J\}}d - \mathbf{1}_{\{c \in J\}}c.$$

In general, the same expression remains correct except when two endpoints of the same type coincide. If  $a = c$ , then the point  $a$  is counted twice as a left endpoint, so one copy must be added back; this produces the term  $\mathbf{1}_{\{a=c\}}a$ . Similarly, if  $b = d$ , then the point  $b$  is counted twice as a right endpoint, so one copy must be subtracted; this produces the term  $-\mathbf{1}_{\{b=d\}}b$ . By contrast, if  $a = d$  or  $b = c$ , the two contributions occur with opposite signs and therefore cancel automatically, so no further correction is needed. This proves (4.11).  $\square$

**Corollary 4.10.** *Let  $\mathbf{b} \in B(\mathbf{G}_\beta)$  and  $\mathbf{w} \in W(\mathbf{G}_\beta)$ . Let  $J = [a, b]$  and  $K = [c, d]$  be any two nontrivial closed intervals in  $\partial N$  with endpoints in  $V(N)$ . Then*

$$\frac{1}{2\pi i} \int_{C(J \cap K)} g_{\mathbf{b}, \mathbf{w}} = -\langle \partial J \cap K, \mathbf{A}_{\mathbf{b}, \mathbf{w}} \rangle - \langle \partial K \cap J, \mathbf{A}_{\mathbf{b}, \mathbf{w}} \rangle - \mathbf{1}_{\{a=c\}} \mathbf{A}_{\mathbf{b}, \mathbf{w}}^a + \mathbf{1}_{\{b=d\}} \mathbf{A}_{\mathbf{b}, \mathbf{w}}^b.$$

*Proof.* We apply the linear functional  $\langle \cdot, \mathbf{A}_{\mathbf{b}, \mathbf{w}} \rangle$  from (4.2) to (4.11); see also Definition 2.17.  $\square$

**Proposition 4.11.** *The formula (4.3) is:*

1. *Invariant under swapping  $\mathbf{x}$  and  $\mathbf{y}$ .*
2. *Invariant under fixing  $I_{\mathbf{w}}$  and expanding  $I_{\mathbf{b}}$  across any edge  $e$  such that:*
  - $\text{sign}_{R_{\mathbf{b}}}(e) = +$ , or
  - $\mathbf{b}$  lies on  $\beta_e$  and  $e$  is black (in particular, this holds whenever  $\mathbf{b}$  is  $\beta_e$ -outer).

*Precisely, if  $I'_{\mathbf{b}} = [\ell'_{\mathbf{b}}, r'_{\mathbf{b}}]$  is obtained from  $I_{\mathbf{b}}$  by a sequence of such expansions, then*

$$\mathbf{K}_{\mathbf{b}, \mathbf{w}}^{-1} = \text{sign}(v) \left( \iint_{C(I_{\mathbf{w}}) \prec C(I'_{\mathbf{b}})} \omega_{\mathbf{b}, \mathbf{w}} - \langle \partial I'_{\mathbf{b}} \cap I_{\mathbf{w}}, \mathbf{A}_{\mathbf{b}, \mathbf{w}} \rangle \right). \quad (4.12)$$

*Similarly, we can fix  $I_{\mathbf{b}}$  and expand  $I_{\mathbf{w}}$  across any edge such that:*

- $\text{sign}_{R_{\mathbf{w}}}(e) = -$ , or
  - $\mathbf{w}$  lies on  $\beta_e$  and  $e$  is white (in particular, this holds whenever  $\mathbf{w}$  is  $\beta_e$ -outer).
3. *Independent of the choice of convex vertex  $v$ .*
  4. *Invariant under changing the choice of  $R_{\mathbf{x}}^{\text{in}}$  for an outer vertex  $\mathbf{x}$  whenever it is not unique.*

*Proof.* 1. It suffices to take  $(\mathbf{x}, \mathbf{y}) = (\mathbf{b}, \mathbf{w})$ . Exchanging the order of integration requires deforming contours so that, at common angles, the circles in  $C(I_{\mathbf{w}})$  lie strictly inside those in  $C(I_{\mathbf{b}})$ . The only new contribution arises from crossing the simple pole on the diagonal  $\zeta = \eta$ , giving

$$\begin{aligned} \mathbf{K}_{\mathbf{b}, \mathbf{w}}^{-1} &:= \text{sign}(v) \left( \iint_{C(I_{\mathbf{w}}) \prec C(I_{\mathbf{b}})} \omega_{\mathbf{b}, \mathbf{w}} - \langle \partial I_{\mathbf{b}} \cap I_{\mathbf{w}}, \mathbf{A}_{\mathbf{b}, \mathbf{w}} \rangle \right) \\ &= \text{sign}(v) \left( \iint_{C(I_{\mathbf{b}}) \prec C(I_{\mathbf{w}})} \omega_{\mathbf{b}, \mathbf{w}} - \frac{1}{2\pi i} \int_{C(I_{\mathbf{b}} \cap I_{\mathbf{w}})} g_{\mathbf{b}, \mathbf{w}} - \langle \partial I_{\mathbf{b}} \cap I_{\mathbf{w}}, \mathbf{A}_{\mathbf{b}, \mathbf{w}} \rangle \right). \end{aligned}$$

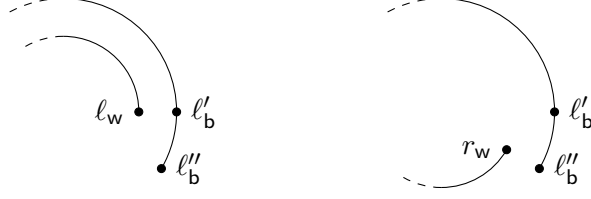


Figure 25: The two possible configurations in the case  $\text{sign}_{R_b^{\text{in}}}(e) = \text{sign}_{R_w}(e)$  considered in the proof of Item 2. The inner (resp. outer) cyclic interval is  $I_w$  (resp.  $I_b''$ ). The left panel shows the case  $l_b' = l_w$ , and the right panel the case  $l_b'' = r_w$ . In both cases  $\text{sign}_{R_w}(e) = -$  by definition of  $I_w$ .

Applying Corollary 4.10 to  $I_b = [\ell_b, r_b]$  and  $I_w = [\ell_w, r_w]$ , we get

$$-\frac{1}{2\pi i} \int_{C(I_b \cap I_w)} g_{b,w} - \langle \partial I_b \cap I_w, A_{b,w} \rangle = \langle \partial I_w \cap I_b, A_{b,w} \rangle + \mathbf{1}_{\{\ell_b = \ell_w\}} A_{b,w}^{\ell_b} - \mathbf{1}_{\{r_b = r_w\}} A_{b,w}^{r_b}.$$

It remains to show that the two terms with indicators vanish. Consider the term  $\mathbf{1}_{\{\ell_b = \ell_w\}} A_{b,w}^{\ell_b}$ . If  $\ell_b = \ell_w$ , let  $e$  be the boundary edge immediately clockwise from the common point  $\ell_b = \ell_w$ . Then, by definition of  $I_b$  and  $I_w$ ,

$$\text{sign}_{R_b^{\text{in}}}(e) = +, \quad \text{sign}_{R_w^{\text{in}}}(e) = -.$$

We have the following cases:

- $\text{sign}_{R_b^{\text{in}}}(e) = \text{sign}_{R_b}(e)$  and  $\text{sign}_{R_w^{\text{in}}}(e) = \text{sign}_{R_w}(e)$ : Lemma 4.8 implies that  $\ell_b = \ell_w \in \mathfrak{N}_{b,w}$  and therefore,  $A_{b,w}^{\ell_b} = 0$ .
- $\text{sign}_{R_b^{\text{in}}}(e) \neq \text{sign}_{R_b}(e)$  and  $\text{sign}_{R_w^{\text{in}}}(e) = \text{sign}_{R_w}(e)$ : Then  $b$  is  $\beta_e$ -outer; in particular it lies on  $\beta_e$ . Therefore, Lemma 4.8 again implies that  $\ell_b = \ell_w \in \mathfrak{N}_{b,w}$  and therefore,  $A_{b,w}^{\ell_b} = 0$ .
- $\text{sign}_{R_b^{\text{in}}}(e) = \text{sign}_{R_b}(e)$  and  $\text{sign}_{R_w^{\text{in}}}(e) \neq \text{sign}_{R_w}(e)$ : This is symmetric to the previous case.
- $\text{sign}_{R_b^{\text{in}}}(e) \neq \text{sign}_{R_b}(e)$  and  $\text{sign}_{R_w^{\text{in}}}(e) \neq \text{sign}_{R_w}(e)$ : This case cannot happen since it would mean that  $b$  and  $w$  are both  $\beta_e$ -outer but  $\beta_e$ -outer vertices must all have the same color.

Similarly,  $\mathbf{1}_{\{r_b = r_w\}} A_{b,w}^{r_b} = 0$ .

2. By Item 1, it suffices to consider the case where  $(x, y) = (b, w)$  and we expand  $I_b$ . Moreover, since moving the right endpoint is analogous, it suffices to show that (4.12) is invariant under moving the left endpoint of  $I_b'$  one step, *i.e.* let

$$I_b'' = [\ell_b'', r_b'], \quad e = [\ell_b'', \ell_b'] \in E(N).$$

We deform the contour  $C(I_b')$  to  $C(I_b'')$ , keeping  $C(I_w)$  fixed. Since  $\alpha_e$  does not contain any poles of the integrand (by Lemma 4.7 when  $\text{sign}_{R_b}(e) = +$ , or because (4.10) = 0 if  $b$  lies on  $\beta_e$  and  $e$  is black), the only poles that the deformation crosses are along the diagonal  $\zeta = \eta$ .

Therefore,

$$\begin{aligned} & \text{sign}(v) \left( \iint_{C(I_w) \prec C(I'_b)} \omega_{b,w} - \langle \partial I'_b \cap I_w, \mathbf{A}_{b,w} \rangle \right) \\ &= \text{sign}(v) \left( \iint_{C(I_w) \prec C(I''_b)} \omega_{b,w} + \frac{1}{2\pi i} \int_{C(e \cap I_w)} g_{b,w} - \langle \partial I'_b \cap I_w, \mathbf{A}_{b,w} \rangle \right). \end{aligned}$$

Thus, it suffices to show

$$\frac{1}{2\pi i} \int_{C(e \cap I_w)} g_{b,w} - \langle \partial I'_b \cap I_w, \mathbf{A}_{b,w} \rangle + \langle \partial I''_b \cap I_w, \mathbf{A}_{b,w} \rangle = \frac{1}{2\pi i} \int_{C(e \cap I_w)} g_{b,w} + \langle \partial e \cap I_w, \mathbf{A}_{b,w} \rangle = 0.$$

Applying Corollary 4.10 to the pair of intervals  $e$  and  $I_w = [\ell_w, r_w]$ ,

$$\begin{aligned} & \frac{1}{2\pi i} \int_{C(e \cap I_w)} g_{b,w} + \langle \partial e \cap I_w, \mathbf{A}_{b,w} \rangle \tag{4.13} \\ &= -\langle \partial I_w \cap e, \mathbf{A}_{b,w} \rangle - \mathbf{1}_{\{\ell''_b = \ell_w\}} \mathbf{A}_{b,w}^{\ell''_b} + \mathbf{1}_{\{\ell'_b = r_w\}} \mathbf{A}_{b,w}^{\ell'_b} \\ &= \left( \mathbf{1}_{\{\ell_w \in e\}} \mathbf{A}_{b,w}^{\ell_w} - \mathbf{1}_{\{\ell''_b = \ell_w\}} \mathbf{A}_{b,w}^{\ell''_b} \right) + \left( -\mathbf{1}_{\{r_w \in e\}} \mathbf{A}_{b,w}^{r_w} + \mathbf{1}_{\{\ell'_b = r_w\}} \mathbf{A}_{b,w}^{\ell'_b} \right). \end{aligned}$$

Since  $e = [\ell''_b, \ell'_b]$  is an edge,

$$\mathbf{1}_{\{\ell_w \in e\}} = \mathbf{1}_{\{\ell_w = \ell'_b\}} + \mathbf{1}_{\{\ell_w = \ell''_b\}}, \quad \mathbf{1}_{\{r_w \in e\}} = \mathbf{1}_{\{r_w = \ell'_b\}} + \mathbf{1}_{\{r_w = \ell''_b\}}.$$

Thus, (4.13) becomes

$$\mathbf{1}_{\{\ell'_b = \ell_w\}} \mathbf{A}_{b,w}^{\ell'_b} - \mathbf{1}_{\{\ell''_b = r_w\}} \mathbf{A}_{b,w}^{\ell''_b}. \tag{4.14}$$

We now have two cases:

- $\text{sign}_{R_w^{\text{in}}}(e) = \text{sign}_{R_w}(e)$ : If  $\ell'_b = \ell_w$  or if  $\ell''_b = r_w$ , then  $e$  is adjacent to an endpoint of  $I_w$ , hence  $\text{sign}_{R_w}(e) = -$  by definition of  $I_w$  (see Figure 25). Thus, the common endpoint lies in  $\mathfrak{N}_{b,w}$  by Cases 1 and 2 of Lemma 4.8, corresponding exactly to the two cases in the hypothesis of Item 2.
- $\text{sign}_{R_w^{\text{in}}}(e) \neq \text{sign}_{R_w}(e)$ : Then  $w$  is  $\beta_e$ -outer; in particular,  $e$  is white so the second hypothesis case that  $b$  lies on  $\beta_e$  and  $e$  is black cannot occur. In the first hypothesis case where  $\text{sign}_{R_b}(e) = +$ , any common endpoints are again, as in the previous bullet point, in  $\mathfrak{N}_{b,w}$  by Case 1 of Lemma 4.8.

Thus, in both cases, both terms in (4.14) are 0, completing the proof of Item 2.

3. By Item 1, it suffices to show that the formula is invariant under replacing  $I_b$  with the other white sign interval  $I'_b$  for  $R_b^{\text{in}}$ . We can enlarge  $I_b$  until it becomes the complement of  $I'_b$ . For every edge  $e$  that we absorb along the way, either

- $\text{sign}_{R_b}(e) = \text{sign}_{R_b^{\text{in}}}(e) = +$ , or
- $\text{sign}_{R_b}(e) \neq \text{sign}_{R_b^{\text{in}}}(e)$  so  $b$  is  $\beta_e$ -outer.

Thus, (4.12) is unchanged by Item 2. We then reverse the orientation, which complements the interval and flips  $\text{sign}(v)$ .

4. By Item 1, we may assume that  $\mathbf{x} = \mathbf{b}$  is a black vertex. By Condition 2 of Definition 3.2, the edges  $e \in E(N)$  for which  $\mathbf{b}$  is  $\beta_e$ -outer form a consecutive sequence

$$e_1 < e_2 < \cdots < e_k$$

along  $\partial N$ ; see Figure 10(right). Let  $v_1 < v_2 < v_3 < v_4$  be the convex vertices of  $N$  such that  $e_1, \dots, e_k \subset [v_2, v_3]$ .

If we choose  $e_i$  to define  $R_{\mathbf{b}}^{\text{in}}$ , then

$$\text{sign}_{R_{\mathbf{b}}^{\text{in}}}(e_i) = +, \quad \text{sign}_{R_{\mathbf{b}}^{\text{in}}}(e_j) = - \quad (j \neq i).$$

Considering the inconsistently oriented cones as in the proof of Lemma 3.18 and Figure 20, we obtain that  $s_{v_2}(R_{\mathbf{b}}^{\text{in}})$  is the left endpoint of  $e_i$  and  $s_{v_3}(R_{\mathbf{b}}^{\text{in}})$  is the right endpoint of  $e_i$ . If  $v \in \{v_1, v_2\}$ , then

$$I_{\mathbf{b}} = [s_{v_1}(R_{\mathbf{b}}^{\text{in}}), s_{v_2}(R_{\mathbf{b}}^{\text{in}})].$$

As  $e_i$  varies, the right endpoint of  $I_{\mathbf{b}}$  moves across the edges  $e_1, \dots, e_k$ . By Item 2, expanding or shrinking  $I_{\mathbf{b}}$  across these edges does not change the value of (4.3). The case  $v \in \{v_3, v_4\}$  is analogous. □

## 4.5 Proof of Theorem 4.3

Our proof is based on the same general ideas as those used for the Aztec diamond in [27]. Proposition 4.11 allows us to simplify parts of the argument from that setting. The added complexity of our proof comes from features of the general setting that do not arise in the Aztec diamond case.

It suffices to check that  $\mathbf{K}\mathbf{K}^{-1} = \mathbf{I}$ ; then  $\mathbf{K}^{-1}\mathbf{K} = \mathbf{I}$  (or, rather, its transpose) follows by Item 1 of Proposition 4.11. Let  $\mathbf{w}, \mathbf{w}' \in W(\mathbf{G}_{\beta})$ . We need to show that

$$\sum_{\mathbf{b} \sim \mathbf{w}'} \mathbf{K}_{\mathbf{w}', \mathbf{b}} \mathbf{K}_{\mathbf{b}, \mathbf{w}}^{-1} = \mathbf{1}_{\{\mathbf{w} = \mathbf{w}'\}}.$$

We distinguish two cases:

1.  $\mathbf{w}'$  is an inner vertex: Suppose first that  $R_{\mathbf{b}}^{\text{in}} = R_{\mathbf{w}'}^{\text{in}}$  for all  $\mathbf{b} \sim \mathbf{w}'$  so that  $I_{\mathbf{b}}$  is the same interval  $J = [\ell, r]$  for all of them. Then, we get

$$\sum_{\mathbf{b} \sim \mathbf{w}'} \mathbf{K}_{\mathbf{w}', \mathbf{b}} \left( \iint_{C(I_{\mathbf{w}}) \prec C(I_{\mathbf{b}})} \omega_{\mathbf{b}, \mathbf{w}} \right) = \iint_{C(I_{\mathbf{w}}) \prec C(J)} \left( \sum_{\mathbf{b} \sim \mathbf{w}'} \mathbf{K}_{\mathbf{w}', \mathbf{b}} \omega_{\mathbf{b}, \mathbf{w}} \right) = 0 \quad (4.15)$$

by Lemma 2.9 applied to  $\mathbf{x} = \mathbf{f}_0$  so that  $\frac{1}{g_{\mathbf{b}}(\zeta)} = g_{\mathbf{b}, \mathbf{f}_0}(\zeta)$ . Moreover, by (4.4) and Theorem 2.16,

$$- \sum_{\mathbf{b} \sim \mathbf{w}'} \mathbf{K}_{\mathbf{w}', \mathbf{b}} \text{sign}(v) \langle \partial I_{\mathbf{b}} \cap I_{\mathbf{w}}, \mathbf{A}_{\mathbf{b}, \mathbf{w}} \rangle = \text{sign}(v) (\mathbf{1}_{\{\ell \in I_{\mathbf{w}}\}} - \mathbf{1}_{\{r \in I_{\mathbf{w}}\}}) \mathbf{1}_{\{\mathbf{w} = \mathbf{w}'\}}.$$

In order for the right-hand side to be nonzero, we need  $w = w'$ . In that case, we show that

$$\mathbf{1}_{\{\ell \in I_w\}} - \mathbf{1}_{\{r \in I_w\}} = \text{sign}(v). \quad (4.16)$$

Indeed, if  $\text{sign}(v) = +1$ , then  $J$  (resp.  $I_w$ ) is the interval immediately counterclockwise (resp. clockwise) of  $s_v(R_b^{\text{in}}) = s_v(R_w^{\text{in}})$ , so  $\ell \in I_w$  and  $r \notin I_w$ , hence, the left-hand side equals 1. If instead  $\text{sign}(v) = -1$ , then  $J$  (resp.  $I_w$ ) is the interval immediately clockwise (resp. counterclockwise) of  $s_v(R_b^{\text{in}})$ , so  $r \in I_w$  and  $\ell \notin I_w$ , hence, the left-hand side equals  $-1$ .

Now we consider the general case where the intervals  $I_b$  may depend on  $b \sim w'$ . We first claim that we can replace  $I_b$  for all  $b \sim w'$  by the single uniform choice of interval

$$J := \bigcup_{b \sim w'} I_b$$

without changing the value of (4.3).

Let  $I$  denote the  $-$  sign interval for  $\text{sign}_{R_{w'}}$  that is adjacent to  $s_v(R_{w'})$ . Let  $e_-$  (resp.  $e_+$ ) denote the edge immediately clockwise (resp. counterclockwise) of  $I$ .

**Lemma 4.12.** *For every  $b \sim w'$ , the interval  $I_b$  is obtained from  $I$  by adjoining a (possibly empty) subset of  $\{e_-, e_+\}$ .*

*Proof.* We have two cases.

- $b$  is an inner vertex: As we move from  $R_{w'}$  to  $R_b$ , we cross exactly the two strands passing through the edge  $bw'$ . If one of the strands is  $\beta_e$  for some  $e \in E(N)$ , then  $\text{sign}_{R_{w'}}(e) = +$  changes to  $\text{sign}_{R_b}(e) = -$ . In particular, if we cross  $\beta_e$  for  $e \in \{e_-, e_+\}$ ,  $I_b$  is obtained from  $I$  by shifting the corresponding endpoint outwards.

We claim that an endpoint cannot shift by more than one step. If it did, then the two crossed strands would correspond to consecutive edges  $e$  and  $e'$  along  $\partial N$ . In that case the edge  $bw'$  would lie on both the zig-zag paths  $\beta_e$  and  $\beta_{e'}$ , hence would lie on  $\partial G_\beta$ . This contradicts the assumption that both  $w'$  and  $b$  are inner. Therefore, each endpoint moves by at most one step.

- $b$  is an outer vertex: Since  $w'$  is inner by assumption and  $b \sim w'$ , the edge  $bw'$  lies on at least one strand  $\beta_e$  for which  $b$  is  $\beta_e$ -outer. We may use such an  $e$  to define  $R_b^{\text{in}}$ . Then, moving from  $R_{w'}$  to  $R_b^{\text{in}}$  crosses exactly one strand, namely the strand through  $bw'$  other than  $\beta_e$ . Hence, at most one endpoint of  $I$  shifts and by at most one step.

□

We claim that every such adjoining step is allowed by Item 2 of Proposition 4.11.

**Lemma 4.13.** *Let  $b \sim w'$  and let  $e \in \{e_-, e_+\}$ . If  $bw'$  lies on  $\beta_e$  and  $e$  is white, then  $e \subset I_b$ .*

*Proof.* Since  $b$  lies on  $\beta_e$ ,  $\text{sign}_{R_b}(e) = -$ . Since  $e$  is white,  $b$  is not  $\beta_e$ -outer, so  $\text{sign}_{R_b^{\text{in}}}(e) = \text{sign}_{R_b}(e) = -$ . On the other hand,  $e$  is one of the two boundary edges adjacent to the interval  $I$ . By Lemma 4.12, the only way to obtain  $I_b$  from  $I$  is by adjoining some subset of  $\{e_-, e_+\}$ . Since  $\text{sign}_{R_b^{\text{in}}}(e) = -$ , the edge  $e$  must be adjoined. Therefore,  $e \subset I_b$ . □

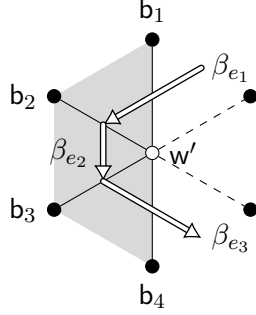


Figure 26: The labeling of black vertices in a neighborhood of an outer white vertex  $w'$  when  $k = 3$ . The shaded region represents faces of  $G_\beta$ . Solid edges belong to  $E(G_\beta)$  while dashed edges do not.

Let  $e \in \{e_-, e_+\}$  with  $e \subset J \setminus I_b$ . Since  $e$  is adjacent to the white sign interval  $I$  for  $\text{sign}_{R_{w'}}$ , we have  $\text{sign}_{R_{w'}}(e) = +$ . If  $b$  does not lie on  $\beta_e$ , then passing from  $R_{w'}$  to  $R_b$  does not cross  $\beta_e$ , so

$$\text{sign}_{R_b}(e) = \text{sign}_{R_{w'}}(e) = +,$$

and the first hypothesis of Item 2 of Proposition 4.11 applies. If instead  $b$  lies on  $\beta_e$  and  $e$  is black, then the second hypothesis of Item 2 of Proposition 4.11 applies. Finally, if  $b$  lies on  $\beta_e$  and  $e$  is white, then Lemma 4.13 implies that  $e \subset I_b$  already, contrary to  $e \subset J \setminus I_b$ . Thus, this case cannot occur.

Hence, every edge needed to enlarge  $I_b$  to  $J$  can be adjoined using Item 2 of Proposition 4.11, and therefore, we may replace every  $I_b$  by the common interval  $J$  without changing the value of  $K_{b,w}^{-1}$ .

To complete the proof for inner  $w'$ , it remains to verify (4.16) when  $w = w'$ . If  $\text{sign}(v) = +1$ , then  $I$  (resp.  $I_w$ ) is the interval immediately counterclockwise (resp. clockwise) of  $s_v(R_w^{\text{in}})$ . Regardless of whether the endpoints of  $J = [\ell, r]$  are the same as  $I$  or pushed out one step,  $\ell \in I_w$  and  $r$  is contained in the other black sign interval, so  $r \notin I_w$  and therefore,  $\mathbf{1}_{\{\ell \in I_w\}} - \mathbf{1}_{\{r \in I_w\}} = 1$ . The case when  $\text{sign}(v) = -1$  is analogous.

2.  $w'$  is an outer vertex: By Condition 2 of Definition 3.2, the edges  $e \in E(N)$  for which  $w'$  is  $\beta_e$ -outer form a consecutive block

$$e_1 < e_2 < \dots < e_k$$

along  $\partial N$ . Let  $v$  be the nearest convex vertex to this block such that  $\text{sign}(v) = 1$ , and set

$$J := \bigcup_{i=1}^k e_i. \tag{4.17}$$

Let  $b_1, \dots, b_{k+1}$  be the black vertices adjacent to  $w'$  that lie on at least one of the strands  $\beta_{e_i}$ ,  $1 \leq i \leq k$ , listed in counterclockwise order around  $w'$ , as in Figure 26. By considering the inconsistently oriented cones, exactly as in the proof of Lemma 3.18 and in Figure 20, we obtain

$$I_{b_1} = e_1, \quad I_{b_2} = e_1 \cup e_2, \quad \dots, \quad I_{b_k} = e_{k-1} \cup e_k, \quad I_{b_{k+1}} = e_k. \tag{4.18}$$

Equivalently, one can argue as follows. Passing from  $w'$  to  $\mathbf{b}_i$ , the number of sign changes increases from two to four by Lemma 3.15. The only signs that change are those on the intervals appearing in (4.18), and they change from  $+$  to  $-$ . These intervals must therefore be precisely the negative sign intervals.

Now suppose that, in enlarging  $I_{\mathbf{b}_i}$  to  $J$ , one has to adjoin an edge  $e_j$ . Figure 26 shows that  $\mathbf{b}_i$  lies to the left of  $\beta_{e_j}$ , hence,  $\text{sign}_{\mathbf{R}_{\mathbf{b}_i}}(e_j) = +$ . We may therefore apply Case 1 of Item 2 in Proposition 4.11.

To use the same calculation as in (4.15), we extend the definition of  $\mathbf{K}_{\mathbf{b},\mathbf{w}}^{-1}$  to the black vertices  $\mathbf{b} \sim \mathbf{w}'$  such that  $\mathbf{b} \in B(\mathbf{G}) \setminus B(\mathbf{G}_\beta)$  (the black vertices joined to  $\mathbf{w}'$  by dashed edges in Figure 26) by setting

$$\mathbf{K}_{\mathbf{b},\mathbf{w}}^{-1} := \text{sign}(v) \left( \iint_{C(J) \prec C(I_w)} \omega_{\mathbf{b},\mathbf{w}} + \langle \partial I_w \cap J, \mathbf{A}_{\mathbf{b},\mathbf{w}} \rangle \right). \quad (4.19)$$

We claim that (4.19) vanishes.

Indeed, since  $\mathbf{b}$  lies to the left of each strand  $\beta_{e_i}$ , the sign function  $\text{sign}_{\mathbf{R}_{\mathbf{b}}}$  is equal to  $+$  on  $J$ . Hence, by Lemma 4.7, the contour  $C(J)$  encloses no poles of  $\omega_{\mathbf{b},\mathbf{w}}$ , and therefore,

$$\iint_{C(J) \prec C(I_w)} \omega_{\mathbf{b},\mathbf{w}} = 0.$$

For the remaining term, we need to know that the sign stays positive also on the two boundary edges adjacent to  $J$ .

**Lemma 4.14.** *Assume  $\mathbf{b} \sim \mathbf{w}'$  such that  $\mathbf{b} \in B(\mathbf{G}) \setminus B(\mathbf{G}_\beta)$ . Let  $e_-$  (resp.  $e_+$ ) denote the edge immediately clockwise (resp. counterclockwise) of  $J$ . Then*

$$\text{sign}_{\mathbf{R}_{\mathbf{b}}}(e_-) = \text{sign}_{\mathbf{R}_{\mathbf{b}}}(e_+) = +.$$

*Proof.* As we move from  $\mathbf{b}_1$  to  $\mathbf{w}'$  to  $\mathbf{b}$ , Condition 2 of Definition 3.2 implies that the only boundary strand crossed is  $\beta_{e_1}$ . Therefore,

$$\text{sign}_{\mathbf{R}_{\mathbf{b}}}(e_-) = \text{sign}_{\mathbf{R}_{\mathbf{b}_1}}(e_-) = +.$$

Applying the same argument to  $\mathbf{b}_{k+1}$  gives

$$\text{sign}_{\mathbf{R}_{\mathbf{b}}}(e_+) = \text{sign}_{\mathbf{R}_{\mathbf{b}_{k+1}}}(e_+) = +.$$

□

Write  $I_w = [\ell_w, r_w]$ . Then,

$$-\langle \partial I_w \cap J, \mathbf{A}_{\mathbf{b},\mathbf{w}} \rangle = \mathbf{1}_{\{\ell_w \in J\}} \mathbf{A}_{\mathbf{b},\mathbf{w}}^{\ell_w} - \mathbf{1}_{\{r_w \in J\}} \mathbf{A}_{\mathbf{b},\mathbf{w}}^{r_w}.$$

We show that both terms on the right-hand side vanish. Consider first the term  $\mathbf{1}_{\{\ell_w \in J\}} \mathbf{A}_{\mathbf{b},\mathbf{w}}^{\ell_w}$ , and assume that  $\ell_w \in J$ . Let  $e'$  be the edge immediately clockwise from  $I_w$ . Since  $e'$  is adjacent to  $I_w$ , we have  $\text{sign}_{\mathbf{R}_w}(e') = -$ . On the other hand,  $\text{sign}_{\mathbf{R}_{\mathbf{b}}}$  is positive on  $J$  and on

the two edges adjacent to  $J$ , so  $\text{sign}_{R_b}(e') = +$ . It follows from Case 1 of Lemma 4.8 that  $\ell_w \in \mathfrak{N}_{b,w}$ , and therefore,

$$\mathbf{1}_{\{\ell_w \in J\}} A_{b,w}^{\ell_w} = 0.$$

A similar argument shows that

$$\mathbf{1}_{\{r_w \in J\}} A_{b,w}^{r_w} = 0.$$

We have thus shown that both terms in (4.19) are 0. Therefore,

$$\sum_{bw' \in E(G_\beta)} K_{w',b} K_{b,w}^{-1} = \sum_{bw' \in E(G)} K_{w',b} K_{b,w}^{-1} = \text{sign}(v) (\mathbf{1}_{\{\ell \in I_w\}} - \mathbf{1}_{\{r \in I_w\}}) \mathbf{1}_{\{w=w'\}}.$$

as in (4.15).

Finally, we need to verify (4.16) when  $w = w'$ . Recall the definition of  $v$  above (4.17). Suppose we use  $e_1$  to define  $R_{w'}^{\text{in}}$ , so that

$$R_{w'}^{\text{in}} = R_{b_1}.$$

Then  $I_w$  and  $I_{b_1} = e_1$  are adjacent sign intervals that meet at the vertex  $s_v(R_b) = \ell$ , so

$$\mathbf{1}_{\{\ell \in I_w\}} - \mathbf{1}_{\{r \in I_w\}} = 1 = \text{sign}(v).$$

The proof of Theorem 4.3 is complete.  $\square$

## 5 Scaling limits and astroidal domains

We now turn to the asymptotic consequences of the exact inverse Kasteleyn formula (4.3). Throughout the remainder of the paper, we assume that the angle function is periodic as in Section 2.4.6.

### 5.1 Astroidal domains

In this section, we consider scaling limits of AZ graphs. Recall from Section 2.1 that  $|e|_{\mathbb{Z}}$  is the lattice length of  $e \in E(N)$  and  $(a_e, b_e) = \vec{e}/|e|_{\mathbb{Z}} \in \mathbb{Z}^2$ . Let

$$c : E(N) \rightarrow \mathbb{R}$$

be a function. For each  $e \in E(N)$ , consider the line

$$\mathbf{l}_e := \{(x, y) \in \mathbb{R}^2 : -b_e x + a_e y + c(e) = 0\}.$$

For a vertex  $v = e_- \cap e_+ \in V(N)$ , let

$$\mathbf{p}_v := \mathbf{l}_{e_-} \cap \mathbf{l}_{e_+}$$

denote the point of intersection.

We first define a cycle in  $\mathbb{R}^2$  analogous to the medial boundary cycle construction. Let  $\partial \mathcal{D}_c$  be the counterclockwise-oriented cycle obtained as follows: it visits the points  $\mathbf{p}_v$  in the clockwise cyclic order of  $v \in V(N)$ , and for each edge  $e = [v_-, v_+] \in E(N)$ , it goes from  $\mathbf{p}_{v_+}$  to  $\mathbf{p}_{v_-}$  along  $\mathbf{l}_e$ . We do not assume that  $\mathbf{p}_{v_+} \neq \mathbf{p}_{v_-}$ .

Assume that  $\partial \mathcal{D}_c$  is a simple cycle, and let  $\mathcal{D}_c \subset \mathbb{R}^2$  be the closed set bounded by  $\partial \mathcal{D}_c$ . We say that  $\mathcal{D}_c$  is *admissible* if

$$\sum_{e \in E(N)} |e|_{\mathbb{Z}} c(e) = 0.$$

**Definition 5.1.** We call  $\mathcal{D}_c$  an *astroidal domain* if:

1. the cycle  $\partial\mathcal{D}_c$  is simple,
2.  $\mathcal{D}_c$  is admissible.

Astroidal domains naturally arise as scaling limits of AZ graphs. Let  $\mathbf{G}_{\beta^n}, n = 1, 2, \dots$ , be a family of AZ graphs such that for each  $e \in E(N)$ ,

$$c_{\beta^n}(e) = n|e|_{\mathbb{Z}}c(e) + O(1) \quad \text{as } n \rightarrow \infty. \quad (5.1)$$

By Lemma 3.8,  $\mathcal{D}_c$  is admissible. If  $\partial\mathcal{D}_c$  is also simple, then  $\mathcal{D}_c$  is an astroidal domain.

**Remark 5.2.** For every Newton polygon  $N$ , one can construct an astroidal domain. Choose two consecutive edges  $e_1, e_2 \in E(N)$ , and write

$$e_1 = [v_0, v_1], \quad e_2 = [v_1, v_2],$$

where  $v_0, v_1, v_2 \in V(N)$  occur consecutively in the cyclic order on  $\partial N$ . We first choose  $c : E(N) \rightarrow \mathbb{R}$  so that  $\partial\mathcal{D}_c$  is simple and so that the three vertices of  $\partial\mathcal{D}_c$  corresponding to  $v_0, v_1, v_2$  are three consecutive convex vertices. To arrange this, first place the oriented lines  $\mathbf{l}_{e_1}$  and  $\mathbf{l}_{e_2}$ . Among the four cones determined by these two oriented lines at  $\mathbf{p}_{v_1} = \mathbf{l}_{e_1} \cap \mathbf{l}_{e_2}$ , choose one of the inconsistently oriented cones. Then choose the remaining lines  $\mathbf{l}_e, e \neq e_1, e_2$ , in their cyclic order, so that the rest of  $\partial\mathcal{D}_c$  is a simple polygonal chain contained in this cone. We color the sides as in the definition of AZ graphs: an edge  $e \in E(N)$  is black if the orientation of  $\mathbf{l}_e$  induced by  $e$  agrees with the orientation of the corresponding side of  $\partial\mathcal{D}_c$ , and white otherwise. The color changes occur precisely at the convex vertices of  $\mathcal{D}_c$ . Hence, the two sides corresponding to  $e_1$  and  $e_2$  have opposite colors. Moreover, either of them can be translated arbitrarily far outward while keeping  $\partial\mathcal{D}_c$  simple. Replacing  $c(e)$  by  $c(e) + t$  for a black side, or by  $c(e) - t$  for a white side, translates that side outward. Thus, by translating  $e_1$  and  $e_2$  outward, we can make the domain admissible.

Astroidal domains can also be obtained from the tropical limit of the Aztec diamond, as in Appendix A.

Once an astroidal domain  $\mathcal{D}_c$  is fixed, one can choose boundary zig-zag paths so as to obtain a sequence of AZ graphs  $\mathbf{G}_{\beta^n}$  satisfying (5.1).

We expect the inverse Kasteleyn formula and the asymptotic analysis developed in this paper to extend to more general domains, including domains whose boundary cycle is not simple. For this reason, we do not pursue a more systematic construction of astroidal domains here beyond the above argument showing that the class is nonempty.

## 5.2 Cell decomposition and sign changes

The collection of lines  $\{\mathbf{l}_e : e \in E(N)\}$  determines a cell decomposition of  $\mathbb{R}^2$ ; see Figure 5. To each (0-, 1- or 2-) cell  $R$ , we associate a *sign function*

$$\text{sign}_R : E(N) \rightarrow \{+, -, 0\},$$

defined exactly as in (3.9), except that we use the oriented line  $\mathbf{l}_e$  (oriented in the direction of  $\vec{e}$ ) in place of the strand  $\beta_e$ . Explicitly, for  $e \in E(N)$ ,

$$\text{sign}_R(e) = \begin{cases} +, & \text{if } R \text{ lies on the left-hand side of } \mathbf{l}_e, \\ -, & \text{if } R \text{ lies on the right-hand side of } \mathbf{l}_e, \\ 0, & \text{if } R \subset \mathbf{l}_e. \end{cases}$$

The *chambers* are precisely those cells for which  $\text{sign}_R(e) \neq 0$  for all  $e \in E(N)$ . We say that a cell  $R$  has a *sign change*

1. at a vertex  $v \in V(N)$  if  $\text{sign}_R$  takes opposite values on the two edges of  $\partial N$  incident to  $v$ .
2. along a closed interval  $[v_-, v_+] \subset \partial N$  with  $v_-, v_+ \in V(N)$  if  $\text{sign}_R = 0$  on  $[v_-, v_+]$  and  $\text{sign}_R$  takes opposite values on the two adjacent boundary intervals.

With this convention, all definitions and results of Section 3.6 admit direct continuous analogues. In particular, we have the following continuous analogue of Lemma 3.15.

**Lemma 5.3.** *For each cell  $R \subset \mathcal{D}_c^\circ$  (resp.  $R \subset \mathbb{R}^2 \setminus \mathcal{D}_c^\circ$ ),  $\text{sign}_R$  has four (resp. two) sign changes.*

### 5.3 Characterization of astroidal domains

The limit shape offers a natural perspective on the astroidal domains, which we discuss in this section. Throughout this discussion, we assume that the points  $\mathbf{p}_v$  are pairwise distinct for all  $v \in V(N)$ .

Let  $N^\perp$  be the polygon obtained (up to translation) by rotating  $N$  by  $-\frac{\pi}{2}$ . For any  $v \in V(N)$  we denote the corresponding vertex of  $N^\perp$  by  $(s^v, t^v) \in V(N^\perp)$ .

**Proposition 5.4.** *Let  $\partial\mathcal{D} \subset \mathbb{R}^2$  be a simple polygon and assume there is an order-preserving bijection  $v \mapsto \mathbf{p}_v$  from  $V(N)$  to  $V(\partial\mathcal{D})$ . Assume further that there is a continuous piecewise linear function  $\mathfrak{h}_0 : \partial\mathcal{D} \rightarrow \mathbb{R}$  such that, for all  $v \in V(N)$ ,*

$$\mathfrak{h}_0(x, y) = (x, y) \cdot (s^v, t^v) - d_v \tag{5.2}$$

*along the two line segments of  $\partial\mathcal{D}$  adjacent to  $\mathbf{p}_v$ , for some constant  $d_v$ . Then the enclosed region  $\mathcal{D}$  is an astroidal domain  $\mathcal{D}_c$  with*

$$c(e) = \frac{d_{v_+} - d_{v_-}}{|e|_{\mathbb{Z}}}$$

*for  $e = [v_-, v_+] \in E(N)$ .*

**Remark 5.5.** We do not assume that the edges of  $\partial\mathcal{D}$  are parallel to the edges of  $N$ , nor that the map  $v \mapsto \mathbf{p}_v$  reverses orientation. Both properties follow from (5.2).

**Remark 5.6.** In Section 7.2 below, we obtain the limit shape for astroidal domains, and that limit shape indeed satisfies the boundary condition (5.2).

*Proof.* Recall the conventions from Section 2.1. We divide the proof into three steps. First, we show that each edge of  $\partial\mathcal{D}$  is parallel to its corresponding edge of  $N$ . Next, we show that the orientation of  $\partial\mathcal{D}$  is reversed compared to that of  $\partial N$ . Finally, we verify admissibility.

1. Let  $e = [v_-, v_+] \in E(N)$  and let  $\mathbf{l}_e$  be the line through  $\mathbf{p}_{v_-}$  and  $\mathbf{p}_{v_+}$ . For any  $(x, y) \in \mathbf{l}_e$ , (5.2) holds with  $v = v_-$  and with  $v = v_+$ . Subtracting the two expressions, we find that the equation for the line  $\mathbf{l}_e$  is

$$(x, y) \cdot (s^{v_+} - s^{v_-}, t^{v_+} - t^{v_-}) - d_{v_+} + d_{v_-} = 0.$$

By the definition of  $(s^v, t^v) \in V(N^\perp)$ ,

$$(s^{v_+} - s^{v_-}, t^{v_+} - t^{v_-}) = |e|_{\mathbb{Z}}(b_e, -a_e),$$

and therefore,

$$(x, y) \cdot |e|_{\mathbb{Z}}(b_e, -a_e) - d_{v_+} + d_{v_-} = 0. \quad (5.3)$$

In particular,  $\vec{e} = |e|_{\mathbb{Z}}(a_e, b_e)$  and  $\mathbf{l}_e$  are parallel.

2. For any  $e = [v_-, v_+] \in E(N)$ , let

$$\vec{u}_e := (s^{v_+} - s^{v_-}, t^{v_+} - t^{v_-}), \quad \text{and} \quad \vec{w}_e := \mathbf{p}_{v_+} - \mathbf{p}_{v_-}.$$

For any  $v = e_- \cap e_+ \in V(N)$ , let  $\theta_v \in (-\pi, \pi)$  be the angle between  $\vec{u}_{e_-}$  and  $\vec{u}_{e_+}$  and let  $\varphi_v \in (-\pi, \pi)$  be the angle between  $\vec{w}_{e_-}$  and  $\vec{w}_{e_+}$ . Since  $N$  is convex and oriented counterclockwise,  $\theta_v \in (0, \pi)$ . It follows, using  $\vec{u}_e \perp \vec{w}_e$ , that

$$\varphi_v = \theta_v \quad \text{or} \quad \varphi_v = \theta_v - \pi.$$

We say that  $\vec{w}_e$  is *reversed* at  $v$  if  $\varphi_v = \theta_v - \pi$ . If  $\vec{w}_e$  is reversed at  $v$ , it means that  $\vec{w}_e$  changes from being an inward pointing normal to an outward pointing normal of  $N^\perp$ , or the other way around.

Since  $\partial N$  is simple and oriented counterclockwise,

$$\sum_{v \in V(N)} \theta_v = 2\pi.$$

So, by summing over the angles  $\varphi_v$ , we get

$$\sum_{v \in V(N)} \theta_v - \pi \#\{v \in V(N) : \vec{w}_e \text{ is reversed at } v\} = 2\pi - \pi \#\{v \in V(N) : \vec{w}_e \text{ is reversed at } v\}.$$

Since  $\partial D$  is simple, the above sum is equal to  $2\pi$  or  $-2\pi$  according to whether  $\partial D$  is oriented counterclockwise or clockwise. We therefore only need to show that  $\vec{w}_e$  is reversed at least once.

By the assumption (5.2),

$$\mathfrak{h}_0(\mathbf{p}_{v_+}) - \mathfrak{h}_0(\mathbf{p}_{v_-}) = \vec{w}_e \cdot (s^{v_+}, t^{v_+}).$$

Summing over all  $e \in E(N)$  yields

$$\sum_{e=[v_-, v_+] \in E(N)} \vec{w}_e \cdot (s^{v_+}, t^{v_+}) = 0.$$

Now fix  $(s, t) \in N^\perp$ . Since  $\sum_{e \in E(N)} \vec{w}_e = 0$ , we have

$$\sum_{e=[v_-, v_+] \in E(N)} \vec{w}_e \cdot (s^{v_+} - s, t^{v_+} - t) = 0.$$

Since  $N^\perp$  is convex, the quantity

$$\vec{w}_e \cdot (s^{v_+} - s, t^{v_+} - t)$$

changes sign at  $v \in V(N)$  only if  $\vec{w}_e$  is reversed at  $v$ . Hence,  $\vec{w}_e$  is reversed at least once and so  $\partial \mathcal{D}$  is oriented opposite to  $\partial N$ .

3. Let  $e = [v_-, v_+]$ . With  $c(e) = (d_{v_+} - d_{v_-})/|e|_{\mathbb{Z}}$  as in the statement, the line  $\mathbf{l}_e$  given by the equation (5.3) coincides with the corresponding line given before Definition 5.1. The domain is admissible since

$$\sum_{e \in E(N)} |e|_{\mathbb{Z}} c(e) = \sum_{e = [v_-, v_+] \in E(N)} (d_{v_+} - d_{v_-}) = 0.$$

□

**Remark 5.7.** The previous proposition shows that, in the language of [8], astroidal domains are natural domains with natural boundary values.

## 6 Phase regions and the arctic curve

In this section, we study the geometry of the phase regions and the arctic curve, building on the methods developed in [10, 12, 18, 34, 48] in the context of dimer models. See also [85] for similar arguments in a different context. Recall that we assume that the angle function is periodic as in Section 2.4.6.

### 6.1 Action 1-form and action function

Fix an astroidal domain  $\mathcal{D}_c$  and a sequence of AZ graphs  $\mathbf{G}_{\beta^n}$ ,  $n = 1, 2, \dots$ , such that

$$c_{\beta^n}(e) = n|e|_{\mathbb{Z}}c(e) + O(1), \quad e \in E(N).$$

Recall the notion of imaginary-normalized differentials from Section 2.3.4. We also recall the (multi-valued) functions

$$z(\zeta) := \prod_{\alpha \in \alpha_{\mathbb{T}}} E(\alpha, \zeta)^{-b_{e(\alpha)}}, \quad w(\zeta) := \prod_{\alpha \in \alpha_{\mathbb{T}}} E(\alpha, \zeta)^{a_{e(\alpha)}},$$

introduced in (2.6). By Lemma 2.6, the logarithmic differentials  $d \log z$  and  $d \log w$  are imaginary-normalized meromorphic 1-forms with residues

$$\operatorname{res}_{\alpha}(d \log z) = -b_{e(\alpha)}, \quad \operatorname{res}_{\alpha}(d \log w) = a_{e(\alpha)}.$$

**Definition 6.1.** For  $(x, y) \in \mathbb{R}^2$ , we define the *action 1-form*  $\lambda_c(x, y)$  to be the unique imaginary-normalized meromorphic 1-form such that

$$\operatorname{res}_{\alpha}(\lambda_c(x, y)) = -b_{e(\alpha)}x + a_{e(\alpha)}y + c(e(\alpha)) \quad \text{for all } \alpha \in \alpha.$$

Let

$$\mathbf{D}_c := \sum_{\alpha \in \alpha} c(e(\alpha))\alpha. \tag{6.1}$$

Then the differential

$$x d \log z + y d \log w + \omega_{\mathbf{D}_c}$$

is imaginary-normalized by Lemma 2.6 and it has residue at  $\alpha$  equal to

$$-b_{e(\alpha)}x + a_{e(\alpha)}y + c(e(\alpha)).$$

Hence, by Lemma 2.5,

$$\lambda_c(x, y) = xd \log z + yd \log w + \omega_{\mathbf{D}_c}. \quad (6.2)$$

The reason for introducing  $\lambda_c(x, y)$  is that it governs the leading exponential factor in the inverse Kasteleyn matrix formula (4.3). To make this explicit, we introduce its multi-valued primitive.

**Definition 6.2.** The *action function* is the multi-valued function on  $\Sigma$  defined by

$$\Lambda_c(\zeta; x, y) := \sum_{\alpha \in \boldsymbol{\alpha}} (-b_{e(\alpha)}x + a_{e(\alpha)}y + c(e(\alpha))) \log E(\alpha, \zeta).$$

By construction,

$$d\Lambda_c(\cdot; x, y) = \lambda_c(x, y).$$

**Lemma 6.3.** Let  $(x, y) \in \mathbb{R}^2$  and let  $(j_n, k_n) \in \mathbb{Z}^2$  be such that

$$(j_n, k_n) = n(x, y) + O(1).$$

For every  $\alpha \in \boldsymbol{\alpha}$ ,

$$\begin{aligned} \frac{1}{n} \left( -\mathbf{d}(\mathbf{f}_0 + (j_n, k_n)) - \mathbf{D}\beta^n \right)_\alpha &= -b_{e(\alpha)}x + a_{e(\alpha)}y + c(e(\alpha)) + O\left(\frac{1}{n}\right) \\ &= \text{res}_\alpha(\lambda_c(x, y)) + O\left(\frac{1}{n}\right). \end{aligned}$$

*Proof.* Lemma 3.8 implies that

$$(\mathbf{D}\beta^n)_\alpha = -\frac{c\beta^n(e(\alpha))}{|e(\alpha)|_{\mathbb{Z}}} + O(1).$$

On the other hand,

$$\mathbf{d}(x + (j, k)) = \mathbf{d}(x) + \sum_{\alpha \in \boldsymbol{\alpha}} (b_{e(\alpha)}j - a_{e(\alpha)}k)\alpha,$$

so

$$\left( -\mathbf{d}(\mathbf{f}_0 + (j_n, k_n)) - \mathbf{D}\beta^n \right)_\alpha = -b_{e(\alpha)}j_n + a_{e(\alpha)}k_n + \frac{c\beta^n(e(\alpha))}{|e(\alpha)|_{\mathbb{Z}}} + O(1).$$

Using

$$(j_n, k_n) = n(x, y) + O(1), \quad c\beta^n(e) = n|e|_{\mathbb{Z}}c(e) + O(1),$$

we obtain

$$\frac{1}{n} \left( -\mathbf{d}(\mathbf{f}_0 + (j_n, k_n)) - \mathbf{D}\beta^n \right)_\alpha = -b_{e(\alpha)}x + a_{e(\alpha)}y + c(e(\alpha)) + O\left(\frac{1}{n}\right).$$

□

The previous lemma identifies the asymptotic exponent of the prime-form factor in the kernel (4.1). More precisely, if  $\mathbf{b}_n$  and  $\mathbf{w}_n$  are vertices whose macroscopic location is  $(x, y)$ , then

$$(\mathbf{d}(\mathbf{b}_n) + \mathbf{D}\beta^n)_\alpha = n \text{res}_\alpha(\lambda_c(x, y)) + O(1), \quad (\mathbf{d}(\mathbf{w}_n) + \mathbf{D}\beta^n)_\alpha = n \text{res}_\alpha(\lambda_c(x, y)) + O(1),$$

and therefore,

$$\frac{\prod_{\alpha \in \boldsymbol{\alpha}} E(\alpha, \eta)^{(\mathbf{d}(\mathbf{w}_n) + \mathbf{D}\beta^n)_\alpha}}{\prod_{\alpha \in \boldsymbol{\alpha}} E(\alpha, \zeta)^{(\mathbf{d}(\mathbf{b}_n) + \mathbf{D}\beta^n)_\alpha}} = \exp\left(n(\Lambda_c(\eta; x, y) - \Lambda_c(\zeta; x, y)) + O(1)\right). \quad (6.3)$$

In other words,  $\Lambda_c$  is the action governing the exponential part of the kernel.

## 6.2 Phase regions and boundary strata of the limit shape

The continuous analogue of Lemma 4.7 is:

**Lemma 6.4.** *Let  $e \in E(N)$  and let  $R$  be a cell. Then, for all  $\alpha \in \boldsymbol{\alpha}$  and  $(x, y) \in R$ ,*

$$\text{sign}(\text{res}_\alpha(\lambda_c(x, y))) = \text{sign}_R(e(\alpha)).$$

*Proof.* By Definition 6.1, for every  $\alpha \in \boldsymbol{\alpha}_e$  we have

$$\text{res}_\alpha(\lambda_c(x, y)) = -b_{e(\alpha)}x + a_{e(\alpha)}y + c(e(\alpha)),$$

which is the defining linear function of  $\boldsymbol{l}_{e(\alpha)}$ . Since  $\boldsymbol{l}_e$  is oriented along  $(a_{e(\alpha)}, b_{e(\alpha)})$ , this linear function is positive on the left-hand side of  $\boldsymbol{l}_{e(\alpha)}$ , negative on the right-hand side, and zero on  $\boldsymbol{l}_{e(\alpha)}$ . Therefore,

$$\text{sign}(\text{res}_\alpha(\lambda_c(x, y))) = \text{sign}_R(e(\alpha))$$

for all  $(x, y) \in R$ . □

Recall that we always count zeroes with multiplicity.

**Lemma 6.5.** *Let  $(x, y) \in \mathbb{R}^2$  and let  $R$  be the cell containing  $(x, y)$ . Then:*

- *For each  $1 \leq j \leq g$ , the number of zeroes of  $\lambda_c(x, y)$  in  $A_j$  is even and at least two.*
- *For consecutive poles  $\alpha, \beta \in A_0$  with  $e(\alpha) \neq e(\beta)$ ,  $\lambda_c(x, y)$  has an odd (resp. even) number of zeroes in  $(\alpha, \beta) \subset A_0$  when the residues at  $\alpha$  and  $\beta$  have the same (resp. opposite) sign. In particular, there is at least one zero in  $(\alpha, \beta)$  when its residues have the same sign.*
- *For consecutive poles  $\alpha, \beta \in A_0$  with  $e(\alpha) = e(\beta)$  and  $\text{res}_\alpha(\lambda_c(x, y)) = \text{res}_\beta(\lambda_c(x, y)) \neq 0$ ,  $\lambda_c(x, y)$  has an odd number of zeroes in  $(\alpha, \beta) \subset A_0$ . In particular, there is at least one zero.*

*The above accounts for all but possibly two zeroes of  $\lambda_c(x, y)$ . If  $R \subset \mathcal{D}_c^\circ$ , there are exactly two remaining zeroes, while if  $R \subset \mathbb{R}^2 \setminus \mathcal{D}_c^\circ$ , then all zeroes are accounted for.*

*Proof.* By Lemma 2.7, together with Lemma 6.4, the zeroes forced by the above conditions account for

$$2g + \sum_{e \in E(N): \text{sign}_R(e) \neq 0} |\boldsymbol{\alpha}_e| - \#\{\text{sign changes of } \text{sign}_R\}$$

zeroes of  $\lambda_c(x, y)$ . Since  $\deg(\lambda_c(x, y)) = 2g - 2$  and  $\lambda_c(x, y)$  has

$$\sum_{e \in E(N): \text{sign}_R(e) \neq 0} |\boldsymbol{\alpha}_e|$$

simple poles, its total number of zeroes is

$$2g + \sum_{e \in E(N): \text{sign}_R(e) \neq 0} |\boldsymbol{\alpha}_e| - 2.$$

Hence, the number of zeroes of  $\lambda_c(x, y)$  not accounted for by the above conditions is exactly

$$\#\{\text{sign changes of } \text{sign}_R\} - 2.$$

The final claim now follows from Lemma 5.3, which gives four sign changes when  $R \subset \mathcal{D}_c^\circ$  and two sign changes when  $R \subset \mathbb{R}^2 \setminus \mathcal{D}_c^\circ$ . □

**Lemma 6.6.** *Let  $(x, y) \in \mathcal{D}_c^\circ$ . With the conclusion of Lemma 6.5, exactly one of the following holds:*

1. *The two remaining zeroes form a  $\sigma$ -conjugate pair*

$$\{\zeta, \sigma(\zeta)\}, \quad \zeta \in \Sigma_+^\circ, \sigma(\zeta) \in \Sigma_-^\circ.$$

2. *The two remaining zeroes are real and lie in the same connected component of*

$$\Sigma(\mathbb{R}) \setminus \{\text{poles of } \lambda_c(x, y)\}.$$

*Proof.* By Lemma 2.6, the zero set of  $\lambda_c(x, y)$  is invariant under  $\sigma$ . Hence, these two remaining zeroes either both lie on  $\Sigma(\mathbb{R})$  or else they form a  $\sigma$ -conjugate pair.

Assume they lie on  $\Sigma(\mathbb{R})$ . By Lemma 6.5, on each connected component of  $\Sigma(\mathbb{R}) \setminus \{\text{poles of } \lambda_c(x, y)\}$ , the number of zeroes of  $\lambda_c(x, y)$  has a prescribed parity, and the parity of the number of zeroes already forced is correct, so the two remaining zeroes must lie in the same connected component of  $\Sigma(\mathbb{R}) \setminus \{\text{poles of } \lambda_c(x, y)\}$ .  $\square$

We now give a partial classification of regions in  $\mathcal{D}_c^\circ$  according to the locations of these two remaining zeroes. For some of these regions, we also define a map  $\Omega$  to  $\Sigma_+$ . The reader is encouraged to look ahead to Section 6.8 where we give a geometric reformulation.

**Definition 6.7.** We say that  $(x, y) \in \mathcal{D}_c^\circ$ :

1. lies in the *rough region*  $\mathcal{L}$  if  $\lambda_c(x, y)$  has a  $\sigma$ -conjugate pair of zeroes. In this case, there is a unique zero  $\zeta \in \Sigma_+^\circ$ , and we define

$$\Omega : \mathcal{L} \rightarrow \Sigma_+^\circ, \quad \Omega(x, y) := \zeta.$$

2. lies in the *frozen region*  $\mathcal{F}_{(\alpha, \beta)}$ , where  $(\alpha, \beta) \subset A_0$  is the open interval between two consecutive angles  $\alpha, \beta$  with  $e(\alpha) \neq e(\beta)$ , such that  $\lambda_c(x, y)$  has exactly two simple zeroes in  $(\alpha, \beta)$  when the vertex  $e(\alpha) \cap e(\beta) \in V(N)$  is convex, and two or three simple zeroes when it is concave.
3. lies in the *quasi-frozen region*  $\mathcal{Q}_{(\alpha, \beta)}$ , where  $(\alpha, \beta) \subset A_0$  is the open interval between two consecutive angles  $\alpha, \beta$  with  $e(\alpha) = e(\beta)$ , if  $\lambda_c(x, y)$  has three simple zeroes in  $(\alpha, \beta)$ .
4. lies in the *smooth region*  $\mathcal{G}_j$ , for  $1 \leq j \leq g$ , if  $\lambda_c(x, y)$  has four simple zeroes in  $A_j$ .
5. lies on the *frozen-rough boundary*  $\mathcal{A}_{(\alpha, \beta)}$ , where  $(\alpha, \beta) \subset A_0$  is the open interval between two consecutive angles  $\alpha, \beta$  with  $e(\alpha) \neq e(\beta)$ , if  $\lambda_c(x, y)$  has a zero  $\zeta$  of order two or three in  $(\alpha, \beta)$ . In this case,  $\zeta$  is unique, and we define

$$\Omega : \mathcal{A}_{(\alpha, \beta)} \rightarrow (\alpha, \beta), \quad \Omega(x, y) := \zeta.$$

If  $\zeta$  is a zero of order three, which can occur only when the vertex  $e(\alpha) \cap e(\beta) \in V(N)$  is concave, we call  $(x, y)$  a *frozen cusp*.

6. lies on the *quasi-frozen-rough boundary*  $\mathcal{A}_{(\alpha,\beta)}$ , where  $(\alpha, \beta) \subset A_0$  is the open interval between two consecutive angles  $\alpha, \beta$  with  $e(\alpha) = e(\beta)$ , if  $\lambda_c(x, y)$  has a zero  $\zeta$  of order two or three in  $(\alpha, \beta)$ . In this case,  $\zeta$  is unique, and we define

$$\Omega : \mathcal{A}_{(\alpha,\beta)} \rightarrow (\alpha, \beta), \quad \Omega(x, y) := \zeta.$$

If  $\zeta$  is a zero of order three, we call  $(x, y)$  a *quasi-frozen cusp*.

7. lies on the *smooth-rough boundary*  $\mathcal{A}_{A_j}$ , for  $1 \leq j \leq g$ , if  $\lambda_c(x, y)$  has a total of four zeroes in  $A_j$  (counted with multiplicity), one of which is a zero  $\zeta$  of order two or three. In this case,  $\zeta$  is unique: if  $\lambda_c(x, y)$  had two double zeroes on  $A_j$ , then it would have constant sign on  $A_j$ , forcing

$$\int_{A_j} \lambda_c(x, y) \neq 0,$$

contradicting imaginary-normalization. We therefore define

$$\Omega : \mathcal{A}_{A_j} \rightarrow A_j, \quad \Omega(x, y) := \zeta.$$

If  $\zeta$  is a zero of order three, we call  $(x, y)$  a *smooth cusp*; an order four zero is excluded by the same argument that rules out two double zeroes.

**Remark 6.8.** The above list is only a partial classification of the various singularities that can occur and does not in general exhaust  $\mathcal{D}_c^\circ$ ; for example, it does not include the frozen–frozen boundary in Figure 6. We do not attempt a complete classification of all possibilities in this paper.

We set

$$\mathcal{A}_{A_0 \setminus \alpha} := \bigsqcup_{(\alpha,\beta) \subset A_0} \mathcal{A}_{(\alpha,\beta)},$$

where the union ranges over all open intervals  $(\alpha, \beta)$  between consecutive angles on  $A_0$ . We also set

$$\mathcal{A}^\circ := \left( \bigsqcup_{j=1}^g \mathcal{A}_{A_j} \right) \sqcup \mathcal{A}_{A_0 \setminus \alpha}.$$

The next few subsections describe the geometry of the map  $\Omega$ . We first show in Section 6.3 that  $\Omega$  is a diffeomorphism from the rough region  $\mathcal{L}$  onto  $\Sigma_+^\circ$ . We then extend the inverse map  $\Omega^{-1}$  to  $\Sigma(\mathbb{R}) \setminus \alpha$  in Section 6.4. Finally, in Section 6.5, we extend  $\Omega^{-1}$  to the points of  $\alpha$ , whose images will be denoted by  $\mathcal{A}_\alpha$ ; these points are called *turning points*. Once  $\mathcal{A}_\alpha$  has been defined, we define the *arctic curve*

$$\mathcal{A} := \mathcal{A}^\circ \sqcup \mathcal{A}_\alpha.$$

We then show in Proposition 6.17 that these constructions glue together to give a smooth map

$$\Omega^{-1} : \Sigma_+ \rightarrow \overline{\mathcal{L}}.$$

In particular, we will prove that the arctic curve is precisely the boundary of the rough region, *i.e.*

$$\mathcal{A} = \partial \mathcal{L},$$

and moreover that the restriction

$$\Omega^{-1}|_{\Sigma(\mathbb{R})} : \Sigma(\mathbb{R}) \rightarrow \mathcal{A}$$

is a smooth parametrization of  $\mathcal{A}$ . By a smooth parameterization, we mean that the components of  $\Omega^{-1}$  are smooth functions on  $\Sigma(\mathbb{R})$ . The image  $\mathcal{A}$  need not be smooth; indeed, it typically has cusps at the critical values of  $\Omega^{-1}$ .

**Remark 6.9.** By general theory [8], the arctic curve is algebraic. The characterization of the arctic curve via the zeros of  $\lambda_c$  fits naturally into the framework developed in [101]. It may provide an alternative proof that the arctic curve is algebraic while simultaneously determining the degree of the curve. We will not pursue this argument here.

### 6.3 The rough region and the interior parametrization

Recall from (2.6) the definition of the functions  $z(\zeta)$  and  $w(\zeta)$ . While these functions may be multi-valued, the logarithmic differentials  $d \log z$  and  $d \log w$  are nevertheless well-defined meromorphic 1-forms.

**Lemma 6.10.** *Let  $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ . All zeroes of the 1-form*

$$x d \log z + y d \log w$$

*lie in  $\Sigma(\mathbb{R}) \setminus \alpha$ . Moreover, all zeroes that lie in the oval  $A_0$  are simple.*

*Proof.* The 1-form has residues

$$\text{res}_\alpha(xd \log z + yd \log w) = -b_{e(\alpha)}x + a_{e(\alpha)}y = \frac{1}{|e(\alpha)|_{\mathbb{Z}}} \langle e(\alpha), (y, -x) \rangle.$$

Since  $N$  is convex, the residues have exactly two sign changes, and therefore, Lemma 2.7 accounts for all of its zeroes. In particular, there are no zeroes in any interval of

$$A_0 \setminus \{\alpha \in \alpha : -b_{e(\alpha)}x + a_{e(\alpha)}y \neq 0\} \tag{6.4}$$

at whose endpoints the residues of  $x d \log z + y d \log w$  have opposite signs. It remains only to exclude zeroes at angles  $\alpha \in \alpha$  where the residue vanishes. Such an angle satisfies

$$-b_{e(\alpha)}x + a_{e(\alpha)}y = 0,$$

so the edge  $e(\alpha)$  is parallel to  $(x, y)$ . By convexity of  $N$ ,  $\alpha$  lies in an interval of (6.4) at whose endpoints the residues have opposite sign, and these intervals contain no zeroes as shown above.  $\square$

Given a local coordinate  $\zeta$  on  $\Sigma_+^\circ$ , write

$$d \log z = f_1(\zeta)d\zeta, \quad d \log w = f_2(\zeta)d\zeta, \quad \omega_{D_c} = f_3(\zeta)d\zeta. \tag{6.5}$$

**Proposition 6.11.** *The map  $\Omega : \mathcal{L} \rightarrow \Sigma_+^\circ$  is a diffeomorphism with inverse  $\Omega^{-1} : \Sigma_+^\circ \rightarrow \mathcal{L}$  given, for  $\zeta \in \Sigma_+^\circ$ , by*

$$x = \frac{\Im(\overline{f_2}f_3)}{\Im(\overline{f_1}f_2)}, \quad y = \frac{\Im(f_1\overline{f_3})}{\Im(\overline{f_1}f_2)}, \tag{6.6}$$

*where all functions are evaluated at  $\zeta$ .*

*Proof.* Let  $\zeta \in \Sigma_+^\circ$ . The vanishing of  $\lambda_c(x, y)$  at  $\zeta$  is equivalent to the system of linear equations

$$\begin{bmatrix} \Re f_1(\zeta) & \Re f_2(\zeta) \\ \Im f_1(\zeta) & \Im f_2(\zeta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} \Re f_3(\zeta) \\ \Im f_3(\zeta) \end{bmatrix}. \quad (6.7)$$

The system has a unique solution  $(x, y) \in \mathbb{R}^2$  provided the determinant

$$\det \begin{bmatrix} \Re f_1(\zeta) & \Re f_2(\zeta) \\ \Im f_1(\zeta) & \Im f_2(\zeta) \end{bmatrix} = \Im(\overline{f_1(\zeta)} f_2(\zeta)) \neq 0. \quad (6.8)$$

If this determinant vanishes, then

$$r := \frac{f_1(\zeta)}{f_2(\zeta)} \in \mathbb{R}.$$

By Lemma 6.10, the 1-form

$$d \log z - r d \log w$$

cannot vanish at  $\zeta$ , a contradiction.

Thus, the determinant is nonzero, so for every  $\zeta \in \Sigma_+^\circ$  the system (6.7) has a unique solution  $(x, y) \in \mathbb{R}^2$ , which is given by (6.6).

We now show that  $(x, y) \in \mathcal{D}_c^\circ$ , and therefore,  $(x, y) \in \mathcal{L}$ . If  $(x, y) \notin \mathcal{D}_c^\circ$ , then there are only two sign changes in the residues of  $\lambda_c(x, y)$  and so Lemma 6.5 accounts for all the zeroes which must therefore all be in  $\Sigma(\mathbb{R})$ , contradicting that  $\lambda_c(x, y)$  has a  $\sigma$ -conjugate pair of zeroes. Thus,  $(x, y) \in \mathcal{D}_c^\circ$ .

Hence, we obtain a map

$$\Omega^{-1} : \Sigma_+^\circ \rightarrow \mathcal{L}$$

sending  $\zeta$  to the unique solution (6.6). By construction,  $\Omega^{-1}$  is a right inverse to  $\Omega$ . Moreover, uniqueness of the solution to (6.7) implies that  $\Omega$  is injective, so  $\Omega$  is a bijection with inverse  $\Omega^{-1}$ .

Let

$$F(x, y; \zeta) := \frac{\lambda_c(x, y)}{d\zeta} = x f_1(\zeta) + y f_2(\zeta) + f_3(\zeta). \quad (6.9)$$

Let  $\zeta = u + iv$  and let

$$G(x, y; u, v) := \begin{bmatrix} \Re F(x, y; u + iv) \\ \Im F(x, y; u + iv) \end{bmatrix}.$$

Implicitly differentiating  $G(x, y; \Omega(x, y)) = 0$ , we obtain

$$\mathcal{J}_{(x,y)} G + \mathcal{J}_{(u,v)} G \mathcal{J} \Omega = 0,$$

where  $\mathcal{J}$  denotes the Jacobian matrix. Hence,

$$\mathcal{J} \Omega = -(\mathcal{J}_{(u,v)} G)^{-1} \mathcal{J}_{(x,y)} G,$$

and, therefore,

$$\det \mathcal{J} \Omega = \frac{\det \mathcal{J}_{(x,y)} G}{\det \mathcal{J}_{(u,v)} G}.$$

The numerator is (6.8) and the denominator is

$$\left| \frac{dF(x, y; \zeta)}{d\zeta} \right|^2$$

since  $F$  is holomorphic in  $\zeta$ . Therefore,

$$\det \mathcal{J} \Omega = \frac{\Im(\overline{f_1} f_2)}{|x f'_1 + y f'_2 + f'_3|^2}.$$

We have already seen that  $\Im(\overline{f_1} f_2) \neq 0$ , and the denominator is nonzero because  $\zeta$  is a simple zero of  $\lambda_c(x, y)$ . Therefore,  $\Omega$  is a local diffeomorphism. Since it is a bijection, it is a diffeomorphism.  $\square$

#### 6.4 Boundary parametrization away from angles

For  $\zeta_0 \in \Sigma(\mathbb{R}) \setminus \alpha$ , choose a local coordinate  $\zeta = u + iv$  centered at  $\zeta_0$  such that  $\sigma(\zeta) = \bar{\zeta}$  and such that  $\Sigma_+$  is locally given by  $v \geq 0$ . By Lemma 2.6, the 1-forms in (6.5) are  $\sigma$ -real, so  $f_j(\bar{\zeta}) = \overline{f_j(\zeta)}$ , and hence,

$$f_j(u) \in \mathbb{R}$$

for  $v = 0$ .

**Proposition 6.12.** *The map  $\Omega : \mathcal{A}^\circ \rightarrow \Sigma(\mathbb{R}) \setminus \alpha$  is a homeomorphism with inverse  $\Omega^{-1} : \Sigma(\mathbb{R}) \setminus \alpha \rightarrow \mathcal{A}^\circ$  given, for  $\zeta = u \in \Sigma(\mathbb{R}) \setminus \alpha$ , by*

$$x = \frac{f_2 f'_3 - f'_2 f_3}{f_1 f'_2 - f'_1 f_2}, \quad y = \frac{f'_1 f_3 - f_1 f'_3}{f_1 f'_2 - f'_1 f_2}, \quad (6.10)$$

where all functions are evaluated at  $u$ .

*Proof.* The condition  $\Omega(x, y) = u$  means that  $\lambda_c(x, y)$  vanishes at  $u$  to order at least two, which is equivalent to the linear system

$$\begin{bmatrix} f_1(u) & f_2(u) \\ f'_1(u) & f'_2(u) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} f_3(u) \\ f'_3(u) \end{bmatrix}. \quad (6.11)$$

We claim that the matrix is invertible. Indeed, if its determinant were zero, there would exist a nonzero vector  $(a, b) \in \mathbb{R}^2$  such that

$$\begin{bmatrix} f_1(u) & f_2(u) \\ f'_1(u) & f'_2(u) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0.$$

Equivalently, the 1-form

$$ad \log z + bd \log w$$

would vanish at  $u$  to order at least two, contradicting Lemma 6.10. Thus, the matrix in (6.11) is invertible, and the system has a unique solution which is given by (6.10).

If  $(x(u), y(u)) \notin \mathcal{D}_c^\circ$ , then there are only two sign changes in the residues of  $\lambda_c(x(u), y(u))$ , and so Lemma 6.5 accounts for all the zeroes. Hence:

- if  $u \in A_0$ , then  $u$  must be a simple zero, contradicting that it has order at least two.

- if  $u \in A_j$  and is a zero of order two, then since there are no other zeroes in  $A_j$ ,  $\lambda_c(x, y)$  would have constant sign on  $A_j$ , forcing

$$\int_{A_j} \lambda_c(x, y) \neq 0,$$

contradicting imaginary-normalization.

Thus,  $(x(u), y(u)) \in \mathcal{D}_c^\circ$ , and hence,  $(x(u), y(u)) \in \mathcal{A}^\circ$ .

Hence, we obtain a map

$$\Omega^{-1} : \Sigma(\mathbb{R}) \setminus \alpha \rightarrow \mathcal{A}^\circ$$

sending  $u$  to this unique solution. By construction,  $\Omega^{-1}$  is a right inverse to  $\Omega$ . Moreover, uniqueness of the solution to (6.11) implies that  $\Omega$  is injective. Therefore,  $\Omega$  is a homeomorphism with inverse  $\Omega^{-1}$ .  $\square$

## 6.5 Turning points

**Definition 6.13.** We say that  $(x, y) \in \mathcal{D}_c$  is a *turning point corresponding to*  $\alpha \in \alpha$  if  $\lambda_c(x, y)$  has a zero at  $\alpha$ . We denote by  $\mathcal{A}_\alpha$  the set of turning points.

For  $\alpha \in \alpha$ , choose a local coordinate  $\zeta$  centered at  $\alpha$  such that  $\sigma(\zeta) = \bar{\zeta}$ , and write

$$f_j(\zeta) = \frac{r_j}{\zeta} + s_j + O(|\zeta|), \quad r_j, s_j \in \mathbb{R}, \quad j = 1, 2, 3,$$

where, by definitions (2.6) of  $z$  and  $w$  and (6.1) of  $\mathbf{D}_c$ ,

$$r_1 = -b_{e(\alpha)}, \quad r_2 = a_{e(\alpha)}, \quad r_3 = c(e(\alpha)).$$

We then define

$$\Omega^{-1} : \alpha \rightarrow \mathcal{A}_\alpha$$

by sending  $\alpha$  to

$$x = \frac{r_2 s_3 - s_2 r_3}{r_1 s_2 - s_1 r_2}, \quad y = \frac{s_1 r_3 - r_1 s_3}{r_1 s_2 - s_1 r_2}. \quad (6.12)$$

**Proposition 6.14.** *For each  $\alpha \in \alpha$ , the point  $(x, y) = \Omega^{-1}(\alpha)$  is the unique turning point corresponding to  $\alpha$ , and it lies on  $\mathbf{l}_e$ .*

*Proof.* The 1-form  $\lambda_c(x, y)$  has a zero at  $\alpha$  if and only if both the pole term and the constant term vanish. Equivalently,  $(x, y)$  satisfies

$$\begin{bmatrix} r_1 & r_2 \\ s_1 & s_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} r_3 \\ s_3 \end{bmatrix}. \quad (6.13)$$

Solving (6.13) gives (6.12), provided the matrix is invertible.

We claim that the matrix in (6.13) is invertible. Indeed, if its determinant were zero, the 1-form

$$a_{e(\alpha)} d \log z + b_{e(\alpha)} d \log w$$

would have a zero at  $\alpha$ , contradicting Lemma 6.10. Therefore, (6.13) has a unique solution.

The point  $(x, y)$  is contained in  $\mathbf{l}_{e(\alpha)}$  since  $\text{res}_\alpha(\lambda_c(x, y)) = -b_{e(\alpha)}x + a_{e(\alpha)}y + c(e(\alpha)) = 0$ , and this is exactly the linear function defining  $\mathbf{l}_{e(\alpha)}$ .

Finally, we show that  $(x, y) \in \mathcal{D}_c$ . Suppose that  $(x, y) \notin \mathcal{D}_c$ , and let  $R$  be the cell containing  $(x, y)$ . Since  $(x, y) \in \mathbf{l}_{e(\alpha)}$ , we have  $\text{sign}_R(e(\alpha)) = 0$ . Let  $I \subset \partial N$  be the maximal closed interval containing  $e(\alpha)$  on which  $\text{sign}_R = 0$ , and let  $e_-$  and  $e_+$  be the nearest edges on the two sides of  $I$  such that

$$\text{sign}_R(e_-) \neq 0, \quad \text{sign}_R(e_+) \neq 0.$$

Since  $R \subset \mathbb{R}^2 \setminus \mathcal{D}_c^\circ$ , Lemma 5.3 implies that  $\text{sign}_R$  has exactly two sign changes. Crossing the line  $\mathbf{l}_{e(\alpha)}$  through  $(x, y)$  changes the sign at  $e(\alpha)$  and leaves all other signs unchanged. The two adjacent cells are still contained in  $\mathbb{R}^2 \setminus \mathcal{D}_c^\circ$ , and hence their sign functions also have exactly two sign changes. Therefore, the number of sign changes cannot change under this crossing, which is possible only if

$$\text{sign}_R(e_-) = -\text{sign}_R(e_+).$$

Thus,  $\text{sign}_R$  has a sign change along the closed interval  $I$ .

By Lemma 6.5, all zeroes of  $\lambda_c(x, y)$  are accounted for. Moreover, the connected component of

$$A_0 \setminus \{\text{poles of } \lambda_c(x, y)\}$$

corresponding to  $I$  has endpoints with opposite residues, and therefore contains no zeroes of  $\lambda_c(x, y)$ . In particular,  $\lambda_c(x, y)$  has no zero at  $\alpha$ , contradicting the definition of  $(x, y)$ . Hence  $(x, y) \in \mathcal{D}_c$ .  $\square$

**Lemma 6.15.** *Let  $\alpha \in \alpha_e$ , and let  $(x, y) = \Omega^{-1}(\alpha) \in \mathcal{A}_\alpha$  be the turning point corresponding to  $\alpha$ . Then:*

1. *If  $(x, y) \in \mathbf{l}_{e(\alpha)} \cap \partial \mathcal{D}_c$ , then  $\alpha$  is a simple zero of  $\lambda_c(x, y)$ .*
2. *If  $(x, y) \in \mathbf{l}_{e(\alpha)} \cap \mathcal{D}_c^\circ$ , then  $\alpha$  is a zero of  $\lambda_c(x, y)$  of order at most two.*

*Proof.* If  $(x, y) \in \mathbf{l}_{e(\alpha)} \cap \partial \mathcal{D}_c$ , then  $(x, y)$  lies on the boundary of  $\mathcal{D}_c$ . By Lemma 6.5,  $\alpha$  is simple.

Now suppose  $(x, y) \in \mathbf{l}_{e(\alpha)} \cap \mathcal{D}_c^\circ$  and let  $R$  be the cell containing  $(x, y)$ . Let  $e_-$  (resp.  $e_+$ ) denote the edges of  $N$  immediately clockwise (resp. counterclockwise) of  $e$ . Then

$$\text{sign}_R(e) = 0, \quad \text{sign}_R(e_-) \neq \text{sign}_R(e_+).$$

By Lemma 6.5, no zero in the connected component of

$$A_0 \setminus \{\text{poles of } \lambda_c(x, y)\}$$

containing  $\alpha$  is forced, so there are either none or two zeroes in that component. In particular,  $\alpha$  is a zero of order at most two.  $\square$

**Definition 6.16.** For  $\alpha \in \alpha$ , we call the turning point  $(x, y) = \Omega^{-1}(\alpha) \in \mathcal{A}_\alpha$  a *turning-point cusp* if  $\alpha$  is a zero of order two of  $\lambda_c(x, y)$ .

## 6.6 Global extension of $\Omega^{-1}$ and the arctic curve

Let  $\bar{\mathcal{L}}$  denote the closure of the rough region  $\mathcal{L}$  in  $\mathbb{R}^2$  and let  $\partial\mathcal{L} := \bar{\mathcal{L}} \setminus \mathcal{L}$  denote the boundary of  $\mathcal{L}$ .

**Proposition 6.17.** *The maps*

$$\Omega^{-1} : \Sigma_+^\circ \rightarrow \mathcal{L}, \quad \Omega^{-1} : \Sigma(\mathbb{R}) \setminus \alpha \rightarrow \mathcal{A}^\circ, \quad \Omega^{-1} : \alpha \rightarrow \mathcal{A}_\alpha$$

*glue together to a smooth map*

$$\Omega^{-1} : \Sigma_+ \rightarrow \bar{\mathcal{L}}.$$

*Moreover,*

$$\Omega^{-1}(\Sigma(\mathbb{R})) = \partial\mathcal{L} = \mathcal{A}.$$

*Proof.* First let  $\zeta_0 \in \Sigma(\mathbb{R}) \setminus \alpha$ . As in Section 6.4, choose a local coordinate  $\zeta = u + iv$  centered at  $\zeta_0$  such that  $\sigma(\zeta) = \bar{\zeta}$  and such that  $\Sigma_+$  is locally given by  $v \geq 0$ . Since  $f_j(u + iv) \in \mathbb{R}$  when  $v = 0$ , the functions

$$\Im(\overline{f_j(\zeta)} f_k(\zeta))$$

vanish on  $v = 0$ . By Taylor's theorem, there exist smooth functions  $B_{jk}(u, v)$  such that locally

$$\Im(\overline{f_j(\zeta)} f_k(\zeta)) = v B_{jk}(u, v).$$

Hence, for  $v > 0$ , (6.6) becomes

$$x = \frac{B_{23}}{B_{12}}, \quad y = \frac{B_{31}}{B_{12}}.$$

Locally expanding  $f_j(\zeta)$  and  $f_k(\zeta)$ , we obtain

$$B_{jk}(u, v) = f_j(\zeta_0) f'_k(\zeta_0) - f'_j(\zeta_0) f_k(\zeta_0) + O(|\zeta|).$$

Thus,

$$B_{12}(0, 0) = f_1(\zeta_0) f'_2(\zeta_0) - f'_1(\zeta_0) f_2(\zeta_0) \neq 0$$

by Proposition 6.12, so the above formulas define a smooth extension of  $\Omega^{-1}$  across  $\zeta_0$ , and its value at  $\zeta_0$  is exactly (6.10).

Now let  $\alpha \in \alpha$ , and choose a local coordinate  $\zeta = u + iv$  centered at  $\alpha$  such that  $\sigma(\zeta) = \bar{\zeta}$ . Write

$$f_j(\zeta) = \frac{r_j}{\zeta} + s_j + O(|\zeta|), \quad r_j, s_j \in \mathbb{R},$$

and let

$$h_j(\zeta) := \zeta f_j(\zeta) = r_j + s_j \zeta + O(|\zeta|^2).$$

Then each  $h_j$  is holomorphic and satisfies

$$h_j(\bar{\zeta}) = \overline{h_j(\zeta)},$$

so  $h_j(u) \in \mathbb{R}$  for  $v = 0$ . Since  $|\zeta|^2$  is real, the formula (6.6) can be rewritten for  $v > 0$  as

$$x = \frac{\Im(\overline{h_2 h_3})}{\Im(\overline{h_1 h_2})}, \quad y = \frac{\Im(h_1 \overline{h_3})}{\Im(\overline{h_1 h_2})}.$$

Again by Taylor's theorem, locally

$$\Im(\overline{h_j(\zeta)}h_k(\zeta)) = v\tilde{B}_{jk}(u, v)$$

for smooth  $\tilde{B}_{jk}$ . Hence,

$$x = \frac{\tilde{B}_{23}}{\tilde{B}_{12}}, \quad y = \frac{\tilde{B}_{31}}{\tilde{B}_{12}}.$$

At  $(u, v) = (0, 0)$ ,

$$\tilde{B}_{12}(0, 0) = r_1s_2 - s_1r_2 \neq 0$$

by Proposition 6.14. Therefore, these formulas extend smoothly across  $\alpha$ , and their value at  $\alpha$  is exactly given by (6.12). We have therefore obtained a smooth map

$$\Omega^{-1} : \Sigma_+ \rightarrow \bar{\mathcal{L}}$$

whose restrictions to  $\Sigma_+^\circ$ ,  $\Sigma(\mathbb{R}) \setminus \alpha$ , and  $\alpha$  are the three maps above.

Since  $\Omega^{-1}(\Sigma(\mathbb{R})) \subset \mathcal{A}^\circ \sqcup \mathcal{A}_\alpha$ , it is disjoint from  $\mathcal{L}$ , and hence,

$$\Omega^{-1}(\Sigma(\mathbb{R})) \subset \partial\mathcal{L}.$$

Conversely, let  $(x, y) \in \partial\mathcal{L}$ , and choose  $(x_n, y_n) \in \mathcal{L}$  with  $(x_n, y_n) \rightarrow (x, y)$ . Set  $\zeta_n := \Omega(x_n, y_n) \in \Sigma_+^\circ$ . By compactness of  $\Sigma_+$ , after passing to a subsequence, we may assume  $\zeta_n \rightarrow \zeta \in \Sigma_+$ . By continuity of  $\Omega^{-1}$ ,

$$(x, y) = \lim_{n \rightarrow \infty} (x_n, y_n) = \lim_{n \rightarrow \infty} \Omega^{-1}(\zeta_n) = \Omega^{-1}(\zeta).$$

If  $\zeta \in \Sigma_+^\circ$ , then  $(x, y) = \Omega^{-1}(\zeta) \in \mathcal{L}$ , a contradiction. Hence,  $\zeta \in \Sigma(\mathbb{R})$ , and therefore

$$\partial\mathcal{L} \subset \Omega^{-1}(\Sigma(\mathbb{R})).$$

Thus,

$$\partial\mathcal{L} = \Omega^{-1}(\Sigma(\mathbb{R})).$$

Since

$$\Omega^{-1}(\Sigma(\mathbb{R}) \setminus \alpha) = \mathcal{A}^\circ, \quad \Omega^{-1}(\alpha) = \mathcal{A}_\alpha,$$

we conclude that

$$\partial\mathcal{L} = \mathcal{A}^\circ \sqcup \mathcal{A}_\alpha = \mathcal{A}.$$

□

## 6.7 Geometry of the arctic curve

The map

$$\gamma : \Sigma \rightarrow \mathbb{CP}^1, \quad \zeta \mapsto \frac{d \log w}{d \log z}(\zeta)$$

is called the *logarithmic Gauss map*.

By Proposition 6.17, the restriction

$$\Omega^{-1}|_{\Sigma(\mathbb{R})} : \Sigma(\mathbb{R}) \rightarrow \mathbb{R}^2$$

is a smooth parametrization of the arctic curve  $\mathcal{A}$ . We now identify its critical points and compute tangent lines to  $\mathcal{A}$ . We phrase the statement about tangent lines in terms of branches since we have not yet established injectivity of the parameterization; we will show this in Corollary 6.21.

**Proposition 6.18.** *Let  $\zeta_0 \in \Sigma(\mathbb{R})$ . Then:*

1. *The point  $\zeta_0$  is a critical point of  $\Omega^{-1}|_{\Sigma(\mathbb{R})}$  if and only if  $\Omega^{-1}(\zeta_0)$  is one of the four types of cusps in Definition 6.7 and Definition 6.16.*
2. *Each of these cusps is an ordinary cusp in the geometric sense, i.e. after a smooth change of coordinates, it is locally given by  $y^2 = x^3$ .*
3. *The tangent line to the branch of  $\mathcal{A}$  at  $\Omega^{-1}(\zeta_0)$  determined by  $\zeta_0$  has slope*

$$-\frac{1}{\gamma(\zeta_0)} = -\frac{f_1(\zeta_0)}{f_2(\zeta_0)}.$$

*In particular, if  $\zeta_0 = \alpha \in \boldsymbol{\alpha}$ , then the tangent line is  $\mathbf{l}_{e(\alpha)}$ .*

*Proof.* First, we consider the case  $\zeta_0 \in \Sigma(\mathbb{R}) \setminus \boldsymbol{\alpha}$ . Choose a local coordinate  $\zeta = u + iv$  centered at  $\zeta_0$  such that  $\sigma(\zeta) = \bar{\zeta}$ . Write

$$(x(u), y(u)) := \Omega^{-1}(u),$$

and set

$$W := f_1 f_2' - f_1' f_2, \quad H := x f_1'' + y f_2'' + f_3'', \quad (6.14)$$

evaluated at  $\zeta_0$ . Since  $\lambda_c(x(u), y(u))$  has a zero of order at least two at  $u$ , we have

$$x(u) f_1(u) + y(u) f_2(u) + f_3(u) = 0, \quad x(u) f_1'(u) + y(u) f_2'(u) + f_3'(u) = 0.$$

Differentiating with respect to  $u$  gives

$$\begin{aligned} f_1 x' + f_2 y' &= 0, \\ f_1' x' + f_2' y' &= -H. \end{aligned}$$

Solving this linear system yields

$$(x', y') = \frac{H}{W}(f_2, -f_1). \quad (6.15)$$

By Proposition 6.12, we have  $W \neq 0$  and  $(f_2, -f_1) \neq 0$ . Hence,  $(x', y') = 0$  if and only if  $H = 0$ , i.e. if and only if  $\lambda_c(x, y)$  has a zero of order at least three at  $\zeta_0$ . This proves Item 1 in this case.

A standard criterion for an ordinary cusp (see, for example, [47, Definition 5.3.9]) is that

$$(x', y') = 0 \quad \text{and} \quad \det \begin{bmatrix} x'' & y'' \\ x''' & y''' \end{bmatrix} \neq 0;$$

Suppose  $\Omega^{-1}(\zeta_0)$  is one of the cusp types from Definition 6.7. Since  $\lambda_c(x, y)$  has a zero of exactly order three at  $\zeta_0$ ,

$$H = 0, \quad x f_1''' + y f_2''' + f_3''' \neq 0.$$

Therefore,

$$(x', y') = \frac{H}{W}(f_2, -f_1) = 0. \quad (6.16)$$

Differentiating (6.15), we get

$$\begin{aligned}(x'', y'') &= \left(\frac{H}{W}\right)' (f_2, -f_1) + \frac{H}{W} (f_2', -f_1'), \\(x''', y''') &= \left(\frac{H}{W}\right)'' (f_2, -f_1) + 2 \left(\frac{H}{W}\right)' (f_2', -f_1') + \frac{H}{W} (f_2'', -f_1'').\end{aligned}\tag{6.17}$$

By (6.16),

$$H' = x' f_1'' + y' f_2'' + x f_1''' + y f_2''' + f_3''' = x f_1''' + y f_2''' + f_3'''.$$

Using  $H = 0$  and simplifying, we get

$$\det \begin{bmatrix} x'' & y'' \\ x''' & y''' \end{bmatrix} = \frac{2(x f_1''' + y f_2''' + f_3''')^2}{W} \neq 0,$$

so  $\Omega^{-1}(\zeta_0)$  is an ordinary cusp in the geometric sense, proving Item 2 in this case.

If  $\zeta_0$  is not a critical point, then the tangent line to the branch of  $\mathcal{A}$  at  $(x, y)$  determined by  $\zeta_0$  has slope

$$\frac{y'}{x'} = -\frac{f_1}{f_2} = -\frac{1}{\gamma(\zeta_0)}.$$

When  $\zeta_0$  is a critical point, the ordinary cusp  $\Omega^{-1}(\zeta_0)$  still has a well-defined tangent line spanned by

$$(x'', y'') = \frac{H'}{W} (f_2, -f_1)$$

which gives the same slope  $-1/\gamma(\zeta_0)$ , completing the proof of Item 3 in this case.

Now we consider the case  $\zeta_0 = \alpha \in \boldsymbol{\alpha}$ . Define

$$h_j(\zeta) := \zeta f_j(\zeta).$$

All items now follow exactly as in the previous case with  $h_j$  in place of  $f_j$ .

Finally, the tangent vector at  $\Omega^{-1}(\alpha)$  is

$$(h_2, -h_1) = (a_{e(\alpha)}, b_{e(\alpha)}),$$

which is the vector along  $\mathbf{l}_{e(\alpha)}$ . Since  $\Omega^{-1}(\alpha) \in \mathbf{l}_{e(\alpha)}$ , the tangent line is  $\mathbf{l}_{e(\alpha)}$ .  $\square$

**Corollary 6.19.** *The map  $\Omega : \Sigma_+^\circ \rightarrow \mathcal{L}$  is orientation reversing.*

*Proof.* Choose  $\zeta_0 \in A_0 \setminus \boldsymbol{\alpha}$  and local coordinates  $\zeta = u + iv$  with  $v \geq 0$  on  $\Sigma_+$ . Since  $f_j$  is real,  $\Omega^{-1}(u + iv)$  given by (6.6) is an even function of  $v$ . Therefore, locally we have

$$(x(u + iv), y(u + iv)) = \Omega^{-1}(u + iv) = (x(u), y(u)) + \frac{1}{2}v^2(\partial_{vv}x(u), \partial_{vv}y(u)) + O(v^4).$$

Expanding

$$F(x(u + iv), y(u + iv); u + iv) = 0$$

in  $v$ , with  $F(x, y; \zeta)$  as in (6.9), and using the vanishing of the  $v^2$ -term, we obtain

$$(f_1(u), f_2(u)) \cdot (\partial_{vv}x(u), \partial_{vv}y(u)) = H \neq 0$$

where  $H$  is as in (6.14). In particular,

$$\partial_v \Omega^{-1}(u + iv) = v(\partial_{vv}x(u), \partial_{vv}y(u)) + O(v^3) \neq 0.$$

Differentiating with respect to  $u$  and using (6.15), we also get

$$\partial_u \Omega^{-1}(u + iv) = (\partial_u x(u), \partial_u y(u)) + O(v^2) = \frac{H}{W}(f_2(u), -f_1(u)) + O(v^2) \neq 0.$$

It follows that

$$\det(\partial_u \Omega^{-1}(u + iv), \partial_v \Omega^{-1}(u + iv)) = v \frac{H^2}{W} + O(v^3).$$

In particular,  $\Omega^{-1}$  is orientation reversing if  $W < 0$ .

We use the associated *amoeba* to show that  $W < 0$ . Define

$$X(\zeta) := \Re \left( \int_{\zeta_0}^{\zeta} d \log z \right), \quad Y(\zeta) := \Re \left( \int_{\zeta_0}^{\zeta} d \log w \right),$$

where  $\zeta_0$  is a fixed point on  $A_0$ . Since  $d \log z$  and  $d \log w$  are imaginary-normalized differentials,  $X$  and  $Y$  are well-defined single-valued functions on  $\Sigma$ . The amoeba is the image of the map  $\Sigma \ni \zeta \mapsto (X(\zeta), Y(\zeta)) \in \mathbb{R}^2$ . Under this map,  $\Sigma(\mathbb{R}) \setminus \alpha$  is mapped to the boundary of the amoeba, while  $\Sigma_+^\circ$  is mapped to its interior. In the algebraic (periodic) setting, this agrees with the classical amoeba of the spectral curve; see [85] for a proper introduction.

The connected components of the complement of the amoeba are convex [85, Proposition 2.7]. This convexity implies that, along the boundary, the tangent vector turns to the right. Consequently, for  $\zeta \in \Sigma(\mathbb{R})$ , the vector  $(-Y'(\zeta), X'(\zeta))$  is an inward-pointing normal, and

$$(-Y'(\zeta), X'(\zeta)) \cdot (X''(\zeta), Y''(\zeta)) < 0.$$

Locally we may write

$$X'(\zeta) = f_1(\zeta), \quad Y'(\zeta) = f_2(\zeta),$$

so that

$$(-Y'(\zeta), X'(\zeta)) \cdot (X''(\zeta), Y''(\zeta)) = (-f_2(\zeta), f_1(\zeta)) \cdot (f_1'(\zeta), f_2'(\zeta)) = W.$$

This proves that  $W < 0$  and, hence, that  $\Omega^{-1}$  is orientation reversing.  $\square$

**Proposition 6.20.** *If  $(x, y) = \Omega^{-1}(\zeta_0)$  is a smooth point of  $\mathcal{A}$ , then the curvature  $\kappa(x, y) < 0$ . In particular,  $\mathcal{L}$  is locally convex at  $(x, y)$ .*

*Proof.* Let  $\zeta_0 \in \Sigma(\mathbb{R}) \setminus \alpha$ . Choose a local coordinate  $\zeta = u + iv$  centered at  $\zeta_0$  such that  $\sigma(\zeta) = \bar{\zeta}$ , so that  $\Sigma_+$  is locally given by  $v \geq 0$ . Let  $W$  and  $H$  be as in (6.14). Using (6.15) and (6.17), we get that the curvature is

$$\kappa(x, y) = \frac{\det \begin{bmatrix} x' & y' \\ x'' & y'' \end{bmatrix}}{|(x', y')|^3} = \frac{|W|^3}{|H|(f_1^2 + f_2^2)^{3/2}} \frac{1}{W}.$$

Note that  $H \neq 0$  since  $(x, y)$  is assumed to be a smooth point and therefore, not a cusp. We saw in the proof of Corollary 6.19 that  $W < 0$ , which proves that  $\kappa(x, y) < 0$ .

Since  $\Omega^{-1}$  is orientation reversing, the boundary is traversed clockwise, and therefore, negative curvature means that  $\mathcal{A}$  bends towards the right, *i.e.* towards the rough region.

The argument for  $\zeta_0 = \alpha \in \alpha$  is similar with  $h_j(\zeta) = \zeta f_j(\zeta)$  in place of  $f_j(\zeta)$  as in Proposition 6.18.  $\square$

**Corollary 6.21.** *The map*

$$\Omega^{-1} : \Sigma(\mathbb{R}) \rightarrow \mathcal{A}$$

*is a bijection.*

*Proof.* Surjectivity follows from Proposition 6.17. It remains to show that  $\Omega^{-1}$  is injective.

By Proposition 6.18, every point of  $\mathcal{A}$  is locally either a smooth point of a branch of  $\mathcal{A}$  or an ordinary cusp. In particular,  $\Omega^{-1}|_{\Sigma(\mathbb{R})}$  is locally injective. Therefore, if  $\Omega^{-1}$  were not injective, then  $\mathcal{A}$  would have a self-intersection.

We first claim that no self-intersection can occur at a point of  $\partial\mathcal{D}_c$ . Indeed, there are no cusps on  $\partial\mathcal{D}_c$ , so such a point would have to be a smooth point of two distinct branches of  $\mathcal{A}$  with a common tangent line. Since  $\bar{\mathcal{L}}$  is locally convex by Proposition 6.20 and the map  $\Omega^{-1}|_{\Sigma_+^{\circ}}$  is a diffeomorphism, the two branches would have to lie on opposite sides of this common tangent line. This is impossible because  $\partial\mathcal{D}_c$  is simple.

Therefore, any self-intersection  $(x, y)$  of  $\mathcal{A}$  would have to lie in the interior  $\mathcal{D}_c^{\circ}$ . But then each branch through  $(x, y)$  imposes a different condition on the location of the two remaining zeroes of  $\lambda_c(x, y)$ , which is impossible.

Thus,

$$\Omega^{-1} : \Sigma(\mathbb{R}) \rightarrow \mathcal{A}$$

is globally injective, and therefore, a bijection. □

## 6.8 Zeroes of $\lambda_c(x, y)$ and tangent lines to the arctic curve

The zeroes of  $\lambda_c(x, y)$  admit a simple geometric interpretation in terms of tangent lines to the arctic curve. This gives a convenient dictionary between the phase classification from Definition 6.7 and the geometry of  $\mathcal{A}$ , generalizing the picture of Kenyon and Okounkov [83, Section 1.8]; see also [18, Proposition 15].

**Proposition 6.22.** *Let  $(x_0, y_0) \in \mathcal{A}$  and let*

$$\zeta_0 := \Omega(x_0, y_0).$$

*For  $(x, y) \in \mathbb{R}^2$ , the following are equivalent:*

1.  $(x, y)$  lies on the tangent line to  $\mathcal{A}$  at  $(x_0, y_0)$ .
2.  $\zeta_0$  is a zero of  $\lambda_c(x, y)$ .

*Proof.* By Proposition 6.18,  $(x, y)$  lies on the tangent line through  $(x_0, y_0)$  if and only if

$$(x - x_0)f_1(\zeta_0) + (y - y_0)f_2(\zeta_0) = 0. \tag{6.18}$$

Since  $\zeta_0 = \Omega(x_0, y_0)$ , we have

$$x_0f_1(\zeta_0) + y_0f_2(\zeta_0) + f_3(\zeta_0) = 0.$$

Thus, (6.18) is equivalent to

$$xf_1(\zeta_0) + yf_2(\zeta_0) + f_3(\zeta_0) = 0,$$

*i.e.* to  $\zeta_0$  being a zero of  $\lambda_c(x, y)$ . □

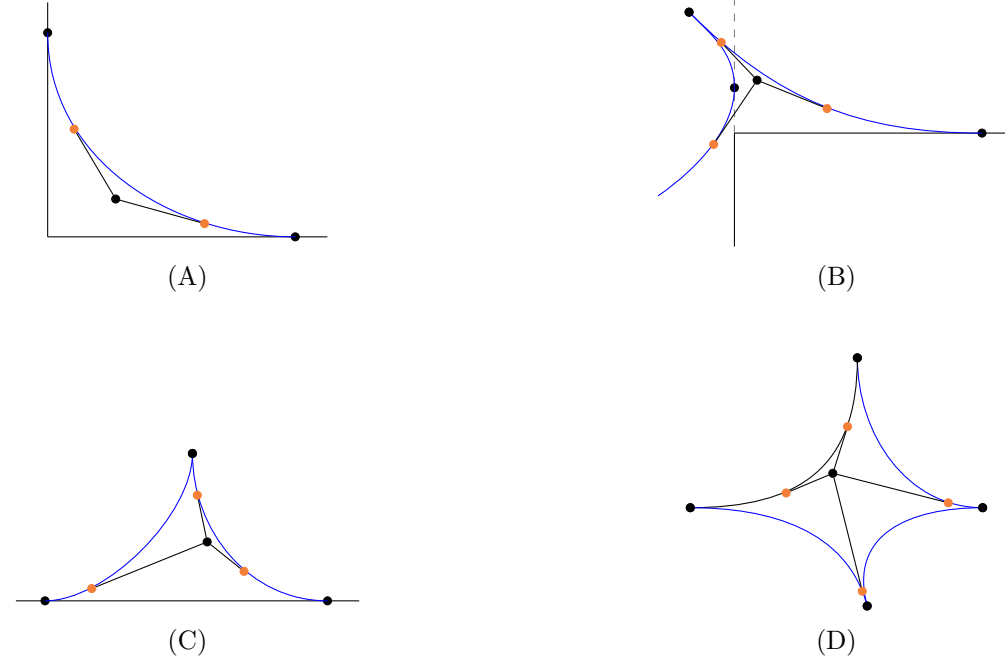


Figure 27: Local tangent configurations associated with the zeroes of  $\lambda_c(x, y)$ . (A)–(D) correspond respectively to the frozen region at a convex vertex, the frozen region at a concave vertex, the quasi-frozen region and the smooth region. In (B), the number of simple zeroes in  $(\alpha, \beta)$  is either two or three, depending on which side of the dashed line  $(x, y)$  lies; here the frozen region is  $\mathcal{F}_{(\alpha, \beta)}$ . Orange points denote zeroes of  $\lambda_c(x, y)$  under the identification  $\Omega^{-1}|_{\Sigma(\mathbb{R})}$ .

Proposition 6.22 translates the conditions on the zeroes of  $\lambda_c(x, y)$  in Definition 6.7 into statements about tangent lines to  $\mathcal{A}$ . The possible degenerations as we approach  $\mathcal{A}$  may be summarized as follows.

1. As  $(x, y)$  approaches the arctic curve from outside, two real tangents merge into a double tangent. When it crosses into the rough region, this double tangent splits into a  $\sigma$ -conjugate pair.
2. As  $(x, y)$  approaches a cusp of  $\mathcal{A}$ , three tangents merge into a triple tangent.
3. As  $(x, y)$  approaches a turning point corresponding to  $\alpha \in \alpha$  that is not a turning-point cusp, two tangents collide with the pole at  $\alpha$  and cancel, producing a simple zero.
4. At a turning-point cusp, a triple tangent collides with the pole, producing a double zero.

## 7 Asymptotics: local statistics and the limit shape

We now turn to the asymptotic analysis of the exact inverse Kasteleyn formula (4.3). The argument follows the general steepest-descent framework developed in [12]; see also [16], with the necessary modifications for our general setting.

## 7.1 Local statistics

Fix an astroidal domain  $\mathcal{D}_c$  and a sequence of AZ graphs  $\mathbf{G}_{\beta^n}$ ,  $n = 1, 2, \dots$ , such that

$$c_{\beta^n}(e) = n|e|_{\mathbb{Z}^2}c(e) + O(1), \quad e \in E(N),$$

see Sections 3.1 and 5.1 for definitions. Let  $\mathbf{C} \subset \mathbf{G}$  be a finite subgraph, and let  $(j_n, k_n) \in \mathbb{Z}^2$  satisfy

$$(j_n, k_n) = n(x, y) + O(1)$$

for some  $(x, y) \in \mathcal{D}_c^\circ$ . Set

$$\mathbf{C}_n := (j_n, k_n) + \mathbf{C}.$$

Recall the phase regions  $\mathcal{L}$ ,  $\mathcal{F}_{(\alpha, \beta)}$ ,  $\mathcal{Q}_{(\alpha, \beta)}$  and  $\mathcal{G}_j$  from Definition 6.7. We define  $A_{(x, y)}$  by

$$A_{(x, y)} := \begin{cases} (\alpha, \beta), & \text{if } (x, y) \text{ lies in a frozen or quasi-frozen region corresponding to } (\alpha, \beta) \subset A_0, \\ A_j, & \text{if } (x, y) \in \mathcal{G}_j, \end{cases}$$

where

$$\Sigma(\mathbb{R}) = A_0 \sqcup \left( \bigsqcup_{j=1}^g A_j \right)$$

as in Section 2.3.

**Definition 7.1.** We say that a point  $(x, y) \in \mathcal{D}_c^\circ$  lies in a *phase region* if it belongs to a rough, smooth, frozen, or quasi-frozen region. For such a point  $(x, y)$ , let

$$\zeta_* = \zeta_*(x, y) := \begin{cases} \Omega(x, y), & \text{if } (x, y) \in \mathcal{L}, \\ \text{any point of } A_{(x, y)}, & \text{otherwise.} \end{cases} \quad (7.1)$$

Convergence of local correlations is captured by the following statement.

**Theorem 7.2.** *Assume that the angle function is periodic. Let  $(x, y) \in \mathcal{D}_c^\circ$  lie in a phase region and let  $\zeta_*$  be as in Definition 7.1. If  $(x, y) \in \mathcal{F}_{(\alpha, \beta)}$ , assume also that*

$$(x, y) \notin \mathbf{l}_{e(\alpha)} \cup \mathbf{l}_{e(\beta)}.$$

For a black vertex  $\mathbf{b}$  and a white vertex  $\mathbf{w}$  of  $\mathbf{C}$ , let

$$\mathbf{b}_n := \mathbf{b} + (j_n, k_n), \quad \mathbf{w}_n := \mathbf{w} + (j_n, k_n)$$

denote their translates in  $\mathbf{C}_n$ . Then

$$\mathbf{K}_{\mathbf{b}_n, \mathbf{w}_n}^{-1} \rightarrow A_{\mathbf{b}, \mathbf{w}}^{\zeta_*} \quad \text{as } n \rightarrow \infty,$$

uniformly over  $\mathbf{b}$  and  $\mathbf{w}$ .

**Remark 7.3.** An equivalent way of stating Theorem 7.2 without formulas is that the local statistics are described by the Gibbs measure of the corresponding slope; see Corollary 7.12 below for a precise statement.

**Remark 7.4.** If  $(x, y) \in \mathcal{F}_{(\alpha, \beta)} \cap (\mathbf{I}_{e(\alpha)} \cup \mathbf{I}_{e(\beta)})$ , then we expect that  $(x, y)$  lies either in the frozen region or on a frozen-frozen boundary. The analysis of these points requires additional technical arguments which we do not pursue here (see Remark 7.8 below); we expect that it does not lead to interesting new phenomena.

The rest of the section is dedicated to the proof of Theorem 7.2. We begin by introducing notation of the relevant sign intervals. For  $(x, y) \in \mathcal{D}_c^\circ$ , let  $I_{(x,y),+}$  (resp.  $I_{(x,y),-}$ ) denote the union of the two + (resp. -) sign intervals for  $\text{sign}_R$  where  $R$  is the cell containing  $(x, y)$  as in Section 6.1. Let  $A_{R,+} \subset A_0$  (resp.  $A_{R,-} \subset A_0$ ) be the union of the two smallest intervals in  $A_0$  containing all angles  $\alpha_e$  for all  $e \in I_{(x,y),+}$  (resp.  $e \in I_{(x,y),-}$ ) and any  $(x, y) \in R$ . Note that if  $(x, y) \in \mathbf{I}_e$  for some edge  $e$ , then some signs vanish, and hence,  $I_{(x,y),+} \cup I_{(x,y),-} \neq \partial N$ .

Similarly, for  $\mathbf{x} \in B(\mathbf{G}_\beta) \sqcup W(\mathbf{G}_\beta)$ , let  $I_{\mathbf{x},+}$  (resp.  $I_{\mathbf{x},-}$ ) denote the union of the two + (resp. -) sign intervals for  $\text{sign}_{R_{\mathbf{x}}}$  as in Section 3.6.

**Lemma 7.5.** *Let  $R \subset \mathcal{D}_c^\circ$  be a cell, and let*

$$\mathbf{x}_n = \mathbf{x} + (j_n, k_n) \in B(\mathbf{C}_n) \sqcup W(\mathbf{C}_n), \quad \mathbf{x} \in B(\mathbf{C}) \sqcup W(\mathbf{C}),$$

*be a sequence such that  $(j_n, k_n) = n(x, y) + O(1)$  for some  $(x, y) \in R$ . Then, for any  $e \in E(N)$  such that  $\text{sign}_R(e) \neq 0$ , we have  $\text{sign}_{R_{\mathbf{x}_n}}(e) = \text{sign}_R(e)$  for all sufficiently large  $n$ . In particular,*

$$I_{(x,y),+} \subseteq I_{\mathbf{x}_n,+}, \quad I_{(x,y),-} \subseteq I_{\mathbf{x}_n,-}$$

*for all sufficiently large  $n$ .*

*Proof.* Since  $\mathbf{C}$  is a fixed finite subgraph,

$$(\mathbf{d}(\mathbf{x}) - \mathbf{d}(\mathbf{f}_0))_\alpha = O(1), \quad \text{as } n \rightarrow \infty,$$

for every  $\alpha \in \alpha$ . Hence,

$$(-\mathbf{d}(\mathbf{x}_n) - \mathbf{D}_{\beta^n})_\alpha = (-\mathbf{d}(\mathbf{f}_0 + (j_n, k_n)) - \mathbf{D}_{\beta^n})_\alpha - (\mathbf{d}(\mathbf{x}) - \mathbf{d}(\mathbf{f}_0))_\alpha.$$

Dividing by  $n$  and using Lemma 6.3, we obtain

$$\frac{1}{n}(-\mathbf{d}(\mathbf{x}_n) - \mathbf{D}_{\beta^n})_\alpha = -b_e x + a_e y + c(e) + O\left(\frac{1}{n}\right) \quad \text{for all } \alpha \in \alpha_e.$$

Therefore, recalling Lemma 6.4, for all  $e \in E(N)$  such that

$$\text{sign}_R(e) = \text{sign}(-b_e x + a_e y + c(e)) \neq 0$$

and all sufficiently large  $n$ ,

$$\text{sign}((-\mathbf{d}(\mathbf{x}_n) - \mathbf{D}_{\beta^n})_\alpha) = \text{sign}(-b_e x + a_e y + c(e)) = \text{sign}_R(e) \quad \text{for all } \alpha \in \alpha_e.$$

As in the proof of Lemma 2.12, this implies that  $\mathbf{x}_n$  lies in one of the strips on the corresponding side of  $\beta_e$ , and hence,

$$\text{sign}_{R_{\mathbf{x}_n}}(e) = \text{sign}_R(e)$$

for all sufficiently large  $n$ . The inclusions

$$I_{(x,y),+} \subseteq I_{\mathbf{x}_n,+}, \quad I_{(x,y),-} \subseteq I_{\mathbf{x}_n,-}$$

follow immediately. □

Fix  $(x, y) \in \mathcal{D}_c^\circ$ . We call a curve  $\gamma \subset \Sigma$  a *steepest descent* (resp. *ascent*) *curve* if:

- it passes through a zero  $\zeta_0$  of  $\lambda_c(x, y)$ ,
- for  $\zeta \in \gamma$ , we have

$$\Im \int_{\zeta_0}^{\zeta} \lambda_c(x, y) = 0,$$

where the integral is taken along  $\gamma$ ,

- the function

$$\gamma \ni \zeta \mapsto \Re \int_{\zeta_0}^{\zeta} \lambda_c(x, y)$$

is locally decreasing (resp. locally increasing) as  $\zeta$  leaves  $\zeta_0$  along  $\gamma$ .

By Lemma 2.6 a steepest descent/ascent curve is invariant (up to orientation) under the involution  $\sigma$ . We therefore first describe the portion of the curve lying in  $\Sigma_+$  and postpone specifying its orientation until Lemma 7.7 below.

We say that  $\zeta_0 \in \Sigma(\mathbb{R})$  is a *local minimum* (resp. *maximum*) if  $\zeta_0$  is a simple zero of  $\lambda_c(x, y)$  and a local minimum (resp. maximum) of the function

$$\Sigma(\mathbb{R}) \ni \zeta \mapsto \Re \int_{\zeta_0}^{\zeta} \lambda_c(x, y).$$

Such a point is a saddle for  $\zeta \mapsto \Re \int_{\zeta_0}^{\zeta} \lambda_c(x, y)$ : along the curve going perpendicular to  $\Sigma(\mathbb{R})$ , it is a local maximum (resp. minimum). Our convention is that the term local minimum (resp. maximum) refers to the behavior along  $\Sigma(\mathbb{R})$ .

**Lemma 7.6.** *Let  $(x, y) \in R \subset \mathcal{D}_c^\circ$  and assume that  $(x, y)$  lies in a phase region.*

1. *If  $(x, y)$  is in the frozen, quasi-frozen, or smooth region, and  $\zeta_0 \in A_{(x,y)}$  is a local minimum (resp. maximum), there exists a simple curve  $C_{\zeta_0, \text{desc}}^+ \subset \Sigma_+$  (resp.  $C_{\zeta_0, \text{asc}}^+ \subset \Sigma_+$ ) going between  $\zeta_0$  and  $A_{R,+}$  (resp.  $A_{R,-}$ ), such that  $\Im \int_{\zeta_0}^{\zeta} \lambda_c(x, y) = 0$  and  $\Re \int_{\zeta_0}^{\zeta} \lambda_c(x, y)$  is decreasing (resp. increasing) along the curve.*
2. *If  $(x, y) \in \mathcal{L}$  and  $\zeta_0 = \Omega(x, y) \in \Sigma_+^\circ$ , there exists a simple curve  $C_{\zeta_0, \text{desc}}^+ \subset \Sigma_+$  (resp.  $C_{\zeta_0, \text{asc}}^+ \subset \Sigma_+$ ) with both endpoints in  $A_{R,+}$  (resp.  $A_{R,-}$ ), passing through  $\zeta_0$ , such that  $\Im \int_{\zeta_0}^{\zeta} \lambda_c(x, y) = 0$  and  $\Re \int_{\zeta_0}^{\zeta} \lambda_c(x, y)$  is decreasing (resp. increasing) along the curve as we move away from  $\zeta_0$ .*

Moreover, if  $\zeta_0 \in A_{(x,y)}$  is a local minimum and  $\zeta'_0 \in A_{(x,y)}$  is a local maximum, then

$$C_{\zeta_0, \text{desc}}^+ \cap C_{\zeta'_0, \text{asc}}^+ = \emptyset.$$

If instead  $(x, y) \in \mathcal{L}$  and  $\zeta_0 = \Omega(x, y) \in \Sigma_+^\circ$ , then

$$C_{\zeta_0, \text{desc}}^+ \cap C_{\zeta_0, \text{asc}}^+ = \{\zeta_0\}.$$

See Figure 28 below for examples of the curves.

*Proof.* We first consider the case  $\zeta_0 \in \Sigma(\mathbb{R})$ , that is,  $(x, y)$  is in the frozen, quasi-frozen, or smooth region. Since  $(x, y) \notin \mathcal{A}$ , the zero  $\zeta_0$  of  $\lambda_c(x, y)$  is simple. Hence, there is one level line of

$$\Im \int_{\zeta_0}^{\zeta} \lambda_c(x, y) = 0$$

along  $\Sigma(\mathbb{R})$  and one level line entering  $\Sigma_+^\circ$ . If  $\zeta_0$  is a local minimum (resp. maximum), the function

$$\zeta \mapsto \Re \int_{\zeta_0}^{\zeta} \lambda_c(x, y) \tag{7.2}$$

is decreasing (resp. increasing) along the level line entering  $\Sigma_+^\circ$ . Following that level line, we eventually reach either another zero of  $\lambda_c(x, y)$  which has to be in  $\Sigma(\mathbb{R})$ , since  $(x, y) \notin \mathcal{L}$ , and is a local maximum (resp. minimum) (recall that the maximum (resp. minimum) refers to the behavior along  $\Sigma(\mathbb{R})$ ) or a point where (7.2) tends to  $-\infty$  (resp.  $+\infty$ ). In the latter case, the endpoint must be an angle  $\alpha \in \alpha_e$  with  $e \in I_{R,+}$  (resp.  $e \in I_{R,-}$ ), which means that  $\alpha \in A_{R,+}$  (resp.  $\alpha \in A_{R,-}$ ).

We define the curve  $C_{\zeta_0, \text{desc}}^+ \subset \Sigma_+$  (resp.  $C_{\zeta_0, \text{asc}}^+ \subset \Sigma_+$ ) to be this level line, extended if necessary until it hits  $A_{R,+}$  (resp.  $A_{R,-}$ ); see Figure 28(B)–(D) for an illustration. If it first hits another zero  $\zeta'_0$  of  $\lambda_c(x, y)$ , then either:

1.  $\zeta'_0 \in A_{R,+}$  (resp.  $\zeta'_0 \in A_{R,-}$ ), and we are done.
2.  $\zeta'_0 \in A_{(x,y)}$ , which is impossible. Indeed,  $\zeta_0$  would be a local maximum (resp. minimum) and since the types of extrema alternate along  $A_{(x,y)}$ ,  $\zeta_0$  and  $\zeta'_0$  would be joined by a steepest ascent (resp. descent) curve along  $\Sigma(\mathbb{R})$  with no zero between them. This is a contradiction, since they are also joined by a steepest descent (resp. ascent) curve in  $\Sigma_+^\circ$ .
3.  $\zeta'_0 \in A_j \neq A_{(x,y)}$ , for some  $j = 1, \dots, g$ , in which case we continue the curve along  $A_j$  in either direction until it hits the other zero of  $\lambda_c(x, y)$  in  $A_j$  which is a local minimum (resp. maximum), and then continue again into  $\Sigma_+^\circ$ .

Since a curve constructed by this procedure cannot return to the same  $A_j$  twice, it eventually has to end up at an angle  $\alpha \in A_{R,+}$  (resp.  $\alpha \in A_{R,-}$ ) or a local maximum (resp. minimum) which lies in  $A_{R,+}$  (resp.  $A_{R,-}$ ).

The resulting curve is simple because the function (7.2) is monotone along it (recall that this function is well-defined since  $\lambda_c(x, y)$  is imaginary-normalized). If  $\zeta_0 \in A_{(x,y)}$  is a local minimum and  $\zeta'_0 \in A_{(x,y)}$  is a local maximum, there cannot be another local minimum or maximum in  $A_{(x,y)}$  between  $\zeta_0$  and  $\zeta'_0$ . Thus, the same monotonicity argument shows that

$$C_{\zeta_0, \text{desc}}^+ \cap C_{\zeta'_0, \text{asc}}^+ = \emptyset.$$

Now we consider the case  $\zeta_0 \in \Sigma_+^\circ$ , *i.e.*  $(x, y) \in \mathcal{L}$ . The local level set  $\Im \int_{\zeta_0}^{\zeta} \lambda_c(x, y) = 0$  consists of two curves going through  $\zeta_0$  orthogonally. One of those curves is used to define  $C_{\zeta_0, \text{desc}}^+ \subset \Sigma_+$  and the other is used to define  $C_{\zeta_0, \text{asc}}^+ \subset \Sigma_+$ . To construct  $C_{\zeta_0, \text{desc}}^+ \subset \Sigma_+$  (resp.  $C_{\zeta_0, \text{asc}}^+ \subset \Sigma_+$ ), we follow the level line for which  $\zeta \mapsto \Re \int_{\zeta_0}^{\zeta} \lambda_c(x, y)$  is decreasing (resp. increasing); see Figure 28(A) for an illustration. The same argument as in the first case now proves the statement for this second case.  $\square$

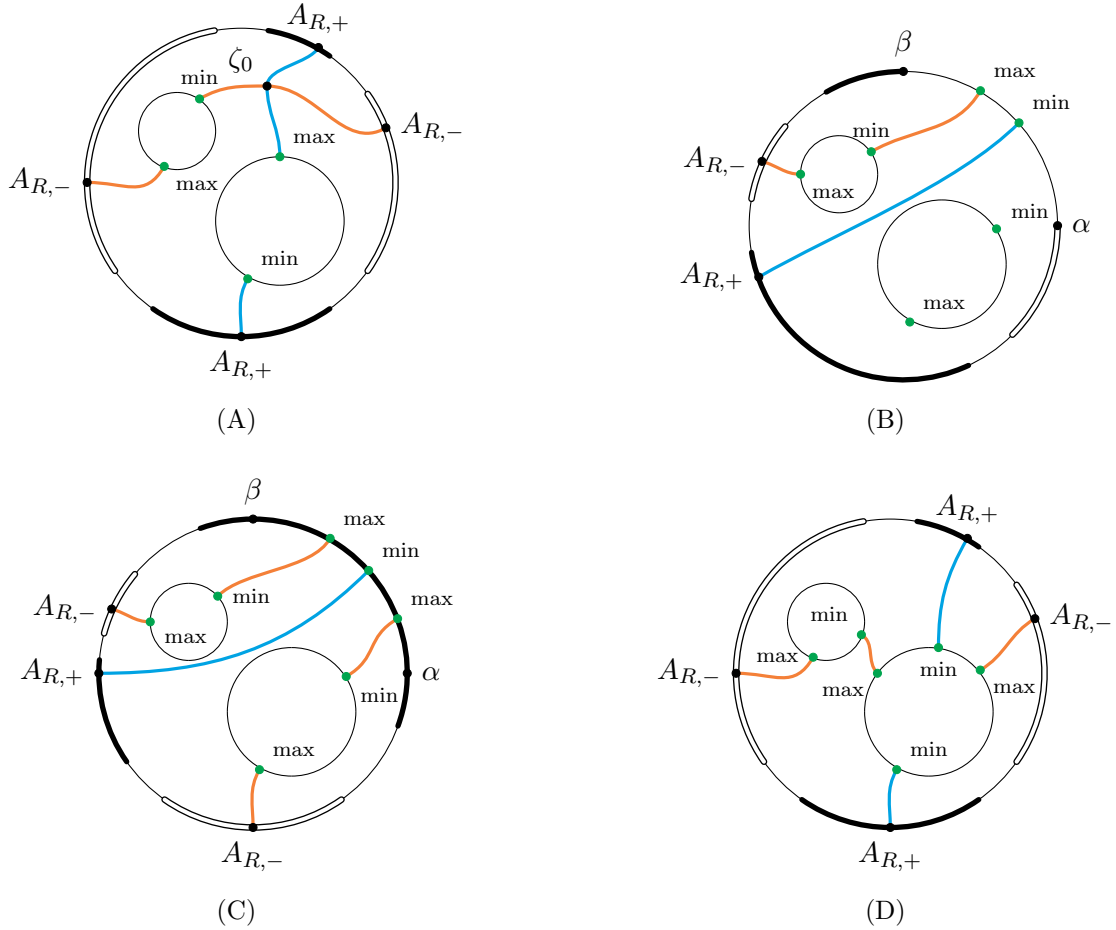


Figure 28: Steepest descent (blue) and steepest ascent (orange) curves in  $\Sigma_+$  for a genus-2 example. The outer circle is  $A_0$ , while the inner circles are  $A_1$  and  $A_2$ . Green points indicate zeros of  $\lambda_c(x, y)$ . (A)–(D) correspond respectively to  $(x, y)$  lying in the rough region, frozen region  $\mathcal{F}_{(\alpha, \beta)}$  such that  $\lambda_c(x, y)$  has two simple zeroes in  $(\alpha, \beta)$ , quasi-frozen region (when  $e(\alpha) = e(\beta)$ ) or frozen region  $\mathcal{F}_{(\alpha, \beta)}$  such that  $\lambda_c(x, y)$  has three simple zeroes in  $(\alpha, \beta)$  (when  $e(\alpha) \neq e(\beta)$ ), and smooth region. In keeping with our sign/color convention, the black and white intervals in  $A_0$  correspond to  $A_{R,+}$  and  $A_{R,-}$  respectively.

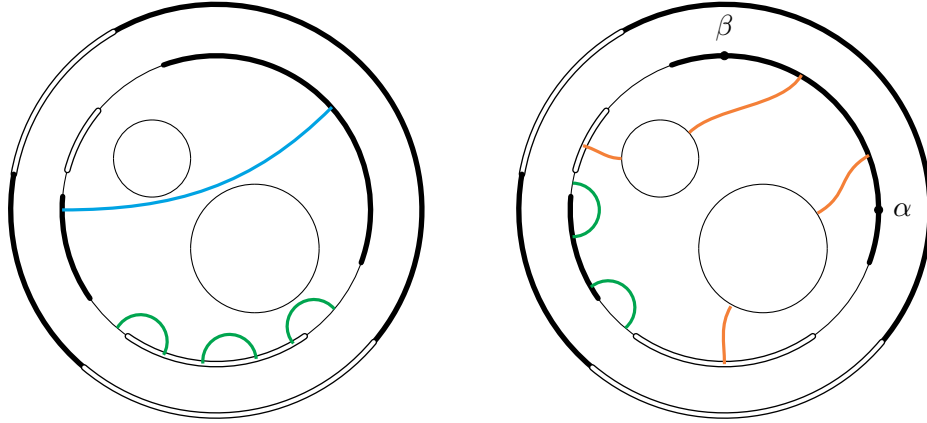


Figure 29: On the left-hand side, the union of the two black (resp. white) intervals in the outer circle corresponds to the angles in  $I_{b_n,+}$  (resp.  $I_{b_n,-}$ ), while union of the two black (resp. white) intervals in  $A_0$  corresponds to  $A_{R,+}$  (resp.  $A_{R,-}$ ) as in Figure 28. The steepest descent curve (blue) in Figure 28(C) can be deformed to the curve  $C(I_{b_n})$  (green) without crossing any angles in  $I_{b_n,-}$ . Similarly, the right-hand side shows that the steepest ascent curve (orange) can be deformed to  $C(I_{w_n})$  (green) without crossing any angles in  $I_{w_n,+}$  (note that  $\alpha$  and  $\beta$  are consecutive angles).

We now extend the curves defined in Lemma 7.6 to  $\Sigma$ . Define  $C_{\zeta_0,\text{desc}} \subset \Sigma$  (resp.  $C_{\zeta_0,\text{asc}} \subset \Sigma$ ) to be the  $\sigma$ -invariant extension of  $C_{\zeta_0,\text{desc}}^+ \subset \Sigma_+$  (resp.  $C_{\zeta_0,\text{asc}}^+ \subset \Sigma_+$ ) such that

$$\sigma_*(C_{\zeta_0,\text{desc}}) = -C_{\zeta_0,\text{desc}} \quad (\text{resp. } \sigma_*(C_{\zeta_0,\text{asc}}) = -C_{\zeta_0,\text{asc}}).$$

Note that we have not specified the orientation of the curves yet, however, we assume they are given some orientation. Any segment of the curve along  $\Sigma(\mathbb{R})$  is traversed in both directions, so its contribution to the integral cancels. By Lemma 7.6, the curve  $C_{\zeta_0,\text{desc}}$  (resp.  $C_{\zeta_0,\text{asc}}$ ) is a curve of steepest descent (resp. ascent) and intersects  $A_0$  in at most two points.

We then define  $C_{(x,y),\text{desc}}$  and  $C_{(x,y),\text{asc}}$  by

$$C_{(x,y),\text{desc}} := \begin{cases} C_{\zeta_0,\text{desc}}, & \text{if } \zeta_0 = \Omega(x,y) \in \Sigma_+^\circ, \\ \bigsqcup_{\zeta_0 \text{ local minimum in } A_{(x,y)}} C_{\zeta_0,\text{desc}}, & \text{otherwise,} \end{cases}$$

and

$$C_{(x,y),\text{asc}} := \begin{cases} C_{\zeta_0,\text{asc}}, & \text{if } \zeta_0 = \Omega(x,y) \in \Sigma_+^\circ, \\ \bigsqcup_{\zeta_0 \text{ local maximum in } A_{(x,y)}} C_{\zeta_0,\text{asc}}, & \text{otherwise.} \end{cases}$$

The following lemma shows that the contours appearing in the inverse Kasteleyn formula (4.3) may be deformed to the steepest descent and steepest ascent curves respectively without crossing the poles of the integrand as specified in the following claim (see Figure 29).

**Lemma 7.7.** *Let  $(x,y) \in \mathcal{D}_c^\circ$  lie in a phase region, and let  $\zeta_0 = \zeta_*(x,y)$ . If  $(x,y) \in \mathcal{F}_{(\alpha,\beta)}$ , where  $(\alpha,\beta) \subset A_0$  is the interval between consecutive angles  $\alpha, \beta$ , assume that  $(x,y) \notin \mathbf{l}_{e(\alpha)} \cup \mathbf{l}_{e(\beta)}$ .*

Then we may choose orientations of  $C_{(x,y),\text{desc}}$  and  $C_{(x,y),\text{asc}}$  such that, for all sufficiently large  $n$  and all  $\mathbf{b}_n \in B(\mathbf{C}_n), \mathbf{w}_n \in W(\mathbf{C}_n)$ , the curve  $C_{(x,y),\text{desc}}$  is homologous to  $C(I_{\mathbf{b}_n})$  in

$$\Sigma \setminus \bigcup_{e \in I_{\mathbf{b}_n, -}} \alpha_e,$$

and the curve  $C_{(x,y),\text{asc}}$  is homologous to  $C(I_{\mathbf{w}_n})$  in

$$\Sigma \setminus \bigcup_{e \in I_{\mathbf{w}_n, +}} \alpha_e.$$

*Proof.* We distinguish cases according to the location of  $(x, y)$ .

1. Assume first that  $(x, y) \in \mathcal{G}_j$  for some  $j = 1, \dots, g$ , so that  $A_{(x,y)} = A_j$ . The curves  $C_{(x,y),\text{desc}}$  and  $C_{(x,y),\text{asc}}$  divide  $\Sigma$  into four parts. The four zeros of  $\lambda_c(x, y)$  on  $A_j$  alternate between local minima and local maxima. By Lemma 7.6,  $C_{(x,y),\text{desc}}$  and  $C_{(x,y),\text{asc}}$  do not intersect, and so the same lemma implies that  $C_{(x,y),\text{desc}}$  intersects  $A_0$  once in each component of  $A_{R,+}$ , and  $C_{(x,y),\text{asc}}$  intersects  $A_0$  once in each component of  $A_{R,-}$ .
2. If  $(x, y) \in \mathcal{L}$ , so that  $\Omega(x, y) \in \Sigma_+^\circ$ , a similar argument shows that  $C_{(x,y),\text{desc}}$  intersects  $A_0$  once in each component of  $A_{R,+}$ , and  $C_{(x,y),\text{asc}}$  intersects  $A_0$  once in each component of  $A_{R,-}$ .
3. Finally, if  $(x, y) \in \mathcal{F}_{(\alpha,\beta)}$  or  $(x, y) \in \mathcal{Q}_{(\alpha,\beta)}$  for some  $(\alpha, \beta) \subset A_0$ , there are two possibilities:
  - (a) If  $A_{(x,y)} \subset A_{R,+}$  (resp.  $A_{(x,y)} \subset A_{R,-}$ ), then there is one local minimum (resp. local maximum) and two local maxima (resp. local minima) in  $A_{(x,y)}$ , and the types of extrema alternate along  $A_{(x,y)}$ . By the same argument as above, the curve  $C_{(x,y),\text{desc}}$  (resp.  $C_{(x,y),\text{asc}}$ ) goes between  $A_{(x,y)}$  and the opposite component of  $A_{R,+}$  (resp.  $A_{R,-}$ ), and  $C_{(x,y),\text{asc}}$  (resp.  $C_{(x,y),\text{desc}}$ ) consists of two components, going between the two components of  $A_{R,-}$  (resp.  $A_{R,+}$ ) and  $A_{(x,y)}$ .
  - (b) If  $A_{(x,y)}$  is contained in neither  $A_{R,+}$  nor  $A_{R,-}$ , then there is one local minimum and one local maximum in  $A_{(x,y)}$ . By the same argument as above,  $C_{(x,y),\text{desc}}$  (resp.  $C_{(x,y),\text{asc}}$ ) joins  $A_{(x,y)}$  and the opposite component of  $A_{R,+}$  (resp.  $A_{R,-}$ ).

In each case, the description of the curves given above together with Lemma 7.5 implies that, if we choose the orientation correctly,  $C_{(x,y),\text{desc}}$  (resp.  $C_{(x,y),\text{asc}}$ ) can be deformed to  $C(I_{\mathbf{b}_n})$  (resp.  $C(I_{\mathbf{w}_n})$ ) in  $\Sigma \setminus \bigcup_{e \in I_{\mathbf{b}_n, -}} \alpha_e$  (resp.  $\Sigma \setminus \bigcup_{e \in I_{\mathbf{w}_n, +}} \alpha_e$ ); see Figure 28. □

**Remark 7.8.** Note that if  $(x, y)$  lies in a frozen region  $\mathcal{F}_{(\alpha,\beta)}$  at a concave vertex as in Figure 27(B), the steepest descent and ascent curves are fundamentally different depending on which side of the dashed line the point  $(x, y)$  lies on; see Figure 28(B)–(C). When  $(x, y) \in \mathcal{F}_{(\alpha,\beta)} \cap (I_{e(\alpha)} \cup I_{e(\beta)})$ , the analysis therefore becomes sensitive to the precise locations of  $\mathbf{b}_n$  and  $\mathbf{w}_n$ . This is the reason that we exclude this case in our analysis.

*Proof of Theorem 7.2.* We use the convention in Remark 4.2, so that

$$\mathbf{K}_{\mathbf{b}_n, \mathbf{w}_n}^{-1} = \iint_{C(I_{\mathbf{w}_n}) \prec C(I_{\mathbf{b}_n})} \omega_{\mathbf{b}_n, \mathbf{w}_n} - \langle \partial I_{\mathbf{b}_n} \cap I_{\mathbf{w}_n}, \mathbf{A}_{\mathbf{b}_n, \mathbf{w}_n} \rangle. \quad (7.3)$$

Let  $I_{\mathbf{b}_n} = [\ell_{\mathbf{b}_n}, r_{\mathbf{b}_n}]$ . Then, for sufficiently large  $n$ ,

$$-\langle \partial I_{\mathbf{b}_n} \cap I_{W_n}, \mathbf{A}_{\mathbf{b}_n, W_n} \rangle = \mathbf{1}_{\{\ell_{\mathbf{b}_n} \in I_{W_n}\}} \mathbf{A}_{\mathbf{b}_n, W_n}^{\ell_{\mathbf{b}_n}} - \mathbf{1}_{\{r_{\mathbf{b}_n} \in I_{W_n}\}} \mathbf{A}_{\mathbf{b}_n, W_n}^{r_{\mathbf{b}_n}} = \mathbf{A}_{\mathbf{b}_n, W_n}^{\ell_{\mathbf{b}_n}}.$$

Indeed, by the choice of  $I_{W_n}$  and  $I_{\mathbf{b}_n}$  in Remark 4.2, and by Lemma 7.5,  $r_{\mathbf{b}_n} \notin I_{W_n}$ . If  $\ell_{\mathbf{b}_n} \notin I_{W_n}$ ,  $\text{sign}_{R_{\mathbf{b}_n}}(e) = +$  and  $\text{sign}_{R_{W_n}}(e) = -$ , where  $e$  is the edge immediately to the left of  $\ell_{\mathbf{b}_n}$ , so  $\ell_{\mathbf{b}_n} \in \mathfrak{N}_{\mathbf{b}_n, W_n}$  by Lemma 4.8, and, hence,  $\mathbf{A}_{\mathbf{b}_n, W_n}^{\ell_{\mathbf{b}_n}} = 0$ .

By Lemma 7.7, we may deform  $C(I_{W_n})$  to  $C_{(x,y), \text{asc}}$  and  $C(I_{\mathbf{b}_n})$  to  $C_{(x,y), \text{desc}}$  in the double contour integral in (7.3) without crossing any poles of  $\omega_{\mathbf{b}_n, W_n}$  other than the diagonal  $\zeta = \eta$ . Thus, the contribution of this deformation is precisely the diagonal residue along the curve  $C_{\ell_{\mathbf{b}_n}}^{\zeta^*}$  defined in Definition 2.15. Therefore,

$$\mathbf{K}_{\mathbf{b}_n, W_n}^{-1} = \int_{C_{(x,y), \text{asc}}} \int_{C_{(x,y), \text{desc}}} \omega_{\mathbf{b}_n, W_n} + \frac{1}{2\pi i} \int_{C_{\ell_{\mathbf{b}_n}}^{\zeta^*}} g_{\mathbf{b}_n, W_n} + \mathbf{A}_{\mathbf{b}_n, W_n}^{\ell_{\mathbf{b}_n}}.$$

Since

$$\frac{1}{2\pi i} \int_{C_{\ell_{\mathbf{b}_n}}^{\zeta^*}} g_{\mathbf{b}_n, W_n} + \mathbf{A}_{\mathbf{b}_n, W_n}^{\ell_{\mathbf{b}_n}} = \mathbf{A}_{\mathbf{b}_n, W_n}^{\zeta^*} = \mathbf{A}_{\mathbf{b}, W}^{\zeta^*},$$

it remains to prove that the double contour integral tends to zero as  $n \rightarrow \infty$ .

Recall the action function  $\Lambda_c(\cdot; x, y)$  from Definition 6.2. Expanding all the factors, we write the kernel (4.1) as

$$\omega_{\mathbf{b}_n, W_n} = \frac{1}{(2\pi i)^2} \frac{\theta(-t + \mathbf{D}_{\beta^n} + \eta - \zeta) \theta(t + \eta + \mathbf{d}(w_n)) \theta(-t + \zeta - \mathbf{d}(b_n))}{E(\zeta, \eta) \theta(-t + \mathbf{D}_{\beta^n})} \frac{\prod_{\alpha \in \alpha} E(\alpha, \eta)^{(\mathbf{d}(w_n) + \mathbf{D}_{\beta^n})_\alpha}}{\prod_{\alpha \in \alpha} E(\alpha, \zeta)^{(\mathbf{d}(b_n) + \mathbf{D}_{\beta^n})_\alpha}}. \quad (7.4)$$

By (6.3), the prime-form factor in (7.4) is

$$\exp\left(n(\Lambda_c(\eta; x, y) - \Lambda_c(\zeta; x, y)) + O(1)\right).$$

Since  $d\Lambda_c(\cdot; x, y) = \lambda_c(x, y)$  and  $\lambda_c(x, y)$  is imaginary-normalized, the real part  $\Re\Lambda_c(\cdot; x, y)$  is well-defined on  $\Sigma$ , and

$$\Re\Lambda_c(\zeta; x, y) - \Re\Lambda_c(\eta; x, y) = \Re \int_{\zeta_0}^{\zeta} \lambda_c(x, y) - \Re \int_{\zeta_0}^{\eta} \lambda_c(x, y).$$

It follows that the curves  $C_{(x,y), \text{asc}}$  and  $C_{(x,y), \text{desc}}$  are steepest ascent and descent contours for the action function  $\Lambda_c(\cdot; x, y)$ . Standard arguments (that go back to [94]), using that the first factor in (7.4) is bounded in  $n$  since the ovals are compact, show that

$$\int_{C_{(x,y), \text{asc}}} \int_{C_{(x,y), \text{desc}}} \omega_{\mathbf{b}_n, W_n} \rightarrow 0$$

as  $n \rightarrow \infty$ , which proves the theorem.  $\square$

## 7.2 The limit shape

Let  $h_M$  be the height function corresponding to a dimer cover  $M$  defined on  $\mathbf{G}_{\beta^n}$  with the extremal dimer cover  $M_0 := M_{u_0}$  corresponding to  $u_0 \in V(N)$  as reference dimer cover. We choose the reference face  $f_0 \in F(\mathbf{G}) \setminus F(\mathbf{G}_{\beta^n})$  to be adjacent to  $e_{u_0} \in E(\mathbf{G}_{\beta^n})$ , where  $e_{u_0}$  is the intersection of two boundary zig-zag paths as in Section 3.1. The limit shape is the appropriate scaled large  $n$  limit of the height function: Let  $(j_n, k_n) = n(x, y) + O(1)$  for some  $(x, y) \in \mathcal{D}_c^\circ$  and set  $f_n = f_0 + (j_n, k_n) \in F(\mathbf{G}_{\beta^n})$ . The limit shape  $\mathfrak{h}$  is defined as

$$\mathfrak{h}(x, y) := \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{n} h_M(f_n) \right].$$

For dimer models with periodic edge weights, it is well known that the limit shape exists, is the solution to a variational problem, and that  $\frac{1}{n} h_M \rightarrow \mathfrak{h}$  in probability as  $n \rightarrow \infty$  [38, 64, 84, 86]. This is also expected to be true for Fock's weights with periodic angles [19].

We now give an integral formula for the limit shape.

**Theorem 7.9.** *Assume that the angle function is periodic. For  $(x, y) \in \mathcal{D}_c^\circ$ ,*

$$\mathfrak{h}(x, y) = -\frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_*}} \lambda_c(x, y), \quad (7.5)$$

where  $C_{u_0}^{\zeta_*}$  is the curve going from  $\sigma(\zeta_*)$  to  $\zeta_*$  crossing  $A_0$  only once in the component of  $A_0 \setminus \alpha$  corresponding to  $u_0$  as in Definition 2.15. The point  $\zeta_* \in \Sigma_+$  is defined as in Definition 7.1 if  $(x, y)$  lies in a phase region and defined by continuity otherwise.

The proof starts by rewriting the expectation of the height function in a form that is suitable for steepest descent analysis, similar to the one carried out in Section 7.1. The key identity that allows for this simplification in our setting is given in the following proposition. This identity was proved in [13] in the study of perfect t-embeddings of Aztec diamonds with Fock's weights. For the convenience of the reader, we include the proof from [13] in Appendix B.

**Proposition 7.10** ([13]). *Let  $f_+, f_- \in F(\mathbf{G}_{\beta^n})$  be adjacent faces with  $\mathbf{wb} \in E(\mathbf{G}_{\beta^n})$  as their common adjacent edge, and let  $\alpha$  and  $\beta$  be the strands containing  $\mathbf{wb}$ , as in Figure 8. Then*

$$E(\zeta, \eta) \frac{g_{\mathbf{w}}(\eta)}{g_{\mathbf{b}}(\zeta)} \mathbf{K}_{\mathbf{w}, \mathbf{b}} = G_{f_+}(\zeta, \eta) - G_{f_-}(\zeta, \eta),$$

where

$$G_f(\zeta, \eta) := \frac{\theta(-t - \mathbf{d}(f) + \zeta - \eta) g_f(\eta)}{\theta(t + \mathbf{d}(f)) g_f(\zeta)},$$

for all faces  $f \in F(\mathbf{G}_{\beta^n})$  and all  $\zeta, \eta \in \Sigma$ .

*Proof of Theorem 7.9.* We will consider points  $(x, y) \in \mathcal{D}_c^\circ$  satisfying the following assumption.

**Assumption 7.11.** *Let  $(x, y) \in \mathcal{D}_c^\circ$  be a point in a phase region. If  $(x, y) \in \mathcal{F}_{(\alpha, \beta)}$ , assume also that  $(x, y) \notin \mathbf{l}_{e(\alpha)} \cup \mathbf{l}_{e(\beta)}$ .*

In particular, we will prove the limit

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{n} h_M(\mathbf{f}_n) \right] = -\frac{1}{2\pi i} \int_{C_{u_0}^{\zeta^*}} \lambda_c(x, y),$$

for  $(x, y)$  satisfying Assumption 7.11. Since  $\mathfrak{h}$  is a continuous function—indeed, Lipschitz continuous—the statement follows.

By definition of the height function; see Section 2.2.1, and by Theorems 2.2 and 2.18, we have

$$\mathbb{E}[h_M(\mathbf{f}_n)] = \sum_{\mathbf{e} \in E(\mathbf{G}_{\beta^n})} (\mathbf{e} \wedge \gamma_n^*) \mathbb{E}[\mathbf{1}_{\mathbf{e} \in M} - \mathbf{1}_{\mathbf{e} \in M_0}] = \sum_{\mathbf{e} = \mathbf{b}\mathbf{w} \in E(\mathbf{G}_{\beta^n})} (\mathbf{e} \wedge \gamma_n^*) \left( \mathbf{K}_{\mathbf{b},\mathbf{w}}^{-1} \mathbf{K}_{\mathbf{w},\mathbf{b}} - \mathbf{A}_{\mathbf{b},\mathbf{w}}^{u_0} \mathbf{K}_{\mathbf{w},\mathbf{b}} \right), \quad (7.6)$$

where  $\gamma_n^*$  is a dual path in  $\mathbf{G}_{\beta^n}^*$  from  $\mathbf{f}_0$  to  $\mathbf{f}_n$ , and  $\mathbf{A}_{\mathbf{b},\mathbf{w}}^{u_0}$  is the whole-plane inverse Kasteleyn matrix defining the Gibbs measure associated with  $u_0$ .

We will first compute the limit of the expected height difference between two faces in the same chamber. We use the convention for  $\mathbf{K}_{\mathbf{b},\mathbf{w}}^{-1}$  from Remark 4.2. Let  $\mathbf{e} = \mathbf{b}\mathbf{w}$  and let  $I_{\mathbf{b}} = [\ell_{\mathbf{b}}, r_{\mathbf{b}}]$ . By Proposition 7.10 and definition of  $\mathbf{A}_{\mathbf{b},\mathbf{w}}^{\zeta}$ ,

$$\begin{aligned} \mathbb{E}[\mathbf{1}_{\mathbf{e} \in M} - \mathbf{1}_{\mathbf{e} \in M_0}] &= \iint_{C(I_{\mathbf{w}}) \prec C(I_{\mathbf{b}})} \omega_{\mathbf{b},\mathbf{w}} \mathbf{K}_{\mathbf{w},\mathbf{b}} + \left( \mathbf{A}_{\mathbf{b},\mathbf{w}}^{\ell_{\mathbf{b}}} - \mathbf{A}_{\mathbf{b},\mathbf{w}}^{u_0} \right) \mathbf{K}_{\mathbf{w},\mathbf{b}} = \frac{1}{2\pi i} \int_{C([\ell_{\mathbf{b}}, u_0])} g_{\mathbf{b},\mathbf{w}} \mathbf{K}_{\mathbf{w},\mathbf{b}} \\ &\quad + \frac{1}{(2\pi i)^2} \iint_{C(I_{\mathbf{w}}) \prec C(I_{\mathbf{b}})} \frac{\theta(-t + \mathbf{D}_{\beta^n} + \eta - \zeta)}{\theta(-t + \mathbf{D}_{\beta^n}) E(\zeta, \eta)^2} (G_{\mathbf{f}_+}(\zeta, \eta) - G_{\mathbf{f}_-}(\zeta, \eta)) \frac{\Psi_{\beta}(\eta)}{\Psi_{\beta}(\zeta)}, \end{aligned}$$

where  $\mathbf{f}_-, \mathbf{f}_+ \in F(\mathbf{G}_{\beta^n})$  are as in Figure 8.

More generally, let  $\mathbf{f}_-, \mathbf{f}_+ \in F(\mathbf{G}_{\beta^n})$  be any two faces in the same chamber  $\mathbf{R}$ , and let  $\gamma^*$  be a dual path in  $\mathbf{R}$  going from  $\mathbf{f}_-$  to  $\mathbf{f}_+$  keeping white vertices to the left. Then

$$\begin{aligned} \mathbb{E}[h_M(\mathbf{f}_+) - h_M(\mathbf{f}_-)] &= \sum_{\mathbf{e} \in E(\mathbf{R})} (\mathbf{e} \wedge \gamma^*) \mathbb{E}[\mathbf{1}_{\mathbf{e} \in M} - \mathbf{1}_{\mathbf{e} \in M_0}] = \sum_{\mathbf{e} = \mathbf{b}\mathbf{w} \in E(\mathbf{R})} \frac{(\mathbf{e} \wedge \gamma^*)}{2\pi i} \int_{C([\ell_{\mathbf{b}_R}, u_0])} g_{\mathbf{b},\mathbf{w}} \mathbf{K}_{\mathbf{w},\mathbf{b}} \\ &\quad + \frac{1}{(2\pi i)^2} \iint_{C(I_{\mathbf{w}_R}) \prec C(I_{\mathbf{b}_R})} \frac{\theta(-t + \mathbf{D}_{\beta^n} + \eta - \zeta)}{\theta(-t + \mathbf{D}_{\beta^n}) E(\zeta, \eta)^2} (G_{\mathbf{f}_+}(\zeta, \eta) - G_{\mathbf{f}_-}(\zeta, \eta)) \frac{\Psi_{\beta}(\eta)}{\Psi_{\beta}(\zeta)}, \quad (7.7) \end{aligned}$$

where  $\mathbf{b}_R \in B(\mathbf{R})$  and  $\mathbf{w}_R \in W(\mathbf{R})$  are fixed vertices, and we use that the  $G_{\mathbf{f}}(\zeta, \eta)$ -terms telescope.

We consider first the single integral. If  $\mathbf{e} = \mathbf{b}\mathbf{w} \in E(\mathbf{R})$ , then Lemma 2.10 implies that

$$\frac{1}{2\pi i} \int_{C([\ell_{\mathbf{b}_R}, u_0])} g_{\mathbf{b},\mathbf{w}} \mathbf{K}_{\mathbf{w},\mathbf{b}} = \frac{1}{2\pi i} \int_{C([\ell_{\mathbf{b}_R}, u_0])} \omega_{\beta-\alpha},$$

where  $\alpha$  and  $\beta$  are the two strands containing  $\mathbf{e}$  as in Figure 8. Summing over  $\gamma^*$  yields

$$\sum_{\mathbf{e} = \mathbf{b}\mathbf{w} \in E(\mathbf{R})} \frac{(\mathbf{e} \wedge \gamma^*)}{2\pi i} \int_{C([\ell_{\mathbf{b}_R}, u_0])} g_{\mathbf{b},\mathbf{w}} \mathbf{K}_{\mathbf{w},\mathbf{b}} = \frac{1}{2\pi i} \int_{C([\ell_{\mathbf{b}_R}, u_0])} \omega_{\mathbf{d}(\mathbf{f}_+) - \mathbf{d}(\mathbf{f}_-)}. \quad (7.8)$$

This is the height difference between  $\mathbf{f}_+$  and  $\mathbf{f}_-$  for the periodic dimer cover  $M_{\ell_{\mathbf{b}_R}}$  of  $\mathbf{G}$  with reference dimer cover  $M_0$ . We note that

$$\omega_{\mathbf{d}(\mathbf{f}_+) - \mathbf{d}(\mathbf{f}_-)} = \omega_{\mathbf{d}(\mathbf{f}_+) + \mathbf{D}_{\beta^n} - \mathbf{d}(\mathbf{f}_-) - \mathbf{D}_{\beta^n}} = \omega_{\mathbf{d}(\mathbf{f}_+) + \mathbf{D}_{\beta^n}} - \omega_{\mathbf{d}(\mathbf{f}_-) + \mathbf{D}_{\beta^n}}. \quad (7.9)$$

Now we turn to the double contour integral in (7.7). For  $f \in F(\mathbf{G}_{\beta^n})$ , let

$$\omega_f := \frac{1}{(2\pi i)^2} \frac{1}{E(\zeta, \eta)^2} \frac{\theta(-t + \mathbf{D}_{\beta^n} + \eta - \zeta)}{\theta(-t + \mathbf{D}_{\beta^n})} \frac{\theta(-t - \mathbf{d}(f) + \zeta - \eta)}{\theta(t + \mathbf{d}(f))} \frac{\prod_{\alpha \in \alpha} E(\alpha, \eta)^{(\mathbf{d}(f) + \mathbf{D}_{\beta^n})_\alpha}}{\prod_{\alpha \in \alpha} E(\alpha, \zeta)^{(\mathbf{d}(f) + \mathbf{D}_{\beta^n})_\alpha}}, \quad (7.10)$$

so that the double integral may be written as

$$\iint_{C(I_{\mathbf{w}_R}) \prec C(I_{\mathbf{b}_R})} (\omega_{f_+} - \omega_{f_-}).$$

For  $n = 1, 2, \dots$ , let  $\mathbf{R}_n \subset \mathbf{G}_{\beta^n}$  be a chamber and assume that  $\text{sign}_{\mathbf{R}_n}(e)$  is independent of  $n$  if  $n$  is large enough. Let  $f_n = f_0 + (j_n, k_n) \in F(\mathbf{R}_n)$  with  $(j_n, k_n) = n(x, y) + O(1)$  as  $n \rightarrow \infty$ , where  $(x, y)$  satisfy Assumption 7.11. Let  $\zeta_* = \zeta_*(x, y)$  be defined by (7.1).

Comparing (7.10) with (7.4) makes it clear that we can reuse the steepest descent analysis performed in Section 7.1. We deform  $C(I_{\mathbf{w}})$  to  $C_{(x,y),\text{asc}}$  and  $C(I_{\mathbf{b}})$  to  $C_{(x,y),\text{desc}}$ , where the curves are the steepest ascent and descent curves as in Section 7.1, and  $\mathbf{w} = \mathbf{w}_{\mathbf{R}_n}$ ,  $\mathbf{b} = \mathbf{b}_{\mathbf{R}_n}$ . Similarly to the proof of Theorem 7.2, Lemma 7.7 implies that the only contribution from such deformation comes from the pole at  $\zeta = \eta$  along the curve  $C_{\ell_b}^{\zeta_*}$ . The pole, however, is now a double pole instead of a simple pole, which leads to a very different contribution along this curve:

$$\begin{aligned} \iint_{C(I_{\mathbf{w}}) \prec C(I_{\mathbf{b}})} \omega_{f_n} &= \frac{1}{2\pi i} \int_{C_{\ell_b}^{\zeta_*}} \omega_{\mathbf{d}(f_n) + \mathbf{D}_{\beta^n}} + \int_{C_{(x,y),\text{asc}}} \int_{C_{(x,y),\text{desc}}} \omega_{f_n} \\ &+ \frac{1}{2\pi i} \sum_{k=1}^g \left( \frac{\partial \log \theta}{\partial z_k}(-t + \mathbf{D}_{\beta^n}) - \frac{\partial \log \theta}{\partial z_k}(-t - \mathbf{d}(f_n)) \right) \int_{C_{\ell_b}^{\zeta_*}} \omega_k. \end{aligned} \quad (7.11)$$

Indeed, the residue at the double pole is computed by differentiating the regular part of the integrand in (7.10) and then setting  $\zeta = \eta$ . The derivative of the factor involving the prime forms gives the first term in (7.11) by (2.2), whereas the derivative of the theta-function factor gives the third term. By the periodicity of  $\theta(z)$  for  $z \in \mathbb{R}^g$ , the third term is bounded in  $n$ . The second term on the right-hand side of (7.11) is likewise bounded in  $n$ , by standard steepest descent arguments. Therefore, as  $n \rightarrow \infty$ , the main contribution to (7.11) comes from the first term on the right-hand side.

Let  $f_{n,+} = f_0 + (i_{n,+}, j_{n,+}) \in F(\mathbf{R}_n)$  and  $f_{n,-} = f_0 + (i_{n,-}, j_{n,-}) \in F(\mathbf{R}_n)$  with  $(i_{n,+}, j_{n,+}) = n(x_+, y_+) + O(1)$  and  $(i_{n,-}, j_{n,-}) = n(x_-, y_-) + O(1)$  as  $n \rightarrow \infty$ , for some  $(x_+, y_+), (x_-, y_-) \in \mathcal{D}_c^\circ$  satisfying Assumption 7.11. The assumption on the limiting points is a technical condition ensuring that we can deform the integration contours to the corresponding steepest descent and ascent curves; see Remark 7.8. Let  $\zeta_+ = \zeta_*(x_+, y_+)$  be defined by (7.1), and similarly  $\zeta_- = \zeta_*(x_-, y_-)$ . Then, by (7.7), (7.8), (7.9), and (7.11),

$$\begin{aligned} \mathbb{E}[h_M(f_{n,+}) - h_M(f_{n,-})] &= \frac{1}{2\pi i} \int_{C([\ell_b, u_0])} \omega_{\mathbf{d}(f_{n,+}) - \mathbf{d}(f_{n,-})} + \frac{1}{2\pi i} \int_{C_{\ell_b}^{\zeta_+}} \omega_{\mathbf{d}(f_{n,+}) + \mathbf{D}_{\beta^n}} \\ &- \frac{1}{2\pi i} \int_{C_{\ell_b}^{\zeta_-}} \omega_{\mathbf{d}(f_{n,-}) + \mathbf{D}_{\beta^n}} + O(1) = \frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_+}} \omega_{\mathbf{d}(f_{n,+}) + \mathbf{D}_{\beta^n}} - \frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_-}} \omega_{\mathbf{d}(f_{n,-}) + \mathbf{D}_{\beta^n}} + O(1), \end{aligned}$$

as  $n \rightarrow \infty$ . By Lemma 6.3 we conclude that, as  $n \rightarrow \infty$ ,

$$\mathbb{E}[h_M(f_{n,+}) - h_M(f_{n,-})] = -n \frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_+}} \lambda_c(x_+, y_+) + n \frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_-}} \lambda_c(x_-, y_-) + O(1). \quad (7.12)$$

Thus, the height difference is consistent with (7.5) when  $(x_+, y_+)$  and  $(x_-, y_-)$  lie in the same chamber. To treat the general case, we connect the two points by a curve and decompose the height difference into a sum of contributions from the chambers crossed by the curve. More precisely, we split the curve into long segments lying inside single chambers and short segments crossing the separating lines between chambers. The long segments are handled by (7.12), and we will show that the short segments contribute negligibly.

Let  $(x, y)$  satisfy Assumption 7.11 and let  $\mathbf{f}_n = \mathbf{f}_0 + (j_n, k_n) \in F(\mathbf{G}_{\beta^n})$  be such that  $(j_n, k_n) = n(x, y) + O(1)$  as  $n \rightarrow \infty$ . Pick a curve  $\Gamma$  in  $\mathcal{D}_c$  from  $(0, 0)$  to  $(x, y)$ , intersecting the lines  $\mathbf{l}_e$ ,  $e \in E(N)$ , at a finite number of locations  $(x_k, y_k)$ ,  $k = 1, \dots, m-1$ . We assume that these intersections are transverse, that each point  $(x_k, y_k)$  lies on exactly one line  $\mathbf{l}_e$  with  $e \in E(N)$ , and that the points  $(x_k, y_k)$  satisfy Assumption 7.11. The last assumption is possible because the rough region is locally convex, so  $\mathcal{A}$  intersects each line  $\mathbf{l}_e$  only finitely many times.

If  $(x_k, y_k) \in \mathbf{l}_e$ , let  $R_k$  and  $R_{k+1}$  be the 2-cells adjacent to the part of  $\mathbf{l}_e$  containing  $(x_k, y_k)$ , so that  $\Gamma$  enters  $(x_k, y_k)$  from  $R_k$  and exits into  $R_{k+1}$ . For large enough  $n$ , Lemma 7.5 implies that, for  $k = 1, \dots, m$ , there is a chamber  $\mathbf{R}_{n,k}$  in  $\mathbf{G}_{\beta^n}$  such that

$$\text{sign}_{\mathbf{R}_{n,k}}(e) = \text{sign}_{R_k}(e),$$

for all  $e \in E(N)$ . Let

$$\mathbf{f}_{n,k,+} = \mathbf{f}_0 + (i_{n,k,+}, j_{n,k,+}) \in F(\mathbf{R}_{n,k}) \quad \text{and} \quad \mathbf{f}_{n,k,-} = \mathbf{f}_0 + (i_{n,k,-}, j_{n,k,-}) \in F(\mathbf{R}_{n,k+1}),$$

where

$$(i_{n,k,+}, j_{n,k,+}) = n(x_k, y_k) + O(1), \quad (i_{n,k,-}, j_{n,k,-}) = n(x_k, y_k) + O(1) \quad \text{as} \quad n \rightarrow \infty. \quad (7.13)$$

We also set  $\mathbf{f}_{n,0,+} = \mathbf{f}_0$ ,  $\mathbf{f}_{n,m,-} = \mathbf{f}_n$ ,

$$\mathbf{f}_{n,0,-} = \mathbf{f}_0 + (i_{n,0,-}, j_{n,0,-}) \in F(\mathbf{R}_{n,1}) \quad \text{with} \quad (i_{n,0,-}, j_{n,0,-}) = O(1) \quad \text{as} \quad n \rightarrow \infty,$$

and

$$\mathbf{f}_{n,m,+} = \mathbf{f}_0 + (i_{n,m,+}, j_{n,m,+}) \in F(\mathbf{R}_{n,m}) \quad \text{with} \quad (i_{n,m,+}, j_{n,m,+}) = n(x, y) + O(1) \quad \text{as} \quad n \rightarrow \infty.$$

For  $k = 1, \dots, m$ , we define the dual path  $\gamma_{n,k,1}^*$  contained in  $\mathbf{R}_{n,k}$  going from  $\mathbf{f}_{n,k-1,-}$  to  $\mathbf{f}_{n,k,+}$ , and for  $k = 0, \dots, m$ , we define the dual path  $\gamma_{n,k,s}^*$  going from  $\mathbf{f}_{n,k,+}$  to  $\mathbf{f}_{n,k,-}$ . Here  $l$  and  $s$  stand for long and short respectively. The concatenation

$$\begin{array}{ccccccccccc} \mathbf{f}_0 = \mathbf{f}_{n,0,+} & \xrightarrow{\gamma_{n,0,s}^*} & \mathbf{f}_{n,0,-} & \xrightarrow{\gamma_{n,1,l}^*} & \mathbf{f}_{n,1,+} & \xrightarrow{\gamma_{n,1,s}^*} & \mathbf{f}_{n,1,-} & \xrightarrow{\gamma_{n,2,l}^*} & \dots & \xrightarrow{\gamma_{n,m,l}^*} & \mathbf{f}_{n,m,+} & \xrightarrow{\gamma_{n,m,s}^*} & \mathbf{f}_{n,m,-} = \mathbf{f}_n \\ & \searrow & \swarrow & & \searrow & \swarrow & & & & & \searrow & \swarrow & \\ & & (0, 0) & & & (x_1, y_1) & & & \dots & & & & (x, y) \end{array}$$

of these paths defines a dual path  $\gamma_n^*$  from  $\mathbf{f}_0$  to  $\mathbf{f}_n$ , where the downward arrows indicate the points to which the faces converge in the scaling limit.

We first bound the part of (7.6) coming from the curves  $\gamma_{n,k,s}^*$ . By (7.13), we may pick  $\gamma_{n,k,s}^*$  so that its length is uniformly bounded for all  $n$ . We obtain

$$\sum_{\mathbf{e} \in E(\mathbf{G}_{\beta^n})} (\mathbf{e} \wedge \gamma_{n,k,s}^*) \mathbb{E}[\mathbf{1}_{\mathbf{e} \in M} - \mathbf{1}_{\mathbf{e} \in M_0}] = O(1)$$

as  $n \rightarrow \infty$ , for  $k = 0, \dots, m$ .

We continue by summing over the curves  $\gamma_{n,k,1}^*$ . For  $k = 1, \dots, m-1$ , let  $\zeta_k = \zeta_*(x_k, y_k)$  be given by (7.1). Let  $\zeta_0 \in A_0 \setminus \alpha$  be any point in the component corresponding to  $u_0$ . Then, by (7.12),

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^m \mathbb{E}[h_M(\mathbf{f}_{n,k,+}) - h_M(\mathbf{f}_{n,k-1,-})] &= \sum_{k=1}^m \left( -\frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_k}} \lambda_c(x_k, y_k) + \frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_{k-1}}} \lambda_c(x_{k-1}, y_{k-1}) + O\left(\frac{1}{n}\right) \right) \\ &= \frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_0}} \lambda_c(0, 0) - \frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_0^*}} \lambda_c(x, y) + O\left(\frac{1}{n}\right), \end{aligned}$$

as  $n \rightarrow \infty$ . Since the curve  $C_{u_0}^{\zeta_0}$  does not enclose any angles, the above equality proves Theorem 7.9.  $\square$

Using the formulation (6.2) for  $\lambda_c(x, y)$ , we see that the slope of  $\mathfrak{h}$  is

$$\nabla \mathfrak{h}(x, y) = \left( -\frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_0^*}} d \log z, -\frac{1}{2\pi i} \int_{C_{u_0}^{\zeta_0^*}} d \log w \right).$$

The right-hand side agrees with the slope (2.7) of the Gibbs measure  $\mathbb{P}^{\zeta_*}$ . We are led to the following corollary of Theorems 7.2 and 7.9.

**Corollary 7.12.** *Assume that the angle function is periodic, and suppose  $(x, y) \in \mathcal{D}_c^\circ$  lies in a phase region and satisfies Assumption 7.11. The local correlations in a neighborhood of  $(x, y)$  of the dimer model on  $\mathcal{G}_{\beta^n}$  defined by the Kasteleyn matrix (2.4) converge to those of the Gibbs measure defined by the corresponding whole-plane Kasteleyn matrix with slope  $\nabla \mathfrak{h}(x, y)$ .*

**Remark 7.13.** In the case of the Aztec diamond, for  $(x, y) \in \mathcal{G}_j$ ,  $j = 1, \dots, g$ , the part of the integral (7.5) coming from the third term in expression (6.2) for  $\lambda_c(x, y)$ , namely

$$-\frac{1}{2\pi i} \int_{B_j} \omega_{D_c},$$

with  $D_c$  as in (6.1), turned out to be important in the description of the global height fluctuations [16, Equation (6)]. It is therefore natural to expect that the same quantity will play a similar role in the global height fluctuations of AZ graphs.

## A Tropical limits and simulations

In this appendix, we explain how certain astroidal domains arise as local pieces of the tropical limit of the Aztec diamond with periodic weights. This allows us to use domino shuffling on the Aztec diamond [102] to simulate limit shapes in more general astroidal domains.

Let  $\mathbf{G}$  be a minimal periodic bipartite graph and let  $p : \mathbf{G} \rightarrow \mathbf{G}_{\mathbb{T}}$  denote the quotient map to the corresponding graph  $\mathbf{G}_{\mathbb{T}}$  on the torus. There is a construction, called the *spectral transform* [82], which allows one to view any periodic choice of edge weights on  $\mathbf{G}$  (equivalently, edge weights on  $\mathbf{G}_{\mathbb{T}}$ ) as gauge-equivalent to Fock's dimer model for an M-curve  $\Sigma$ ,  $t \in \text{Jac}(\Sigma)$ , and angle function  $\nu$  [52, 26]. We apply this to the family of edge weights on  $\mathbf{G}_{\mathbb{T}}$  of the form

$$\text{wt}(\mathbf{e}; \tau) = w(\mathbf{e}) \exp(\tau \mathcal{E}(\mathbf{e})),$$

where

$$w : E(\mathbb{G}_{\mathbb{T}}) \rightarrow \mathbb{R}_{>0}, \quad \mathcal{E} : E(\mathbb{G}_{\mathbb{T}}) \rightarrow \mathbb{R}, \quad \tau \in \mathbb{R}_{>0}.$$

The corresponding M-curve  $\Sigma_{\tau}$  is an explicit algebraic curve given by the vanishing locus of the *characteristic polynomial*

$$P(z, w; \tau) := \sum_{(j,k) \in \mathbb{Z}^2} \pm \left( \sum_{\text{dimer covers } M \text{ on } \mathbb{G}_{\mathbb{T}}: (h_M^y, -h_M^x) = (j,k)} \left( \prod_{e \in M} \text{wt}(e; \tau) \right) z^j w^k \right),$$

which is a Laurent polynomial in the variables  $z, w$  with Newton polygon equal to  $N$  [84]. Here  $\pm$  denotes a sign that depends only on  $(j, k) \in \mathbb{Z}^2$  and  $h_M^x$  and  $h_M^y$  are the height changes of  $M$  along dual paths with homology classes  $(1, 0)$  and  $(0, 1)$  respectively.

The limit as  $\tau \rightarrow \infty$  is called the tropical limit; see [69, 91] for a general introduction to tropical geometry. The *tropical characteristic polynomial* is defined as

$$P^t(x, y) := \max_{(j,k) \in \mathbb{Z}^2} \left( \max_{\text{dimer covers } M \text{ on } \mathbb{G}_{\mathbb{T}}: (h_M^y, -h_M^x) = (j,k)} \left( \sum_{e \in M} \mathcal{E}(e) \right) + jx + ky \right). \quad (\text{A.1})$$

The fundamental theorem of tropical geometry [50] states that, as  $\tau \rightarrow \infty$ , the rescaled *amoeba*

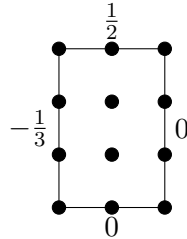
$$\frac{1}{\tau} \{ (\log |z|, \log |w|) \in \mathbb{R}^2 : P(z, w; \tau) = 0 \}$$

converges in the Hausdorff metric to the *tropical spectral curve*

$$\Sigma^t := \{ (x, y) \in \mathbb{R}^2 : \text{the maximum over } (j, k) \in \mathbb{Z}^2 \text{ in (A.1) is achieved at least twice} \}.$$

The graph of the tropical characteristic polynomial induces a regular subdivision of the Newton polygon  $N$  which is dual to  $\Sigma^t$  (see Figure 31). The tropical curve is called *smooth* if all 2-cells of the subdivision are unimodular triangles. We illustrate, by means of an example, how the arctic curve can be computed in the tropical limit by a generalization of the results of [11]. Their setting was for smooth tropical curves, whereas here we consider a non-smooth tropical curve.

Consider the  $(2 \times 3)$ -fundamental domain of the square lattice with the family of edge weights shown in Figure 30(left). The Newton polygon  $N$  and the function  $c : E(N) \rightarrow \mathbb{R}$  for the scaling limit of the Aztec diamond, are given by



The tropical curve for this family of edge weights is shown in Figure 31(left).

A *discrete 1-form*  $\omega$  on  $\Sigma^t$  is an antisymmetric function on the oriented edges of  $\Sigma^t$ . It is called *harmonic* if it is

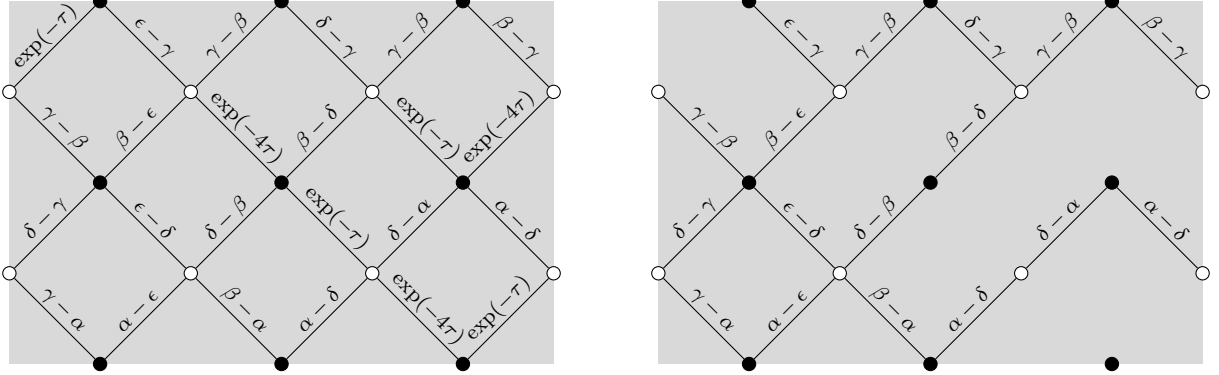


Figure 30: A  $(2 \times 3)$ -fundamental domain of the square lattice with a family of edge weights (left), and the corresponding limiting graph obtained by deleting the edges whose weights vanish in the tropical limit (right). This particular choice of weights was obtained via the tropical inverse spectral transform developed in the forthcoming work [56].

- *closed*, i.e. for every face  $f$  of  $\Sigma^t$ ,

$$\sum_{e \subset \partial f} |e|_{\mathbb{Z}} \omega(e) = 0,$$

where the edges in  $\partial f$  are oriented counterclockwise, and

- *co-closed*, i.e. for every vertex  $u \in V(\Sigma^t)$ ,

$$\sum_{e=uv} \omega(e) = 0,$$

where the sum is over all edges incident to  $u$ , each oriented from  $u$  to  $v$ .

Given any outflows on the unbounded edges of  $\Sigma^t$  that sum to 0, there is a unique harmonic 1-form with those boundary values. Taking the boundary values given by the function  $c$ , we obtain the harmonic 1-form shown in Figure 32(left).

The map  $\Omega^{-1}$  tropicalizes, in the limit  $\tau \rightarrow \infty$ , to the map

$$(\Omega^{-1})^t : \Sigma^t \rightarrow \mathcal{D}_c$$

sending  $v \in V(\Sigma^t)$  to the point  $(x, y) \in \mathbb{R}^2$  such that

$$\frac{1}{|e|_{\mathbb{Z}}} \langle (x, y), e \rangle + \omega(e) = 0 \tag{A.2}$$

for every edge  $e$  of  $\Sigma^t$  incident to  $v$ , and sending non-boundary edges of  $\Sigma^t$  to the corresponding line segments between the images of their endpoints. Its image is shown in Figure 32(right). Note that the degree-5 vertex  $O$  maps to a pentagonal astroidal domain  $\mathcal{D}_O$  bounded by the lines defined by (A.2). The condition that  $\omega$  is co-closed at  $O$  is equivalent to the admissibility condition for  $\mathcal{D}_O$ .

Letting  $\tau \rightarrow \infty$  and deleting edges with zero weight, we obtain the graph shown on the right-hand side of Figure 30, which is a minimal graph for the pentagonal Newton polygon shown shaded in Figure 31(right). Its M-curve  $\Sigma_{\infty}$  is given by the vanishing of the *initial polynomial*

$$\lim_{\tau \rightarrow \infty} P(z, w; \tau).$$

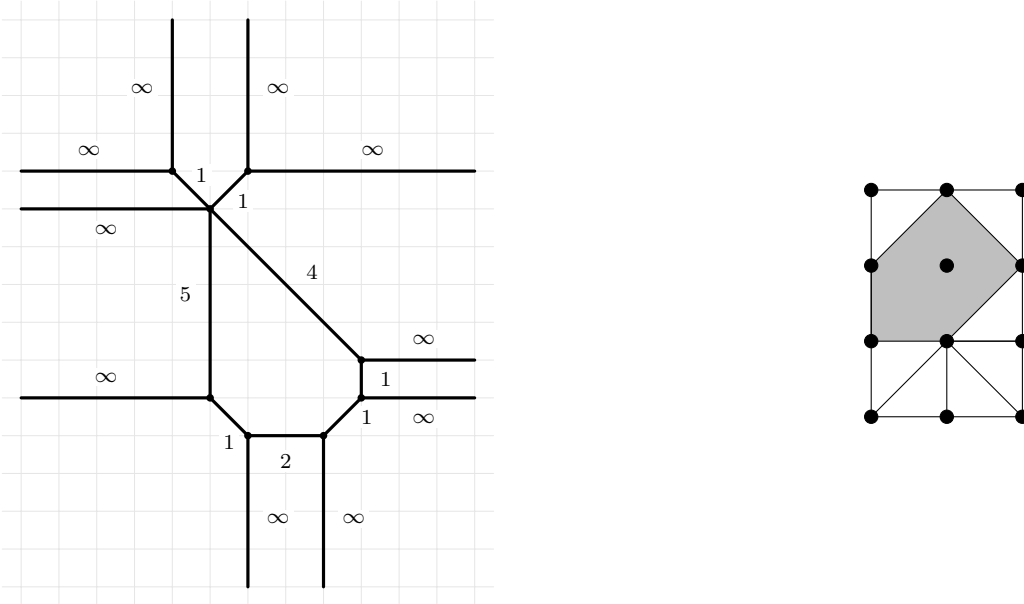


Figure 31: The tropical spectral curve with the integral lengths of its edges (left), and the dual regular subdivision of the Newton polygon with a shaded pentagonal 2-cell (right).

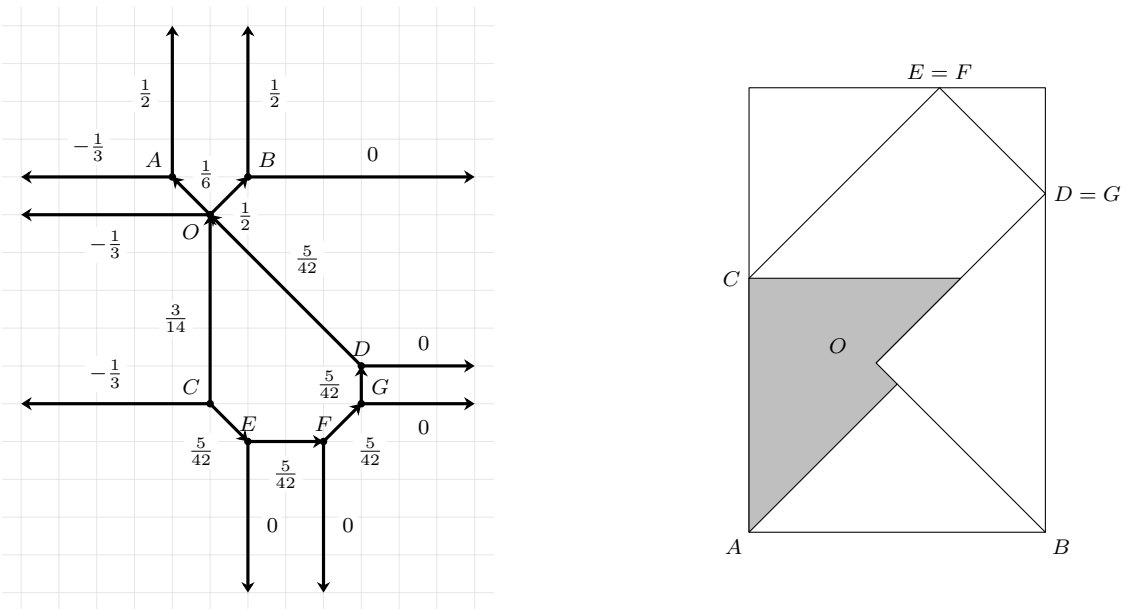


Figure 32: The harmonic 1-form on  $\Sigma^t$  (left), and the image of  $(\Omega^{-1})^t$  in  $\mathcal{D}_c$  (right).

Our choice of weights is such that  $\Sigma_\infty$  has genus 0, and the weights  $\text{wt}(\mathbf{e}; \infty)$  are isoradial weights as in Section 4.3. As in that section, the prime form is given by

$$E(\zeta, \eta) = \frac{\eta - \zeta}{\sqrt{d\zeta}\sqrt{d\eta}}.$$

Thus, the functions in (6.6) can be written explicitly and globally as

$$f_1(\zeta) = - \sum_{\alpha \in \alpha} b_{e(\alpha)} \log(\zeta - \alpha), \quad f_2(\zeta) = \sum_{\alpha \in \alpha} a_{e(\alpha)} \log(\zeta - \alpha), \quad f_3(\zeta) = \sum_{\alpha \in \alpha} c(e(\alpha)) \log(\zeta - \alpha).$$

Choosing the angles

$$\alpha = 1, \quad \beta = e^{\pi i/25}, \quad \gamma = e^{4\pi i/5}, \quad \delta = e^{6\pi i/5}, \quad \epsilon = e^{8\pi i/5},$$

and using the parametrization (6.10), we obtain the arctic curve in  $\mathcal{D}_O$  shown in Figure 6(right). For comparison, a simulation of the Aztec diamond is shown in Figure 6(left).

## B Proof of Proposition 7.10

The following proof is reproduced from [13]. By the product form (2.5) of  $g$ , we have

$$g_x = g_{f_+} g_{f_+, x}, \quad x \in \{\mathbf{b}, \mathbf{w}, \mathbf{f}_-, \mathbf{f}_+\}.$$

Canceling the  $g_{f_+}$ -factors, it suffices to prove that

$$\begin{aligned} E(\zeta, \eta) \frac{g_{f_+, \mathbf{w}}(\eta)}{g_{f_+, \mathbf{b}}(\zeta)} K_{\mathbf{w}, \mathbf{b}} &= \frac{\theta(-t - \mathbf{d}(\mathbf{f}_+) + \zeta - \eta)}{\theta(t + \mathbf{d}(\mathbf{f}_+))} \\ &\quad - \frac{\theta(-t - \mathbf{d}(\mathbf{f}_-) + \zeta - \eta)}{\theta(t + \mathbf{d}(\mathbf{f}_-))} \frac{E(\beta, \zeta)}{E(\alpha, \zeta)} \frac{E(\alpha, \eta)}{E(\beta, \eta)}. \end{aligned} \tag{B.1}$$

By definition,

$$E(\zeta, \eta) \frac{g_{f_+, \mathbf{w}}(\eta)}{g_{f_+, \mathbf{b}}(\zeta)} K_{\mathbf{w}, \mathbf{b}} = \frac{\theta(-t - \mathbf{d}(\mathbf{b}) + \zeta) \theta(t + \mathbf{d}(\mathbf{w}) + \eta) E(\alpha, \beta) E(\zeta, \eta)}{\theta(t + \mathbf{d}(\mathbf{f}_+)) \theta(t + \mathbf{d}(\mathbf{f}_-)) E(\alpha, \zeta) E(\beta, \eta)}.$$

We now apply Fay's trisecant identity [51] (see also [26, Theorem 27]): for  $z \in \mathbb{C}^g$  and  $a, b, c, d \in \Sigma$ ,

$$\begin{aligned} \theta(z + c - a) \theta(z + d - b) E(c, b) E(a, d) + \theta(z + c - b) \theta(z + d - a) E(c, a) E(d, b) \\ = \theta(z + c + d - a - b) \theta(z) E(c, d) E(a, b). \end{aligned}$$

Dividing by  $\theta(z) \theta(z + d - a) E(c, a) E(d, b)$ , we obtain

$$\begin{aligned} \frac{\theta(z + c - a) \theta(z + d - b) E(c, b) E(a, d)}{\theta(z) \theta(z + d - a) E(c, a) E(d, b)} \\ = \frac{\theta(z + c + d - a - b) E(c, d) E(a, b)}{\theta(z + d - a) E(c, a) E(d, b)} - \frac{\theta(z + c - b)}{\theta(z)}. \end{aligned}$$

Now set

$$z = -t - \mathbf{d}(\mathbf{f}_+), \quad a = \alpha, \quad b = \eta, \quad c = \zeta, \quad d = \beta.$$

Using

$$z - a = -t - \mathbf{d}(\mathbf{b}), \quad z + d = -t - \mathbf{d}(\mathbf{w}), \quad z + d - a = -t - \mathbf{d}(\mathbf{f}_-),$$

we get

$$\begin{aligned} & \frac{\theta(-t - \mathbf{d}(\mathbf{b}) + \zeta) \theta(-t - \mathbf{d}(\mathbf{w}) - \eta) E(\zeta, \eta) E(\alpha, \beta)}{\theta(-t - \mathbf{d}(\mathbf{f}_+)) \theta(-t - \mathbf{d}(\mathbf{f}_-)) E(\zeta, \alpha) E(\beta, \eta)} \\ &= \frac{\theta(-t - \mathbf{d}(\mathbf{f}_-) + \zeta - \eta) E(\zeta, \beta) E(\alpha, \eta)}{\theta(-t - \mathbf{d}(\mathbf{f}_-)) E(\zeta, \alpha) E(\beta, \eta)} - \frac{\theta(-t - \mathbf{d}(\mathbf{f}_+) + \zeta - \eta)}{\theta(-t - \mathbf{d}(\mathbf{f}_+))}. \end{aligned}$$

Finally, using that the Riemann theta function is even and that the prime form is antisymmetric, the previous identity simplifies to (B.1). This proves the proposition.  $\square$

## References

- [1] Mark Adler, Sunil Chhita, Kurt Johansson, and Pierre van Moerbeke. Tacnode GUE-minor processes and double Aztec diamonds. *Probab. Theory Related Fields*, 162(1-2):275–325, 2015.
- [2] Mark Adler, Kurt Johansson, and Pierre van Moerbeke. Tilings of non-convex polygons, skew-Young tableaux and determinantal processes. *Comm. Math. Phys.*, 364(1):287–342, 2018.
- [3] Amol Aggarwal. Arctic boundaries of the ice model on three-bundle domains. *Invent. Math.*, 220(2):611–671, 2020.
- [4] Amol Aggarwal. Universality for lozenge tiling local statistics. *Ann. of Math. (2)*, 198(3):881–1012, 2023.
- [5] Amol Aggarwal and Vadim Gorin. Gaussian unitary ensemble in random lozenge tilings. *Probab. Theory Related Fields*, 184(3-4):1139–1166, 2022.
- [6] Amol Aggarwal and Jiaoyang Huang. Edge statistics for lozenge tilings of polygons, II: Airy line ensemble. *Forum Math. Pi*, 13:Paper No. e2, 60, 2025.
- [7] Riccardo Argurio, Matteo Bertolini, Sebastián Franco, Eduardo García-Valdecasas, Shani Meynet, Antoine Pasternak, and Valdo Tatitscheff. The octagon and the non-supersymmetric string landscape. *Physics Letters B*, 815:136153, 2021.
- [8] Kari Astala, Erik Duse, István Prause, and Xiao Zhong. Dimer Models and Conformal Structures. *Comm. Pure Appl. Math.*, 79(2):340–446, 2025.
- [9] Yuliy Baryshnikov and Robin Pemantle. Asymptotics of multivariate sequences, part III: Quadratic points. *Adv. Math.*, 228(6):3127–3206, 2011.
- [10] Tomas Berggren. Domino tilings of the Aztec diamond with doubly periodic weightings. *Ann. Probab.*, 49(4):1965–2011, 2021.
- [11] Tomas Berggren and Alexei Borodin. Crystallization of the Aztec diamond, 2024. arXiv:2410.04187.
- [12] Tomas Berggren and Alexei Borodin. Geometry of the doubly periodic Aztec dimer model. *Commun. Am. Math. Soc.*, 5:475–570, 2025.

- [13] Tomas Berggren, Cédric Boutillier, Marianna Russkikh, and Béatrice de Tilière. *In preparation*.
- [14] Tomas Berggren and Nediialko Bradinoff. Marked GUE-corners process in doubly periodic dimer models, 2026. arXiv:2603.27906.
- [15] Tomas Berggren and Maurice Duits. Correlation functions for determinantal processes defined by infinite block Toeplitz minors. *Adv. Math.*, 356:106766, 48, 2019.
- [16] Tomas Berggren and Matthew Nicoletti. Gaussian Free Field and Discrete Gaussians in Periodic Dimer Models, 2025. arXiv:2502.07241.
- [17] Tomas Berggren, Matthew Nicoletti, and Marianna Russkikh. Perfect t-embeddings of uniformly weighted Aztec diamonds and tower graphs. *Int. Math. Res. Not. IMRN*, (7):5963–6007, 2024.
- [18] Alexander I. Bobenko and Nikolai Bobenko. Dimers and M-Curves: Limit Shapes from Riemann Surfaces, 2024. arXiv:2407.19462. To appear in Duke Math. Journal.
- [19] Alexander I. Bobenko, Nikolai Bobenko, and Yuri B. Suris. Dimers and M-curves. In *The Versatility of Integrability—In Memory of Igor Krichever*, volume 823 of *Contemp. Math.*, pages 1–45. American Mathematical Society, Providence, RI, 2025.
- [20] Alexei Borodin and Maurice Duits. Biased  $2 \times 2$  periodic Aztec diamond and an elliptic curve. *Probab. Theory Related Fields*, 187(1-2):259–315, 2023.
- [21] Alexei Borodin and Patrik L. Ferrari. Random tilings and Markov chains for interlacing particles. *Markov Process. Related Fields*, 24(3):419–451, 2018.
- [22] Alexei Borodin, Vadim Gorin, and Eric M. Rains.  $q$ -distributions on boxed plane partitions. *Selecta Math. (N.S.)*, 16(4):731–789, 2010.
- [23] Mireille Bousquet-Mélou, James Propp, and Julian West. Perfect matchings for the three-term Gale-Robinson sequences. *Electron. J. Combin.*, 16(1):Research Paper 125, 37, 2009.
- [24] Cédric Boutillier, David Cimasoni, and Béatrice de Tilière. Isoradial immersions. *J. Graph Theory*, 99(4):715–757, 2022.
- [25] Cédric Boutillier, David Cimasoni, and Béatrice de Tilière. Elliptic dimers on minimal graphs and genus 1 Harnack curves. *Comm. Math. Phys.*, 400(2):1071–1136, 2023.
- [26] Cédric Boutillier, David Cimasoni, and Béatrice de Tilière. Minimal bipartite dimers and higher genus Harnack curves. *Probab. Math. Phys.*, 4(1):151–208, 2023.
- [27] Cédric Boutillier and Béatrice de Tilière. Fock’s dimer model on the Aztec diamond. *Ann. Inst. Henri Poincaré Comb. Phys. Interact.*, 2026. Published online first.
- [28] Cédric Boutillier, Béatrice de Tilière, and Bishal Deb. Fock’s dimer model on Speyer graphs, 2026. In preparation.

- [29] Cédric Boutillier, Jérémie Bouttier, Guillaume Chapuy, Sylvie Corteel, and Sanjay Ramasamy. Dimers on Rail Yard Graphs. *Ann. Inst. Henri Poincaré Comb. Phys. Interact.*, 4(4):479–539, 2017.
- [30] Raphaël Cerf and Richard Kenyon. The low-temperature expansion of the Wulff crystal in the 3D Ising model. *Comm. Math. Phys.*, 222(1):147–179, 2001.
- [31] Christophe Charlier, Maurice Duits, Arno B. J. Kuijlaars, and Jonatan Lenells. A periodic hexagon tiling model and non-Hermitian orthogonal polynomials. *Comm. Math. Phys.*, 378(1):401–466, 2020.
- [32] Qiyue Chen and Gregg Musiker. Gale-Robinson quivers and principal coefficients, 2026. arXiv:2602.19402.
- [33] Sunil Chhita, Duncan Dauvergne, and Thomas Finn. The Airy line ensemble at the rough-smooth boundary, 2026. arXiv:2602.24063.
- [34] Sunil Chhita and Kurt Johansson. Domino statistics of the two-periodic Aztec diamond. *Adv. Math.*, 294:37–149, 2016.
- [35] Sunil Chhita, Kurt Johansson, and Benjamin Young. Asymptotic domino statistics in the Aztec diamond. *Ann. Appl. Probab.*, 25(3):1232–1278, 2015.
- [36] Sunil Chhita and Benjamin Young. Coupling functions for domino tilings of Aztec diamonds. *Adv. Math.*, 259:173–251, 2014.
- [37] Judy Hsin-Hui Chiang, Gregg Musiker, and Son Nguyen. Minimal Matchings for  $dP_3$  Cluster Variables. 2024. arXiv:2409.05803.
- [38] Henry Cohn, Richard Kenyon, and James Propp. A variational principle for domino tilings. *J. Amer. Math. Soc.*, 14(2):297–346, 2001.
- [39] F. Colomo and A. Sportiello. Arctic curves of the six-vertex model on generic domains: the tangent method. *J. Stat. Phys.*, 164(6):1488–1523, 2016.
- [40] Cyndie Cottrell and Benjamin Young. Domino shuffling for the Del Pezzo 3 lattice, 2010. arXiv:1011.0045.
- [41] Philippe Di Francesco and Emmanuel Guitter. Arctic curves for paths with arbitrary starting points: a tangent method approach. *J. Phys. A*, 51(35):355201, 45, 2018.
- [42] Philippe Di Francesco and Emmanuel Guitter. The arctic curve for Aztec rectangles with defects via the tangent method. *J. Stat. Phys.*, 176(3):639–678, 2019.
- [43] Philippe Di Francesco and Matthew F. Lapa. Arctic curves in path models from the tangent method. *J. Phys. A*, 51(15):155202, 55, 2018.
- [44] Philippe Di Francesco and Rodrigo Soto-Garrido. Arctic curves of the octahedron equation. *J. Phys. A*, 47(28):285204, 34, 2014.
- [45] Philippe Di Francesco and Hieu Trung Vu. Arctic curves of the  $T$ -system with slanted initial data. *J. Phys. A*, 57(33):Paper No. 335201, 57, 2024.

- [46] R. Dobrushin, R. Kotecký, and S. Shlosman. *Wulff construction*, volume 104 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1992. A global shape from local interaction, Translated from the Russian by the authors.
- [47] J. J. Duistermaat and J. A. C. Kolk. *Multidimensional real analysis. I. Differentiation*, volume 86 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2004. Translated from the Dutch by J. P. van Braam Houckgeest.
- [48] Maurice Duits and Arno B. J. Kuijlaars. The two-periodic Aztec diamond and matrix valued orthogonal polynomials. *J. Eur. Math. Soc. (JEMS)*, 23(4):1075–1131, 2021.
- [49] Richard Eager and Sebastián Franco. Colored BPS pyramid partition functions, quivers and cluster transformations. *J. High Energy Phys.*, (9):038, 2012. arXiv:1112.1132.
- [50] Manfred Einsiedler, Mikhail Kapranov, and Douglas Lind. Non-Archimedean amoebas and tropical varieties. *J. Reine Angew. Math.*, 601:139–157, 2006.
- [51] John D. Fay. *Theta functions on Riemann surfaces*. Lecture Notes in Mathematics, Vol. 352. Springer-Verlag, Berlin-New York, 1973.
- [52] Vladimir V. Fock. Inverse spectral problem for GK integrable system., 2015. arXiv:1503.00289.
- [53] Vladimir V. Fock and Andrey Marshakov. Loop groups, clusters, dimers and integrable systems. In *Geometry and quantization of moduli spaces*, Adv. Courses Math. CRM Barcelona, pages 1–66. Birkhäuser/Springer, Cham, 2016.
- [54] Sebastian Franco, Amihay Hanany, Kristian D. Kennaway, David Vegh, and Brian Wecht. Brane dimers and quiver gauge theories. *J. High Energy Phys.*, (1):096, 2006.
- [55] Sebastian Franco, Amihay Hanany, Dario Martelli, James Sparks, David Vegh, and Brian Wecht. Gauge theories from toric geometry and brane tilings. *J. High Energy Phys.*, (1):128, 2006.
- [56] Pavel Galashin and Terrence George. *In preparation*.
- [57] Yibo Gao, Zhaoqi Li, Thuy-Duong Vuong, and Lisa Yang. Toric mutations in the  $dP_2$  quiver and subgraphs of the  $dP_2$  brane tiling. *Electron. J. Combin.*, 26(2):Paper No. 2.19, 46, 2019.
- [58] Terrence George, Alexander B. Goncharov, and Richard Kenyon. The inverse spectral map for dimers. *Math. Phys. Anal. Geom.*, 26(3):Paper No. 24, 51, 2023.
- [59] Terrence George and Giovanni Inchiostro. The cluster modular group of the dimer model. *Ann. Inst. Henri Poincaré D*, 11(1):147–198, 2024.
- [60] Terrence George and Sanjay Ramassamy. Discrete dynamics in cluster integrable systems from geometric  $R$ -matrix transformations. *Comb. Theory*, 3(2):Paper No. 12, 29, 2023.
- [61] Tomack Gilmore. Inverting the Kasteleyn matrix for holey hexagons, 2017. arXiv:1701.07092.
- [62] Alexander B. Goncharov and Richard Kenyon. Dimers and cluster integrable systems. *Ann. Sci. Éc. Norm. Supér. (4)*, 46(5):747–813, 2013.

- [63] Vadim Gorin. Bulk universality for random lozenge tilings near straight boundaries and for tensor products. *Comm. Math. Phys.*, 354(1):317–344, 2017.
- [64] Vadim Gorin. *Lectures on random lozenge tilings*, volume 193 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2021.
- [65] Harald Helfgott. Edge Effects on Local Statistics in Lattice Dimers: A Study of the Aztec Diamond (Finite Case), 2000. arXiv:math/0007136.
- [66] Jiaoyang Huang. Edge statistics for lozenge tilings of polygons, I: concentration of height function on strip domains. *Probab. Theory Related Fields*, 188(1-2):337–485, 2024.
- [67] Jiaoyang Huang, Fan Yang, and Lingfu Zhang. Pearcey universality at cusps of polygonal lozenge tilings. *Comm. Pure Appl. Math.*, 77(9):3708–3784, 2024.
- [68] Akira Ishii and Kazushi Ueda. Dimer models and the special McKay correspondence. *Geom. Topol.*, 19(6):3405–3466, 2015.
- [69] Ilia Itenberg, Grigory Mikhalkin, and Eugenio Shustin. *Tropical algebraic geometry*, volume 35 of *Oberwolfach Seminars*. Birkhäuser Verlag, Basel, second edition, 2009.
- [70] In-Jee Jeong, Gregg Musiker, and Sicong Zhang. Gale-Robinson sequences and brane tilings. In *25th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2013)*, Discrete Math. Theor. Comput. Sci. Proc., AS, pages 707–718. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2013.
- [71] William Jockusch, James Propp, and Peter Shor. Random Domino Tilings and the Arctic Circle Theorem, 1998. arXiv:math/9801068.
- [72] Kurt Johansson. Non-intersecting paths, random tilings and random matrices. *Probab. Theory Relat. Fields*, 123(2):225–280, 2002.
- [73] Kurt Johansson. The arctic circle boundary and the Airy process. *Ann. Probab.*, 33(1):1–30, 2005.
- [74] Kurt Johansson and Eric Nordenstam. Eigenvalues of GUE Minors. *Electron. J. Probab.*, 11(0):1342–1371, 2006.
- [75] Pieter W. Kasteleyn. The statistics of dimers on a lattice: I. the number of dimer arrangements on a quadratic lattice. *Physica*, 27:1209–1225, 1961.
- [76] Pieter W. Kasteleyn. Dimer statistics and phase transitions. *J. Mathematical Phys.*, 4:287–293, 1963.
- [77] Kristian D. Kennaway. Brane tilings. *Int. J. Mod. Phys. A*, 22(18):2977–3038, 2007.
- [78] R. Kenyon. The Laplacian and Dirac operators on critical planar graphs. *Invent. Math.*, 150(2):409–439, 2002.
- [79] Richard Kenyon. Local statistics of lattice dimers. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(5):591–618, 1997.

- [80] Richard Kenyon. Conformal invariance of domino tiling. *Annals of Probability*, 28:759–795, 1999.
- [81] Richard Kenyon. Lectures on dimers. In *Statistical mechanics*, volume 16 of *IAS/Park City Math. Ser.*, pages 191–230. Amer. Math. Soc., Providence, RI, 2009.
- [82] Richard Kenyon and Andrei Okounkov. Planar dimers and Harnack curves. *Duke Math. J.*, 131(3):499–524, 2006.
- [83] Richard Kenyon and Andrei Okounkov. Limit shapes and the complex Burgers equation. *Acta Math.*, 199(2):263–302, 2007.
- [84] Richard Kenyon, Andrei Okounkov, and Scott Sheffield. Dimers and amoebae. *Ann. Math.*, 163(3):1019–1056, 2006.
- [85] Igor M. Krichever. Amoebas, Ronkin function, and Monge-Ampère measures of algebraic curves with marked points. In *Topology, geometry, integrable systems, and mathematical physics*, volume 234 of *Amer. Math. Soc. Transl. Ser. 2*, pages 265–278. Amer. Math. Soc., Providence, RI, 2014.
- [86] Nikolai Kuchumov. Limit shapes for the dimer model, 2017. arXiv:1712.08396.
- [87] Arno BJ Kuijlaars. Matrix valued orthogonal polynomials arising from hexagon tilings with  $3 \times 3$ -periodic weightings. *Journal of Approximation Theory*, 311:106202, 2025.
- [88] Tri Lai and Gregg Musiker. Beyond Aztec castles: toric cascades in the  $dP_3$  quiver. *Comm. Math. Phys.*, 356(3):823–881, 2017.
- [89] Tri Lai and Gregg Musiker. Dungeons and dragons: combinatorics for the  $dP_3$  quiver. *Ann. Comb.*, 24(2):257–309, 2020.
- [90] Megan Leoni, Gregg Musiker, Seth Neel, and Paxton Turner. Aztec castles and the  $dP_3$  quiver. *J. Phys. A*, 47(47):474011, 32, 2014.
- [91] Diane Maclagan and Bernd Sturmfels. *Introduction to tropical geometry*, volume 161 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2015.
- [92] Matthew Nicoletti. Local statistics and shuffling for dimers on a square-hexagon lattice. *Ann. Inst. Henri Poincaré Probab. Stat.*, 61(4):2776–2803, 2025.
- [93] Bernard Nienhuis. Critical behavior of two-dimensional spin models and charge asymmetry in the Coulomb gas. *J. Stat. Phys.*, 34(5-6):731–761, 1984.
- [94] Andrei Okounkov. Symmetric functions and random partitions. In *Symmetric functions 2001: surveys of developments and perspectives*, volume 74 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 223–252. Kluwer Acad. Publ., Dordrecht, 2002.
- [95] Andrei Okounkov and Nicolai Reshetikhin. The birth of a random matrix. *Mosc. Math. J.*, 6(3):553–566, 588, 2006.
- [96] Andrei Okounkov and Nicolai Reshetikhin. Random skew plane partitions and the Pearcey process. *Comm. Math. Phys.*, 269(3):571–609, 2007.

- [97] Andrei Okounkov and Nikolai Reshetikhin. Correlation function of Schur process with application to local geometry of a random 3-dimensional Young diagram. *J. Amer. Math. Soc.*, 16(3):581–603, 2003.
- [98] Robin Pemantle, Mark C. Wilson, and Stephen Melczer. *Analytic combinatorics in several variables*, volume 212 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2024. Second edition [of 3088495].
- [99] T. Kyle Petersen and David Speyer. An arctic circle theorem for Groves. *J. Combin. Theory Ser. A*, 111(1):137–164, 2005.
- [100] Leonid Petrov. Asymptotics of random lozenge tilings via Gelfand-Tsetlin schemes. *Probab. Theory Related Fields*, 160(3-4):429–487, 2014.
- [101] Mateusz Piorkowski. Arctic curves of periodic dimer models and generalized discriminants. 2024. arXiv:2410.17138.
- [102] James Propp. Generalized domino-shuffling. *Theoret. Comput. Sci.*, 303(2-3):267–301, 2003.
- [103] Günter Rote, Francisco Santos, and Ileana Streinu. Pseudo-triangulations—a survey. In *Surveys on discrete and computational geometry*, volume 453 of *Contemp. Math.*, pages 343–410. Amer. Math. Soc., Providence, RI, 2008.
- [104] Philippe Ruelle. Double tangent method for two-periodic Aztec diamonds. *J. Stat. Mech. Theory Exp.*, (12):Paper No. 123103, 44, 2022.
- [105] Meredith Shea. Split Two-Periodic Aztec Diamond. *Ann. Henri Poincaré*, 2026.
- [106] David E. Speyer. Perfect matchings and the octahedron recurrence. *J. Algebraic Combin.*, 25(3):309–348, 2007.
- [107] H. N. V. Temperley and Michael E. Fisher. Dimer problem in statistical mechanics—an exact result. *Philos. Mag. (8)*, 6:1061–1063, 1961.
- [108] William P. Thurston. Conway’s tiling groups. *Amer. Math. Monthly*, 97(8):757–773, 1990.
- [109] Craig A. Tracy and Harold Widom. The Pearcey process. *Comm. Math. Phys.*, 263(2):381–400, 2006.
- [110] G. Wulff. Zur Frage der Geschwindigkeit des Wachstums und der Auflösung der Kristallflächen. *Z. Krystallogr. Mineral.*, 34:449–530, 1901.
- [111] Sicong Zhang. Cluster Variables and Perfect Matchings of Subgraphs of the  $dP_3$  Lattice, 2015. arXiv:1511.06055.