Université catholique de Louvain

Institut de recherche en
mathématique et physique

# Exploration of the arctic curve phenomenon in lattice statistical models via the tangent method

Bryan Debin

Thesis submitted in partial fulfilment of the requirements for the degree of Docteur en Sciences

**Dissertation committee**

| | | |
|---|---|---|
| Prof. Philippe Ruelle | *(Supervisor)* | UCLouvain |
| Prof. Jan Govaerts | *(President)* | UCLouvain |
| Prof. Tom Claeys | *(Secretary)* | UCLouvain |
| Prof. Filippo Colomo | | INFN, Sezione di Firenze |
| Prof. Philippe Di Francesco | | Université Paris-Saclay and University of Illinois |
| Prof. Christian Hagendorf | | UCLouvain |

October 2021

## Remerciements

Je tiens tout d'abord à adresser mes plus vifs remerciements à mon promoteur, le Professeur Philippe Ruelle, pour m'avoir transmis le goût pour la physique statistique et la combinatoire. Cette affinité, née durant les magnifiques cours qu'il dispense, n'a fait que se développer durant mes quatre années de doctorat jusqu'à finir par se transformer en cette thèse. Sa patience, sa disponibilité, son expertise et sa bonne humeur inébranlable ont permis la réalisation de cette thèse dans les meilleures conditions.

Je remercie également mes lecteurs, les professeurs Christian Hagendorf, Jan Govaerts, Tom Claeys, Philippe Di Francesco et Filippo Colomo qui ont accepté de consacrer du temps à la lecture de mon manuscrit et dont les remarques ont permis d'améliorer cette thèse. Je les remercie également de m'avoir chacun exposé à des idées passionnantes, soit par le biais de cours ou—de façon plus indirecte—par articles interposés. Je remercie les professeurs Nauts, Bricmont et Ringeval pour leurs enseignements, qui m'ont également fait forte impression.

Cette thèse est le fruit de nombreuses collaborations. Je remercie vivement Philippe Ruelle, Philippe Di Francesco, Emmanuel Guitter, Etienne Granet ainsi que Jean-François de Kemmeter d'avoir collaboré avec moi pour explorer l'univers riche et mystérieux des courbes arctiques.

De nombreuses personnes fascinantes ont croisé mon chemin durant ces quatre années de thèse à l'IRMP. Je remercie Adrien Poncelet d'avoir porté à mon attention la méthode de la tangente; ce fut un tournant décisif, comme le titre de cette thèse peut en attester. Merci à Antoine Doeraene de m'avoir initié aux joies de la génération de pavages aléatoires. Durant mes—désormais nombreuses—années d'étude, Gilles P. n'était jamais très loin. Je le remercie pour ses jeux de mots, son humour littéral, les soirées et les pauses café—dont la plupart restent inachevées à ce jour ! Je le remercie également, à parts égales avec Gilles B., d'avoir partagé et nourri mon enthousiasme pour la physique. Merci à Alexandre L. pour nos conversations « télescopantes », pour sa bonne humeur et pour les cookies. Je tiens aussi à remercier Jean de m'avoir fait découvrir les meilleurs passages de la Bible, pour son exploitation impitoyable de ma crédulité et de m'avoir fait découvrir les spécialités de son Italie natale, telle que sa fameuse ≪ couronne de choux ≫. Merci à Sandrine pour son énergie débordante et pour sa méticuleuse défense de la langue française face aux anglicismes qui la menace, en particulier aux heures tardives. Je tiens à remercier les doctorants, post-doctorants et secrétaires de l'IRMP pour toutes les conversations et les litres de café partagés : Abel, Alexi, Antoine, Corentin, Elia, Fathi, Florence, Lancelot, Léo, Cathy, Carine et Martine. Enfin, puisque l'UCLouvain ne se limite pas à l'IRMP, je remercie Lilian pour notre quasi-cohabitation, Koffi pour son Togo-gin et Raphael ≪ papa ≫ pour sa motivation à réunir les physiciens de tous horizons !

Durant ces années de doctorat, j'ai eu la chance de participer à divers colloques et écoles doctorales qui m'ont permis de rencontrer de jeunes chercheurs, auxquels j'adresse mes remerciements : I would like to thank Noah, Sam and Yuan, for the Florentine beers, Pavel, Etienne and Alex for the mind-bending riddles, Jason, Jules, Jonathan and Stefano for the hikes in les Houches (which, I am afraid, have irreversibly damaged my knee !), Louis for his daily culinary reviews of les Houches' menu and finally the Balkan club: Lenart, Katja and Tony as a honorary member.

Si la monotonie ne s'est jamais installée dans ma vie de doctorant, je le dois aussi à Lilian, Arthur, Anthony, Mathilde, Juliette, Caroline, Alexandre D., Charlotte, Laurence et Lara qui ont été bien plus que des colocataires. Merci pour les soirées jeux de société, les murder parties, pour la patience d'écouter mes anecdotes scientifiques de ≪ père castor ≫ et plus généralement pour l'excellente ambiance. Merci aussi à mes éternels camarades Arlonais, Tom, Zahi, Antoine et Valentin qui m'ont apporté un soutien indispensable durant ces quatre années—ainsi que durant une bonne part de celles qui les ont précédées.

Enfin, je remercie ma famille : Maman et Papa, merci pour votre amour, pour les encouragements et pour vous être efforcé de rendre ma vie la plus belle possible. Merci aussi à Mamy, à ma sœur, à Louise et au reste de ma famille pour leur bienveillance.

Lauranne, merci pour ton indéfectible soutien, pour ton enthousiasme communicatif et pour tous ces voyages que tu as rendu si beaux. Tu rends le quotidien magnifique !

À mon grand-père, Raymond Debin.

# Contents

## List of Abbreviations

**6V** six-vertex

**20V** twenty-vertex

**AD** Aztec diamond

**AF** anti-ferroelectric

**ASM** alternating sign matrices

**BCH** Baker-Campbell-Hausdorff

**CDF** cumulative distribution function

**D** disordered

**DBC** detailed balance condition

**DPP** descending plane partitions

**DV** all diagonal and vertical

**DWBC** domain wall boundary conditions

**E** empty

**EFP** emptiness formation probability

**FE** ferroelectric

**GBC** global balance condition

**GOE** gaussian orthogonal ensemble

**GUE** gaussian unitary ensemble

**H** all horizontal

**HD** all horizontal and diagonal

**HDV** all horizontal, diagonal and vertical

**HTSASM** half-turn symmetric alternating sign matrices

**LGV** Lindström-Gessel-Viennot

**NILP** non-intersecting lattice paths

**pDWBC** partial domain wall boundary conditions

**PDF** probability density function

**QISM** quantum inverse scattering method

**QTHADT** quarter turn symmetric holey Aztec domino tiling

**V** all vertical

**VSASM** vertically symmetric alternating sign matrices

# Introduction

## Motivations

The intricacies of fundamental interactions between elementary particles or massive objects are respectively understood within the framework of quantum physics (more specifically quantum field theory) and general relativity. Statistical mechanics is the study of the collective behaviour of a large number of objects, stemming from their individual interactions. Figuratively, if quantum physics and general relativity respectively study the infinitely small and the infinitely large, statistical mechanics studies the *infinitely many*. In view of the formidable complexity of fundamental interactions, this task seems hopeless. Fortunately, when a large number of constituents is involved, most details of these interactions are irrelevant and very crude models suffice to correctly predict collective macroscopic behaviours.

A prototypical example of statistical system is a ferromagnet in the absence of an external magnetic field like for example a piece of iron. Its macroscopic magnetic properties result from the interactions between the magnetic moments of its constituent atoms. The *Ising model* is a simplified version of this system. At each site $i = 1, \cdots, N$ of a given regular lattice resides a classical spin $\sigma_i$ that can only point in two opposite directions, it is either up $\sigma_i = +$ or down $\sigma_i = -$. When the system is in thermal equilibrium with its environment at a temperature $T$, the probability of observing a microscopic configuration $\sigma = \{\sigma_i\}_{i=1,\cdots,N}$ is given by

$$P(\sigma) = \frac{\mathrm{e}^{-\beta E(\sigma)}}{Z}, \quad E(\sigma) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j, \quad Z = \sum_{\sigma} \mathrm{e}^{-\beta E(\sigma)},$$

where $\beta^{-1} = k_B T$ with $k_B$ the Boltzmann constant. The energy of a configuration $E(\sigma)$ only involves pairs $\langle i, j\rangle$ of nearest neighbours and favours aligned spins when the constant describing the strength of the interaction is positive $J > 0$, which we now assume. For the Ising model, the choice of *boundary conditions* is not important in the thermodynamic limit $N \to \infty$. The *partition function* $Z$ plays a central role in statistical mechanics as it allows to compute various quantities of interest. At sufficiently low temperature, energetic considerations prevail and spins align so as to produce a macroscopic magnetic field, even in the absence of an external one. At high temperature, this collective behaviour is destroyed by thermal fluctuations. In dimension strictly greater than one, the transition between these limiting regimes occurs abruptly at a critical temperature $T = T_c$ in what is called a phase transition. In dimension two, the model is *exactly solvable* and is understood in much details. In particular, one can exactly compute the critical exponents that characterize the abrupt changes of behaviour of various physical quantities, such as the magnetization for example. These exponents are *universal* in the sense that including additional interaction terms[1] in $E(\sigma)$ (e.g. next-nearest-neighbours interactions) do not change their value, as understood in the framework of the renormalization group. Hence, the study of the (three-dimensional) Ising model is directly relevant to understand the critical properties of real-world materials undergoing a ferromagnetic phase transition.

In addition to being a typical example of statistical lattice model showing a phase transition, the Ising model also exhibits a *limit shape phenomenon*, though it manifests itself in a peculiar way in this instance [1]. Let us consider the two-dimensional Ising model on a large square grid with a fixed and large number $m$ of down spins. The boundary of the domain is considered sufficiently far away so as to have no influence on the following discussion. We might think of this set-up as a model of an equilibrium two-dimensional crystal. Up to an irrelevant additive factor, the energy of a configuration $E(\sigma)$ is proportional to the length of the interface separating up and down spins. When $T < T_c$, driven by

[1] These terms should be sufficiently small and also irrelevant in the renormalization group sense.

the tendency to minimize the length of the interface, down spins accumulate[2] into a large cluster we may call the Ising crystal. Its average shape, which depends on $T$, is exactly known [3], see [4] for nice simulations. When $T = 0$, the system strictly minimizes the energy and hence the Ising crystal strictly minimizes its (lattice) perimeter: it is then a unit square, once distances are rescaled by $\sqrt{m}$. If $m$ is not a perfect square i.e. $m = (\lceil \sqrt{m} \rceil)^2 - p$ with $p > 0$, the crystal cannot be a square on the lattice and numerous configurations minimize the energy. The $p$ "missing sites" accumulate around the four corners of the unit square where they exactly form Young diagrams each composed of $A \sim p/4$ unit squares.

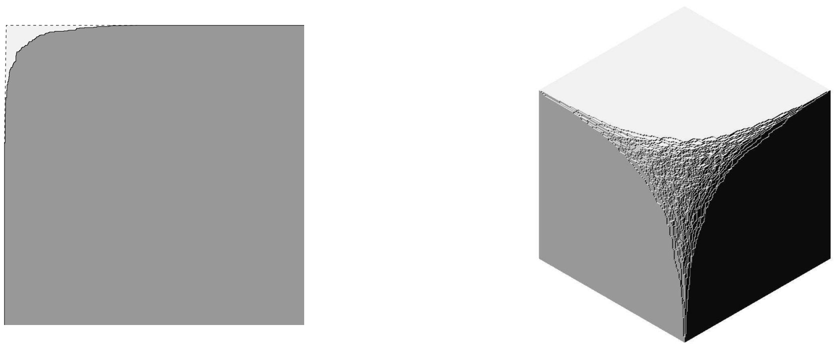


Figure 1: *Left:* Zoom on a corner of a two-dimensional Ising crystal, in the zero-temperature limit. Dark grey regions contain down spins. The missing forming Young diagrams are drawn in light grey. *Right:* Corner of a three-dimensional Ising crystal, in the zero-temperature limit, adapted from [1].

One might wonder what happens when the number of missing sites gets large. When $A \to \infty$ and once distances are rescaled by $\sqrt{A}$, almost all large Young diagrams composed of $A$ unit squares have exactly the same shape, see Figure 1-left. The curve toward which tends the (microscopically) random boundary of the Young diagram constitutes an example of one-dimensional *limit shape*. Its precise determination can be achieved via a *variational approach*. Roughly speaking the idea consists in estimating—using coarse graining arguments—the number of

[2] In dimension $d$, condensation occurs when $m$ is of order $V^{\frac{d}{d+1}}$ with $V$ the volume of the domain [2].

Young diagrams that have the same shape. The limit shape is the one corresponding to the maximal number of microscopic configurations. The problem is then reduced to finding the extremum of a functional. The above considerations can be extended to the three-dimensional Ising model. In the zero-temperature limit and after rescaling, the Ising crystal is then a cube. Again, in a typical situation, the corners of this cube are not perfect on the lattice [5], see Figure 1-right. The random surface that forms at a corner[3] might be described in terms of a discrete *height function*, which again converges to a limit shape. Its computation can again be rephrased in a variational framework [6] as the determination of the *two-dimensional surface* corresponding to the maximum number of discrete *height functions*. A novel feature appearing in this two-dimensional limit shape is the formation of *frozen* and of *disordered* regions. The two phases are spatially separated by an interface which becomes sharper and sharper when the number of missing sites is increased. This interface is called an *arctic curve* and could be seen as some kind of *spatial phase transition*. Limit shapes are found in a large variety of models, including tiling models [6–9], large Young tableaux [10], sandpile models [11, 12], alternating sign matrices [13] and vertex models [14–17]. Computing the functional that has to be minimized, given the definition of a model requires some work and has been realised for dimer models on planar bipartite graphs [6,18], and its computation for other models remains an active field of research [16,19]. Even when the functional is known, its extremization for given boundary conditions may prove challenging.

## Outline

This thesis is devoted to the study of arctic curves. We consider various exactly solvable two-dimensional lattice models whose configurations may be described in terms of *non-crossing lattice paths*. These non-crossing paths may be seen as discrete level curves of a height function, but we will build our understanding of the arctic phenomenon solely in the paths picture. Recently Colomo and Sportiello proposed a novel approach to compute arctic curves based on this path description called

[3] Allowed configurations of this surface are in bijection with plane partitions.

the *tangent method* [20]. This method was checked in a number of situations where it indeed reproduced known arctic curves [20–25]. It has been subsequently used to make predictions in cases where the analytic shape of the arctic curve was not known, see [17,20,22,25–30]. The object at the heart of this method is the partition function of the model, or rather a slightly more general *refined* version of it. A statistical physicist might not be surprised that relevant information might be extracted from a partition function but the implementation of such a program requires some insight, which will turn out to be geometrical in its nature, as suggested by the name of the method. The crux of the problem is hence to compute refined partition functions. This task consists in the evaluation of the (possibly weighted) enumeration of non-crossing paths and is therefore a problem that is essentially combinatorial.

In Chapter 1, we further motivate the study of arctic curves by considering the historical example of domino tilings and show that contrary to the Ising model, boundary conditions play a crucial role. In particular we consider the model that launched the study of the arctic phenomenon, namely domino tilings of the Aztec diamond [7,31].

In Chapter 2, we apply the tangent method to compute the arctic curves of two different tiling models. We first consider fixed boundary conditions on a domain whose shape tends to a rectangle in the scaling limit but nonetheless contains a lot of freedom, as its definition includes a piecewise differentiable function that can be chosen arbitrarily. This model is especially well suited for the application of the tangent method. The boundary conditions of the second model mix fixed and periodic boundary conditions. It can be seen as a domino tiling of an Aztec-diamond-like domain having a central hole on which a $90°$ rotation symmetry is imposed on the configurations.

Chapter 3 is devoted to numerical simulations of statistical models. Arctic curves can be seen as spatial phase transitions and as ordinary phase transitions they exhibit some form of *universality*. Whereas the shape of the arctic curve heavily depends on the model under consideration and on the chosen boundary conditions, fluctuations are believed—and sometimes proved—to be described by universal distributions, the most famous being the Tracy-Widom distribution. We numerically compute various statistics around a particular point of the arctic curve of the Aztec diamond and compare them with known results and conjectures.

Chapter 4 is dedicated to the application of the tangent method to two instances of another important class of two-dimensional statistical models, namely vertex models. The first one is the free-fermionic six-vertex model with partial domain wall boundary conditions that mix fixed and free boundaries. The second model is the twenty-vertex model with domain wall boundary conditions of type 2, this time not limited to the free-fermion case. Following an idea dating back to Baxter, we extensively exploit the integrable structure of this model to compute the required refined partition functions.

The tangent method is very efficient, but as of today, its validity has been proved only in one particular case [21] and hence remains not fully rigorous. In Chapter 5, we investigate the two main underlying assumptions of the tangent method in a variational framework. We give a compelling argument explaining the tangency assumption at the heart of the method. Moreover, we formulate a factorization conjecture which relates the asymptotic behaviour of refined partition functions and the corresponding arctic curves. Using exact lattice results, this conjecture is verified in full generality for domino tilings of the Aztec diamond and also for particular refined enumerations of alternating sign matrices. Finally we use the factorization conjecture to give a novel formulation of the tangent method.

## List of publications

This thesis gathers results published in the following articles:

- B. Debin and P. Ruelle, *Tangent method for the arctic curve arising from freezing boundaries*, J. Stat. Mech.: Theory Exp. **2019 12** (2019) 123105.
- B. Debin, E. Granet and P. Ruelle, *Concavity analysis of the tangent method*, J. Stat. Mech.: Theory Exp. **2019 11** (2019) 113107.
- B. Debin, P. Di Francesco and E. Guitter, *Arctic curves of the twenty-vertex model with domain wall boundaries*, J. Stat. Phys. **179** (2020) 33–89.
- B. Debin and P. Ruelle, *Factorization in the multirefined tangent method*, arXiv preprint `arXiv:2105.02257 [math-ph]` (2021), accepted in J. Stat. Mech.: Theory Exp.

Another publication by the author of this thesis was not included:

- B. Debin and E. Granet, *Stroboscopic exclusion process: A first-moment–driven dynamics*, EPL **133** (2021) 40002.

# Chapter 1

# Tiling models

This chapter is based on (unpublished) joint work with Philippe Ruelle.

---

The purpose of this chapter is to gain a first insight into tiling models and introduce the notion of arctic curve in this context. Our introductory example of random surface at a corner of the three-dimensional Ising crystal (see Figure 1-right) was already a tiling model in disguise [18]. Tiling models were originally introduced by physicists [32] and are relevant in describing surface adsorption of molecules, as some experimental set-ups provide rather faithful realizations of such models [33]. They also naturally arise in chemistry when studying the Lewis structure of graphene [34].

We define a tiling of a finite domain to be the complete covering of this domain using non-overlapping tiles. The tiles can be of finitely many types. We do not allow rotations, so that two tiles that differ only by their orientations are considered of different types. In this chapter and from now on we focus more specifically on tilings of domains composed of $1 \times 1$ unit squares using rectangular $1 \times 2$ or $2 \times 1$ tiles called *dominoes* and occasionally we will also insert $1 \times 1$ squares named *monomers*. Such a domain may be conveniently represented by a subgraph of $\mathbb{Z}^2$ which we (slightly abusively) call its dual graph and whose vertices are the centre of the unit squares linked by an edge for adjacent squares. As exemplified on Figures 1.1 or A.1, there is an obvious bijection between tilings of a domain and *perfect matchings* on its dual graph, namely subsets of edges which cover every vertex exactly once. The term "dimer" sometimes refers exclusively to an edge of a perfect matching [35], but we will also use it interchangeably with "domino".

When studying tiling models, a sequence of increasingly complicated questions naturally arises. Given a domain, an obvious first question concerns the tileability of this domain. Exposing a valid tiling is sufficient to ensure tileability. Likewise, the graph-theoretical version of Hall's marriage theorem [36] ensures that if a region is not tileable by dominoes, one can always find a simple counting argument to prove the non-tileability. Namely, if we colour the unit squares black and white in a chequerboard fashion, one can always find a set of black unit squares having too few or too many adjacent white squares, hence proving the non-tileability by virtue of each domino covering a white and a black square. Finding a tiling or such a set of black squares may be hard. There exists a general algorithm due to Thurston that determines tileability based on the height function representation [37], see also [38] for a neat exposition in the case of lozenge tilings. Secondly, if a domain is tileable one could wonder what the number of tilings is. By itself, this question actually constitutes a vast field of research, see for example [39–47]. There is a plethora of tools to tackle this problem, some of which were developed very recently [48] while others are more ancient. Amongst those is the celebrated Kasteleyn matrix [40], that can be used to enumerate perfect matchings on any planar bipartite graph or even some non-planar bipartite graphs after some tweaking. Depending on the problem at hand, some tools might lead to significantly simpler computations. In this chapter, we use two different tools, namely the *transfer matrix* and the *Lindström-Gessel-Viennot lemma*, to enumerate tilings of two different domains. In doing so we expose the crucial importance of *boundary conditions* and introduce the notion of arctic curve. This brings about even more questions. It is natural to try and compute the shape of arctic curves and study fluctuations around it. Another question concerns the way of generating tilings at random according to a prescribed probability distribution. These topics are addressed in the following chapters.

## 1.1 Rectangle

When it comes to tilings of finite domains, the problem of domino tilings of a rectangular domain is amongst the first example that comes to mind and was also one of the first studied [40, 41, 49]. If the width and the height are simultaneously odd, the rectangle is not tileable since each domino occupies two unit squares. If at least one edge has an even length, the domain is tileable by placing strips of aligned dominoes parallel to this edge. Without loss of generality we consider a $2n \times m$ rectangle with $n, m \in \mathbb{N}$. For simplicity we further restrict ourselves to *even* $m$, as the odd case is nearly identical. We consider a slightly more general problem with a weight $\alpha \in \mathbb{R}_*^+$ for horizontal dimers. The weight of vertical dimers can be set to 1, without loss of generality. Indeed, a weight for vertical dimers can be absorbed in a redefinition of $\alpha$, since the total number $mn$ of dimers is fixed. In this section, we compute the partition function

$$Z_{2n,m}^{\mathrm{rec}} = \sum_{\tau} \alpha^{h(\tau)}, \tag{1.1}$$

where $h(\tau)$ is the number of horizontal dimers in the tiling $\tau$. We also evaluate the probability to observe a horizontal dimer at an arbitrary location. The main tool we use is a transfer matrix introduced by Lieb [39], but the computations are also tractable by the Kasteleyn matrix approach [50]. We use a special case of the Baker-Campbell-Hausdorff (BCH) formula discussed in [51] to evaluate powers of the transfer matrix.

### 1.1.1 Transfer matrix

The transfer matrix is an ubiquitous tool in the study of two-dimensional classical statistical systems [39, 52–54]. It is an operator that acts on vectors of a Hilbert space $\mathcal{H}$. The core idea of this approach is to restrict the two-dimensional configuration to a *fixed row* and map it to a given vector $|v\rangle \in \mathcal{H}$. The matrix $T$ is constructed in such a way that $T|v\rangle$ is a linear combination of all vectors corresponding to allowed configurations on the *next row*. The coefficients of the linear combination are the associated statistical weights. In the case at hand, the coefficients involve a power of $\alpha$ which counts the number of horizontal dimers. Likewise, the power $T^n$ iteratively constructs all possible arrays of $n$ rows built atop $|v\rangle$ which satisfy the rules prescribed by the model. The transfer matrix is at the core of the profound connection between instances of two-dimensional classical statistical systems and one-dimensional quantum systems [54, 55]. This relationship is used in [56] to relate some tiling models and quantum quenches.

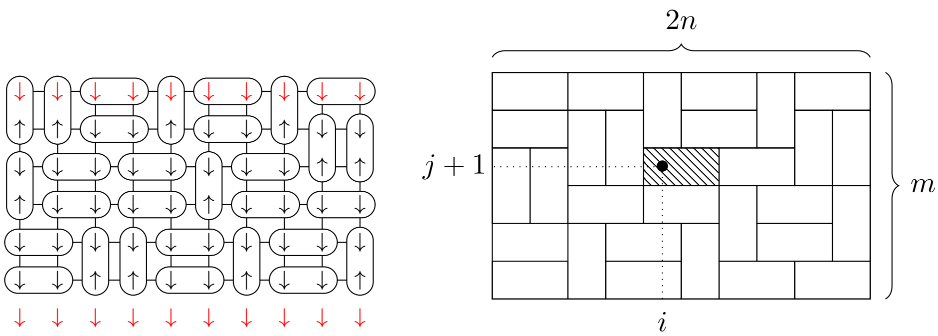



Figure 1.1: *Left*: Illustration of the bijection between tilings of the rectangle, perfect matchings of the dual graph and arrow configuration at the vertices of this graph. The red arrows on the top row are fixed by the boundary conditions while the ones on the lower supplementary row are added to impose the boundary conditions. *Right*: Tiling of the $2n \times m$ rectangle having a horizontal dimer with a left corner at position $(i, j+1)$.

There are several ways to construct the transfer matrix for domino tilings [56, 57]. We follow Lieb's original approach [39]. Each row of dimers is mapped to a string of up and down arrows $s_i \in \{\uparrow, \downarrow\}$ called spins at each site $i = 1, \cdots, 2n$. A site $i$ is assigned an up arrow if a vertical dimer covers $i$ and its top neighbour. It is assigned a down arrow in all other cases. The correspondence is shown on Figure 1.1. One readily notes that two neighbouring down arrows could correspond to a horizontal dimer covering two neighbouring sites on a row, or else to the top of two vertical dimers side by side. However, this ambiguity is lifted by looking at the preceding row and eventually to the bottom boundary. A row is mapped to the canonical basis vector $|s\rangle = |s_1 s_2 \cdots s_{2n}\rangle \equiv |s_1\rangle \otimes |s_2\rangle \otimes \cdots \otimes |s_{2n}\rangle$ of the Hilbert space $(\mathbb{C}^2)^{\otimes 2n}$, where $|\uparrow\rangle = \binom{1}{0}$ and $|\downarrow\rangle = \binom{0}{1}$.

Let us now show how to impose the boundary conditions in this formalism. The left and right boundary conditions will be built directly in $T$. To ensure that dimers do not stick out of the bottom boundary, we place below it a full row of horizontal dimers, see Figure 1.1. $T|\downarrow\downarrow\ldots\downarrow\rangle$ indeed produces all the allowed states on the first row, correctly weighted. After applying $m$ times $T$ we arrive at the top row whose arrows must all be down due to the top boundary. Hence, once an appropriate $T$ is constructed, the partition function can be computed as

$$Z^{\text{rect}}_{2n,m} = \langle\downarrow\downarrow\ldots\downarrow|T^m|\downarrow\downarrow\ldots\downarrow\rangle\,. \tag{1.2}$$

Notice that this approach can be adapted for domains whose shape is not rectangular, such as the one of the next section. The idea is to impose arrow orientations on the top and bottom boundaries in such a way as to *deterministically* freeze regions of the rectangular domain, see for example [56].

It remains to define $T$ and show that it acts as required. In the canonical basis, it is expressed in terms of operators $\sigma_i^a$ with $a \in \{x, y, z, +, -\}$ acting only on a site $i$, defined by the tensor product

$$\sigma_i^a = \underbrace{\mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_2}_{i-1} \otimes \sigma^a \otimes \underbrace{\mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_2}_{2n-i}, \tag{1.3}$$

where $\mathbb{I}_2$ is the $2\times 2$ identity matrix and the usual Pauli matrices are given explicitly by

$$\begin{gathered}\sigma^x = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad \sigma^y = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \qquad \sigma^z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\\ \sigma^- = \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \quad \sigma^+ = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}.\end{gathered} \tag{1.4}$$

The matrix $T$ will generally act on linear combinations of row configurations, but being linear, one may consider its action on a single row configuration, noted $|s\rangle$. Following the notation of [39], we express $T$ as the product of two operators $V_1$ and $V_3$. Acting on $|s\rangle$, the first operator $V_1$ reverses all the arrows, and thereby makes sure that an up arrow in $|s\rangle$ becomes a down arrow in $|s'\rangle = V_1|s\rangle$, on the next row up. It is given by

$$V_1 = \prod_{i=1}^{2n} \sigma_i^x. \tag{1.5}$$

The second operator $V_3$ acts on $|s'\rangle$ by producing a linear combination of *all the states* obtained by flipping (or not) pairs of neighbouring up spins of $|s'\rangle$, with the appropriate coefficient. At the dimer level, the flip of pairs of up arrows transforms two vertical dimers into a single horizontal one. For fixed (straight) boundary conditions on the left and right boundaries it is given by

$$V_3 = \prod_{i=1}^{2n-1} (\mathbb{I} + \alpha\, \sigma_i^- \sigma_{i+1}^-). \tag{1.6}$$

The ordering of this product has no importance. The property $(\sigma_i^-)^2 = 0$ allows to write $V_3$ as an exponential. The transfer matrix is hence

$$T = V_3 V_1 = \exp\left(\alpha \sum_{i=1}^{2n-1} \sigma_i^- \sigma_{i+1}^-\right) \prod_{i=1}^{2n} \sigma_i^x. \tag{1.7}$$

Although it is possible to deal with $T$ itself, its square

$$T^2 = \exp\,(\alpha \sum_{i=1}^{2n-1} \sigma_i^- \sigma_{i+1}^-)\, \exp\,(\alpha \sum_{i=1}^{2n-1} \sigma_i^+ \sigma_{i+1}^+) \tag{1.8}$$

is simpler and sufficient for what we want to do. We start by block-diagonalizing it. The first trick is to use the Jordan-Wigner transform so as to trade the matrices $\sigma_i^\pm$ for new operators, defined as

$$C_i = (-1)^{i-1}\, (\prod_{k=1}^{i-1} \sigma_k^z)\, \sigma_i^-\,, \qquad C_i^\dagger = (-1)^{i-1}\, (\prod_{k=1}^{i-1} \sigma_k^z)\, \sigma_i^+\,, \tag{1.9}$$

for $1 \leqslant i \leqslant 2n$. Notice that this transformation is non-local, $C_i$ and $C_i^\dagger$ no longer act on site $i$ alone. Their algebra follows easily from the algebra of the Pauli matrices,

$$\{C_i, C_j\} = \{C_i^\dagger, C_j^\dagger\} = 0, \qquad \{C_i, C_j^\dagger\} = \mathbb{I}\, \delta_{i,j}\,. \tag{1.10}$$

Because they satisfy anticommutation relations, they are referred to as fermionic operators. From (1.9) and by using $\sigma^z \sigma^- = -\sigma^-$, we easily obtain $C_i\, C_{i+1} = -\sigma_i^-\, \sigma_{i+1}^-$, implying

$$T^2 = \exp\Big(-\alpha \sum_{i=1}^{2n-1} C_i C_{i+1}\Big)\, \exp\Big(\alpha \sum_{i=1}^{2n-1} C_i^\dagger C_{i+1}^\dagger\Big). \tag{1.11}$$

The operators appearing in the exponentials are coupled throughout the whole row via the nearest neighbour interactions. A substantial decoupling is achieved by introducing new fermions via Fourier transformation,

$$\begin{aligned} C_j &= \sqrt{\frac{2}{2n+1}}\, \mathrm{e}^{\mathrm{i}\gamma} \sum_{k=1}^{2n} \mathrm{i}^{j+k} \, \sin \frac{\pi k j}{2n+1}\, \eta_k \,, \\ C_j^\dagger &= \sqrt{\frac{2}{2n+1}}\, \mathrm{e}^{-\mathrm{i}\gamma} \sum_{k=1}^{2n} (-\mathrm{i})^{j+k} \, \sin \frac{\pi k j}{2n+1}\, \eta_k^\dagger \,, \end{aligned} \tag{1.12}$$

with $\mathrm{e}^{2\mathrm{i}\gamma} = (-\mathrm{i})^{2n} = (-1)^n$. This transformation is unitary, which ensures that the new fermionic operators $\eta, \eta^\dagger$ satisfy the same algebra as the $C, C^\dagger$,

$$\{\eta_k, \eta_\ell\} = \{\eta_k^\dagger, \eta_\ell^\dagger\} = 0, \qquad \{\eta_k, \eta_\ell^\dagger\} = \mathbb{I}\, \delta_{k,\ell} \,. \tag{1.13}$$

The quadratic forms in (1.11) become

$$\sum_{i=1}^{2n-1} C_i\, C_{i+1} = -2 \sum_{a=1}^{n} \cos q_a \; \eta_a \, \eta_{2n+1-a} \,, \qquad q_a = \frac{\pi a}{2n+1}, \tag{1.14}$$

and recast $T^2$ into

$$T^2 = \exp\Big\{ 2\alpha \sum_{a=1}^{n} \cos q_a \; \eta_a \, \eta_{2n+1-a} \Big\} \exp\Big\{ 2\alpha \sum_{a=1}^{n} \cos q_a \; \eta_{2n+1-a}^\dagger \, \eta_a^\dagger \Big\}. \tag{1.15}$$

The main advantage of this new form is that the $\eta, \eta^\dagger$ are no longer all coupled together but are coupled within small groups $\{\eta_a, \eta_{2n+1-a}, \eta_a^\dagger, \eta_{2n+1-a}^\dagger\}$. In particular, since the quadratic terms $\eta_a^{(\dagger)}\, \eta_{2n+1-a}^{(\dagger)}$ and $\eta_b^{(\dagger)}\, \eta_{2n+1-b}^{(\dagger)}$ commute for $a \neq b$, we can also write

$$T^2 = \prod_{a=1}^{n} \exp\Big\{ 2\alpha \cos q_a \eta_a \, \eta_{2n+1-a} \Big\} \exp\Big\{ 2\alpha \cos q_a \; \eta_{2n+1-a}^\dagger \, \eta_a^\dagger \Big\}. \tag{1.16}$$

In the fermionic basis $\{(\eta_1^\dagger)^{m_1} (\eta_2^\dagger)^{m_2} \cdots (\eta_{2n}^\dagger)^{m_{2n}} |0\rangle \,|\, m_i = 0, 1\}$ where $|0\rangle = |{\downarrow\downarrow} \cdots {\downarrow}\rangle$ (satisfying $\eta_k |0\rangle = 0$), the matrix $T^2$ is block-diagonal. Each term in the product acts in the four-dimensional space $\mathcal{H}_a \equiv (\mathbb{C}^2)_a \otimes (\mathbb{C}^2)_{2n+1-a}$. Moreover, all terms in the product commute and $T^2$ can therefore be written as $T^2 = \bigotimes_{a=1}^{n} A_a$ with

$$A_a = \exp\Big\{ \gamma_a \; \eta_a \, \eta_{2n+1-a} \Big\} \; \exp\Big\{ \gamma_a \; \eta_{2n+1-a}^\dagger \, \eta_a^\dagger \Big\}, \tag{1.17}$$

where $\gamma_a = 2\alpha \cos q_a$. In the case of the rectangle, the vectors that enforce the boundary conditions are naturally translated into the fermionic basis $|0\rangle = |{\downarrow\downarrow} \cdots {\downarrow}\rangle$. For other fixed boundary conditions, their decomposition into this basis involves a determinant. If one of the boundary vector differs form $|0\rangle$ at $O(n)$ sites, this usually leads to difficulties when computing asymptotic results unless this determinant admits a simple formula.

### 1.1.2 Baker-Campbell-Hausdorff

All blocks $A_a$ commute amongst themselves, so that higher powers of $T^2$ simply factor accordingly, $T^m = \bigotimes A_a^{m/2}$. However, the two exponential factors making up each $A_a$ do not commute, so that computing $A_a^m$ is not straightforward. The idea to use the BCH formula seems obvious and natural, but one soon realizes that in the present case, the infinite series of nested commutators will not stop at any finite order, making the explicit computation non-trivial if possible at all. However, the recent article [58], building on previous works [51], present new and explicit BCH formulae involving operators satisfying certain specific algebras. The case at hand, in (1.17), is covered by this scheme.

We follow the notations $X, Y, Z$ of [58] and omit for the moment the label $a$. We define

$$X = \gamma\, \eta_1 \eta_2, \quad Y = \gamma^2\, (\eta_1 \eta_1^\dagger - \eta_2^\dagger \eta_2) = \gamma^2\, (\mathbb{I} - N), \quad Z = \gamma\, \eta_2^\dagger \eta_1^\dagger. \tag{1.18}$$

The $X$ and $Z$ operators arise directly from (1.17) while $Y$ appear when computing their commutator. The operator $N = \eta_1^\dagger \eta_1 + \eta_2^\dagger \eta_2$ counts the number of excitations in the 4-dimensional Fock space. It is straightforward to compute the algebra satisfied by these three operators. One finds

$$[X, Y] = uX, \qquad [Y, Z] = uZ, \qquad [X, Z] = Y, \qquad u = -2\gamma^2. \tag{1.19}$$

The idea of [58] to rewrite $\exp X \cdot \exp Z$ as a single exponential is to use the results of [51] twice. First step relies on the observation that the couples $\{X, Y\}$ and $\{Y, Z\}$ form *two-dimensional algebras*. On general grounds, this implies that one write $\exp{(X)} \cdot \exp{(tY)}$ as the exponential of a linear combination of $X$ and $Y$, whose coefficients are functions of

$t$ and $u$, and similarly for $\exp(-tY)\cdot\exp(Z)$. The coefficients of these linear combinations were explicitly given in [51],

$$\begin{aligned} \exp(X)\cdot\exp(\frac{t}{u}Y) &= \exp\Big\{\frac{t}{u}Y + \frac{t}{1-\mathrm{e}^{-t}}X\Big\} \equiv \exp\widetilde{X}, \\ \exp(-\frac{t}{u}Y)\cdot\exp(Z) &= \exp\Big\{-\frac{t}{u}Y - \frac{t}{1-\mathrm{e}^{t}}Z\Big\} \equiv \exp\widetilde{Y}. \end{aligned} \tag{1.20}$$

The second step is to *fix* $t$ so that the commutator of $\widetilde{X}$ and $\widetilde{Y}$ *closes* on themselves. One finds

$$[\widetilde{X},\widetilde{Y}] = -t\widetilde{X} + t\widetilde{Y}, \tag{1.21}$$

provided the parameter $t$ satisfies $\cosh t = 1-\frac{u}{4}$. Since $\widetilde{X}$ and $\widetilde{Y}$ satisfy a two-dimensional algebra, one can use [51] once more to re-express $\exp(\widetilde{X})\cdot\exp(\widetilde{Y}) = \exp(X)\cdot\exp(Z)$ as a single exponential,

$$\begin{aligned} \exp(X)\cdot\exp(Z) &= \exp\Big\{\frac{t}{\sinh t}[X+Z+\tfrac{1}{2}Y]\Big\}, \\ t &= \operatorname{arccosh}(1-\frac{u}{4}) = \operatorname{arccosh}(1+\frac{\gamma^2}{2}). \end{aligned} \tag{1.22}$$

One may also check that the operators $X, Y, Z$ satisfy simple anticommutation relations, namely

$$\{X,Y\} = \{Z,Y\} = 0, \qquad \{X,Z\} = Y^2/\gamma^2. \tag{1.23}$$

They imply in particular, since $X^2 = Z^2 = 0$,

$$\big(X+Z+\tfrac{1}{2}Y\big)^2 = (\frac{1}{4}+\frac{1}{\gamma^2})Y^2 = \Big(\sinh t\cdot\frac{Y}{\gamma^2}\Big)^2. \tag{1.24}$$

This last identity allows to write

$$\big[\exp X\,\exp Z\big]^k = \cosh\big[kt(\mathbb{I}-N)\big] + \frac{\sinh\big[kt(\mathbb{I}-N)\big]}{\sinh t\cdot(\mathbb{I}-N)}\big(X+Z+\tfrac{1}{2}Y\big). \tag{1.25}$$

After an appropriate identification of the parameters ($\gamma_a = 2\alpha\cos q_a$), it allows to write the $m$-th power of the transfer matrix as $T^m = \prod_a A_a^{m/2}$, where

$$\big(A_a\big)^{m/2} = \cosh\big[\frac{m}{2}t_a(\mathbb{I}-N_a)\big] + \frac{\sinh\big[\frac{m}{2}t_a(\mathbb{I}-N_a)\big]}{\sinh t_a\cdot(\mathbb{I}-N_a)}\big(X_a+Z_a+\tfrac{1}{2}Y_a\big), \tag{1.26}$$

with $X_a = \gamma_a \eta_a \eta_{2n+1-a}$, $Y_a = \gamma_a^2 \,(\eta_a \eta_a^\dagger - \eta_{2n+1-a}^\dagger \eta_{2n+1-a}) = \gamma_a^2 \,(\mathbb{I} - N_a)$, and $Z_a = \gamma_a \,\eta_{2n+1-a}^\dagger \eta_a^\dagger$, for $1 \leqslant a \leqslant n$. We note a number of useful identities on the functions of $t_a$ appearing in the previous formula.

The definition $t = \operatorname{arccosh}(1 + \frac{\gamma^2}{2})$ implies $\mathrm{i}t = \arccos(1 + \frac{\gamma^2}{2})$, so that

$$\frac{\sinh \frac{m}{2} t}{\sinh t} = \frac{\sin \mathrm{i} \frac{m}{2} t}{\sin \mathrm{i}t} = \frac{\sin \left[\frac{m}{2} \arccos(1 + \frac{\gamma^2}{2})\right]}{\sin \left[\arccos(1 + \frac{\gamma^2}{2})\right]} = U_{\frac{m}{2}-1}(1 + \frac{\gamma^2}{2}), \quad (1.27)$$

where $U_n(x)$ is the Chebyshev polynomial of the second kind. These polynomials satisfy $2xU_n(1 - 2x^2) = (-1)^n U_{2n+1}(x)$, and replacing $x \to \mathrm{i}x$, we obtain

$$2xU_{\frac{m}{2}-1}(1+2x^2) = (-\mathrm{i})^{m-1} U_{m-1}(x) = \frac{(x+\sqrt{x^2+1})^m - (x-\sqrt{x^2+1})^m}{2\sqrt{x^2+1}}. \quad (1.28)$$

Setting $x = \alpha \cos q_a$ (or $2x = \gamma_a$) yields directly

$$\beta_{m,a} \equiv \gamma_a \frac{\sinh \frac{m}{2} t_a}{\sinh t_a} = \frac{\mu_a^m - \mu_a^{-m}}{\mu_a + \mu_a^{-1}}, \quad (1.29)$$

where $\mu_a \equiv \sqrt{1 + \alpha^2 \cos^2 \frac{\pi a}{2n+1}} + \alpha \cos \frac{\pi a}{2n+1}$. Likewise, we have

$$\cosh \tfrac{m}{2} t = \cos \mathrm{i} \tfrac{m}{2} t = \cos \left[\tfrac{m}{2} \arccos(1 + \tfrac{\gamma^2}{2})\right] = T_{\frac{m}{2}}(1 + \frac{\gamma^2}{2}). \quad (1.30)$$

Using the identities for the Chebyshev polynomial of first type $T_n$ similar to the ones for the $U_n$, one readily finds

$$\alpha_{m,a} \equiv \cosh \tfrac{m}{2} t_a + \frac{\gamma_a^2}{2} \frac{\sinh \frac{m}{2} t_a}{\sinh t_a} = \frac{\mu_a^{m+1} + \mu_a^{-(m+1)}}{\mu_a + \mu_a^{-1}}. \quad (1.31)$$

**Partition function.** Combining all of these results we get the partition function

$$\begin{aligned} Z_{2n,m}^{\text{rect}} &= \langle 0 | T^m | 0 \rangle = \langle 0 | \prod_{a=1}^{n} \left\{ \cosh\left(\frac{m t_a}{2}\right) \mathbb{I} + \frac{\gamma_a^2}{2} \frac{\sinh\left(\frac{m t_a}{2}\right)}{\sinh t_a} \mathbb{I} \right\} | 0 \rangle \\ &= \prod_{a=1}^{n} \alpha_{m,a} = \prod_{a=1}^{n} \left[ \frac{\mu_a^{m+1} + \mu_a^{-m-1}}{\mu_a + \mu_a^{-1}} \right]. \end{aligned} \quad (1.32)$$

The best-known form for this partition function involves a double product. Let us show that we recover such a formula by using the identity (already used in [40])

$$\begin{aligned}&\prod_{b=1}^{m}\left[4x^2+4\cos^2\left(\frac{\pi b}{m+1}\right)\right]^{1/2}=\prod_{b=1}^{m/2}\left[4x^2+4\cos^2\left(\frac{\pi b}{m+1}\right)\right]\\&=\frac{1}{2\sqrt{x^2+1}}\left[(x+\sqrt{1+x^2})^{m+1}-(x-\sqrt{1+x^2})^{m+1}\right].\end{aligned}\tag{1.33}$$

To prove this identity (for even $m$) simply notice that the l.h.s. and the r.h.s. are polynomials in $x$ of degree $m$ with an identical leading coefficient and the same zeros. If we set $x=\alpha\cos q_a$ we recover the well–known formula [40],

$$Z^{\text{rect}}_{2n,m}=\prod_{a=1}^{n}\prod_{b=1}^{m}\left[4\alpha^2\cos^2\left(\frac{\pi a}{2n+1}\right)+4\cos^2\left(\frac{\pi b}{m+1}\right)\right]^{1/2}.\tag{1.34}$$

A similar analysis as above shows that the exact same formula is also valid for odd $m$. Let us comment on this result, which has some interesting features. Some properties that are obvious in the tiling picture are not directly apparent in (1.34). For example, $Z^{\text{rect}}_{2n,m}$ must be a polynomial in $\alpha$ whose coefficients are integers. It is also not hard to see that for $\alpha=1$ and $n=1$, $Z^{\text{rect}}_{2n,m}$ is the $(m+1)$-th Fibonacci number. Conversely, some properties that are not obvious in the tiling picture can be accessed using (1.34). In the case of the square $m=2n$ for $\alpha=1$, one can show that $Z^{\text{rect}}_{2n,m}$ is either a perfect square or double a perfect square [43]. This can be understood combinatorially [43] and also follows from Ciucu's factorization theorem [44], which is a general result for perfect matchings on graphs that possesses a reflection symmetry. This theorem indeed implies that $Z^{\text{rect}}_{2n,2n}=2^n b_n^2$ where $b_n$ enumerates tilings of a domain obtained from the square by cutting it along its diagonal following a $2\times 2$ zig-zag. It is quite remarkable that $b_n$ is (conjecturally) closely related to the *twenty-vertex model* [59] that is discussed in Chapter 4. Finally, using (1.32) we compute (see for example Appendix B of [60]) the *number of arrangements* per dimer $z$ defined by

$$\log(z)=\lim_{m,n\to\infty}\frac{1}{nm}\log(Z^{\text{rect}}_{2n,m}).\tag{1.35}$$

Roughly speaking we could say that if each domino was independent from the others and had a partition function $z$ then the global partition function would be $z^{mn}$ asymptotically. In the case $\alpha = 1$, we get $z = \mathrm{e}^{\frac{2G}{\pi}} = 1.7916228..$ where $G = 1 - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \dots$ is the *Catalan constant*. Notice that $z < 2$ due to the hard-core interaction between dimers.

**Dimer density.** So far we enumerated tilings. We now take a tiling $\tau$ at random according to the probability distribution $P(\tau) = \alpha^{h(\tau)}/Z^{\text{rect}}_{2n,m}$, where $h(\tau)$ is the number of horizontal dimers in $\tau$ and $\alpha > 0$. Keeping the $\alpha$ dependence in (1.35) enables one to compute the average number of horizontal dimers per dimer which is $\rho_h = \alpha\, \partial_\alpha \log(z) = 2/\pi \arctan(\alpha)$. In the case $\alpha = 1$ this leads to $\rho_h = 1/2$. It seems to indicate that there is no preferred orientation in most of the rectangle. We finish our analysis of rectangle tilings by verifying this assertion.

The information given by the partition function on the typical features of a random tiling is limited. We will show in detail how some of these features can be extracted from a *refinement* of the partition function. For now however, we make some explicit computations using the BCH approach in the case of the rectangle *with* $\alpha = 1$ *and* $m$ *even*. Let us evaluate the probability of observing a horizontal dimer whose left unit square is centred at $(i, j+1)$, see Figure 1.1. Applying the operator $\sigma_i^+$ on the line $j$ effectively creates a monomer at $(i, j+1)$. It indeed kills all configurations of line $j$ that would cover the monomer and ensures that a vertical dimer cannot be put with its lower square at $(i, j+1)$. Equivalently—and perhaps more intuitively—a monomer can be created at $(i, j+1)$ by applying $\sigma_i^-$ on line $j+1$. Hence,

$$P^{\boxed{\bullet\ }}_{i,j+1} = \frac{\langle 0|T^{m-j}\sigma_i^+\sigma_{i+1}^+ T^j|0\rangle}{\langle 0|T^m|0\rangle}. \tag{1.36}$$

We temporarily set $j$ to be even. The denominator reads

$$\begin{aligned} \langle 0|T^m|0\rangle &= \langle 0|\prod_{a=1}^{n}(\alpha_{m-j,a} + \frac{\beta_{m-j,a}}{\gamma_a}\, X_a)(\alpha_{j,a} + \frac{\beta_{j,a}}{\gamma_a}\, Z_a)|0\rangle \\ &= \prod_{a=1}^{n}\{\alpha_{m-j,a}\alpha_{j,a} + \beta_{m-j,a}\beta_{j,a}\}. \end{aligned} \tag{1.37}$$

We now compute the numerator. Using $\sigma_i^+\sigma_{i+1}^+ = C_i^\dagger C_{i+1}^\dagger$ and (1.12) we get

$$
\begin{aligned}
&\langle 0|T^{m-j}\sigma_i^+\sigma_{i+1}^+ T^j|0\rangle \\
&\qquad = \frac{2(-1)^{n+i}}{2n+1}\sum_{k,\ell=1}^{2n}(-i)^{k+\ell+1}\sin\left(\frac{\pi k i}{2n+1}\right)\sin\left(\frac{\pi \ell(i+1)}{2n+1}\right)M_{k,\ell}.
\end{aligned}
\tag{1.38}
$$

There remains to compute the matrix element $M_{k,\ell} \equiv \langle 0|T^{m-j}\eta_k^\dagger\eta_\ell^\dagger T^j|0\rangle$. We first notice that $M_{k,\ell} = 0$ if $\ell \neq 2n+1-k$. The cases $k \leq n$ and $k > n$ are analysed separately in a similar way as (1.37). We get

$$
\begin{aligned}
P^{\boxed{\bullet\;}}_{i,j+1} = \frac{4}{2n+1}\sum_{k=1}^{n}\sin\left(\frac{\pi k i}{2n+1}\right)\sin\left(\frac{\pi k(i+1)}{2n+1}\right)& \\
\times\frac{\alpha_{j,k}\beta_{m-j,k}}{\alpha_{m-j,k}\alpha_{j,k}+\beta_{m-j,k}\beta_{j,k}}.&\quad (j\text{ even})
\end{aligned}
\tag{1.39}
$$

The case of odd $j$ can be treated similarly, using the identity $T\sigma_i^+ = \sigma_i^- T$ and $\sigma_i^-\sigma_{i+1}^- = -C_i^\dagger C_{i+1}^\dagger$. We get

$$
\begin{aligned}
P^{\boxed{\bullet\;}}_{i,j+1} = \frac{4}{2n+1}\sum_{k=1}^{n}\sin\left(\frac{\pi k i}{2n+1}\right)\sin\left(\frac{\pi k(i+1)}{2n+1}\right)& \\
\times\frac{\alpha_{m-j-1,k}\beta_{j+1,k}}{\alpha_{m-j-1,k}\alpha_{j+1,k}+\beta_{m-j-1,k}\beta_{j+1,k}}.&\quad (j\text{ odd})
\end{aligned}
\tag{1.40}
$$

Similar finite size formulae may be obtained for odd $m$. In Figure 1.2, we plot the exact density of horizontal dimers in finite size, using (1.39) and (1.40). Let us extract the large size behaviour at different points. We take $m, n \to \infty$ with $m/n$ fixed and compute the probability of having a horizontal dimer on the lower boundary or in the bulk. Using the Euler-Maclaurin formula and the Riemann-Lebesgue lemma, we get for $i, j = \mathcal{O}(n)$

$$
\begin{aligned}
\lim_{m,n\to\infty} P^{\boxed{\bullet\;}}_{i,j+1} &= \frac{1}{\pi}\int_0^{\pi/2}\frac{\cos(t)}{\sqrt{1+\cos^2(t)}}\,\mathrm{d}t = \frac{1}{4},\\
\lim_{m,n\to\infty} P^{\boxed{\bullet\;}}_{i,1} &= \frac{2}{\pi}\int_0^{\pi/2}\cos(t)\left[\sqrt{1+\cos^2(t)}-\cos(t)\right]\,\mathrm{d}t = \frac{1}{\pi}.
\end{aligned}
\tag{1.41}
$$

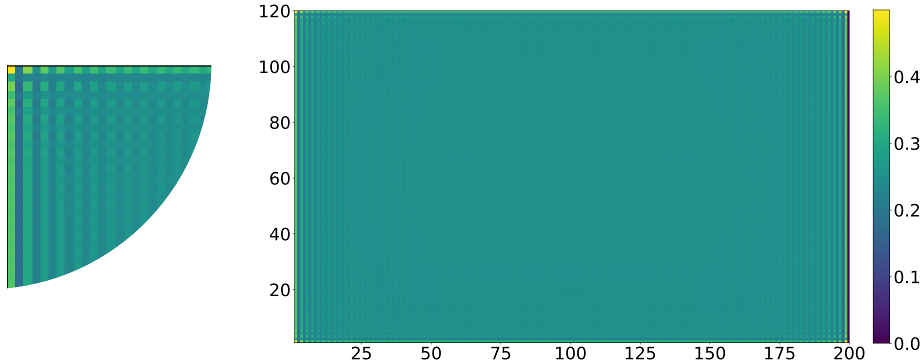



Figure 1.2: Exact density of horizontal dimers on the $200 \times 120$ rectangle, according to formulae (1.39) and (1.40). The density is homogeneous excepted for the oscillations near the boundary where the density oscillates with the parity of the site (see the zoom on the upper left corner). This result is in good agreement with Monte Carlo simulations.

From this last result we deduce that the probability of observing a horizontal dimer on the left or right edge is $1 - 2/\pi$ (far from the corners). This is a particular case of a result by Fisher and Stephenson [61]. We conclude that the boundary has an influence that quickly decays as soon as we get away from it by a macroscopic distance, in which case, the dimers do not have a preferred direction.

## 1.2 Aztec diamond

Boundary conditions were originally believed to play a minor role in the thermodynamic limit $n \to \infty$, as is the case for other classical 2d statistical models such as the Ising model. Kasteleyn provided substance to this statement in [40] by showing that the number of arrangements per dimer is $z = \mathrm{e}^{2G/\pi}$ for periodic boundary conditions as well, although he also stated that "*The effect of boundary conditions is, however, not entirely trivial and will be discussed in more detail in a subsequent paper.*". An obvious limitation is that some domains are not tileable. For example, if one mutilates the $2n \times 2n$ square by removing two unit squares at opposite corners, it is not tileable anymore. Similarly, it is not hard to devise a rectangular domain with two "crenellated" edges that only admits a single tiling. One may still think that "non-trivial" boundary conditions should behave similarly in the thermodynamic limit, at least far away from the boundary. Grensing, Carlsen and Zapp [62] were the first to question this belief by exhibiting boundary conditions such that the number of tilings grows with the size of the domain in a different way than the rectangular case. They found examples where this number grows exponentially with the linear size of the domain ($z = 1$), or with the area of the domain but with a different growth rate than the rectangle case. Their boundary conditions were designed to drastically reduce the number of tilings—while still allowing for some freedom—through the intermediary of their *rough boundaries*[1].

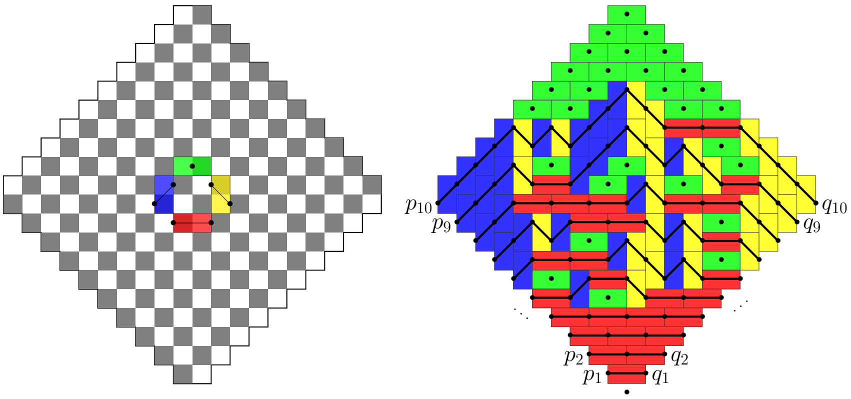



Figure 1.3: *Left*: Aztec diamond of order $n = 10$, whose unit squares are coloured black and white in a chequerboard fashion, in such a way that squares on the lower left edge are black. The four dominoes are distinguished according to the colours of the underlying squares and their orientation. Each of them is associated to different portions of lattice paths. *Right*: Tiling of an Aztec diamond of order $n = 10$. The corresponding NILP start at $p_i$ and end at $q_i$, with $i = 1, \cdots, n$ and are composed of elementary steps in $S = \{(1, 1), (1, -1), (2, 0)\}$.

In this section, we study one of these domains, later coined in [42] the *Aztec diamond*, due to it resemblance with Aztec art patterns. An Aztec diamond of order $n \in \mathbb{N}_0$ is defined as the domain formed by unit squares whose corners $(x, y)$ satisfy the condition $|x| + |y| \leq n + 1$. Figure 1.3 shows a tiling of the Aztec diamond (AD) with dominoes. A formula for the number $\mathrm{AD}_n$ of such tilings was originally conjectured in [62]

[1]In this context, "rough" means that, in the dual graph, vertices at the boundary have a coordinate number less than three [62].

and first proved in four different ways by Elkies, Kuperberg, Larsen and Propp in [42, 63]. It is astonishingly simple,

$$\mathrm{AD}_n = 2^{n(n+1)/2}. \tag{1.42}$$

There is a profusion of proofs, none of which as simple as the formula might suggest. These various proofs reflect the numerous ways of thinking about this particular tiling problem. One of them is the transfer matrix formalism that can be used to establish this result [56]. Two of the original proofs of [42, 63] are based on the connection with the *shuffling algorithm* and alternating sign matrices (ASM) respectively discussed in chapters 3 and 5. Another proof due to Ciucu [45] exploits a general theorem on perfect matchings of "cellular graphs" of which the Aztec diamond is an instance. Another approach uses a general result on perfect matching called "graphical condensation" [64] a particular case of which can be deduced by applying Dodgson[2] condensation to a Kasteleyn matrix [65]. In this section, we take yet another point of view and exploit the bijection with some non-intersecting lattice paths (NILP). A quite simple proof [66] can be achieved using the NILP picture. We provide another proof similar to that of [22] that is slightly longer but conveniently introduces tools and results used throughout this thesis.

### 1.2.1 Lindström-Gessel-Viennot lemma

In domino tiling problems it is often a good idea to colour the underlying unit squares in white and black in a chequerboard manner. For example, the mutilated square mentioned above can be shown not to be tileable by counting the total number of white and black squares. Four types of dominoes can be distinguished according to the colour of the unit squares on which they lie on, as shown on Figure 1.3. For rectangular domains, we have shown that the dimer density is not fundamentally different for the two kinds of horizontal (or vertical) dominoes, except possibly in the close vicinity of the boundary. The situation turns out to be quite different for the Aztec diamond and it is crucial to make the distinction to fully apprehend the problem. The NILP are defined in terms of these four dominoes as shown on Figure 1.3. Notice that

[2] Charles Dodgson is better known under the pseudonym Lewis Carroll.

this construction is not unique and it might sometimes prove useful to exploit other NILP representations, as will be explained in Chapter 2.

For completeness we reproduce the proof of this bijection from [67]. Any tiling of the Aztec diamond of order $n$ corresponds to a NILP configuration with starting points $p_i$ and ending points $q_i$ $i = 1, \cdots, n$ such as in Figure 1.3. There are only two kinds of dimers that can cover black squares at the lower left edge, each of which has a starting point at the corresponding point $p_i$, and similarly for the white squares on the lower right boundary and the ending points $q_i$. Every elementary step starts on a black square and ends on a white square. This implies that each path is continuous, and they are non-intersecting by construction. Conversely, each NILP configuration is associated to a tiling. To see this, it suffices to first place the three kinds of dominoes that convey an elementary step. It is always possible due to the non-intersecting property and the allowed elementary steps. White squares that are not yet covered by a domino must have an empty black square to its right, since paths are continuous. Similarly, empty black squares have an empty white square to their left. Hence, all squares that were not yet covered can be covered by the remaining kind of domino. This finishes the proof of the bijection. Using the correspondence between tilings and NILP configurations, it is not hard to show that any tiling can be obtained by a sequence of *flips*  $\leftrightarrow$  or  $\leftrightarrow$  from any other tiling. It suffices to notice that any NILP configuration can be obtained by these moves starting from the "all-horizontal" NILP configuration.

There is a very elegant way of counting the number of NILP configurations. It was pioneered[3] by Lindström [70] and later rediscovered independantly by Gessel and Viennot in full generality [71], by Fisher [72] and in the context of combinatorial chemistry [73].

**Lemma 1.2.1.** *(Lindström-Gessel-Viennot (LGV) [71]) Let $G$ be a finite directed acyclic graph, and $A = \{p_1, \ldots, p_n\}$, $B = \{q_1, \ldots, q_n\}$ two (non-necessarily disjoint) sets of distinct vertices of $G$. Let also denote by $P_\sigma = (P_1, \ldots, P_n)$ a $n$-tuple of paths such that $P_i$ connects $p_i$ to $q_{\sigma(i)}$.*

[3]Even earlier Karlin and Mc-Gregor [68,69] exposed a similar idea in a probabilistic framework.

*If $A_{i,j}$ denotes the number of paths on $G$ from $p_i$ to $q_j$, then*

$$\det_{1\leqslant i,j\leqslant n}\left[A_{i,j}\right] = \sum_{\sigma\in S_n} \epsilon_\sigma \,|\{P_\sigma \text{ with } P_i\cap P_j=\emptyset\}|. \tag{1.43}$$

*In particular, if the paths of $P_\sigma$ necessarily intersect for $\sigma\neq id$, the determinant on the l.h.s. counts the number of NILP configurations connecting $A$ to $B$.*

Notice that in the above lemma, a path $P_i$ is a sequence of vertices, or stopping points, two successive vertices being separated by an elementary step. Two paths then intersect if and only if they share a common passage point. Two paths can therefore be non-intersecting even though their linear continuous interpolations do intersect. The proof of the LGV lemma is quite elegant, so much so that it appears in "the Book" [74]. Notice that this result is straightforwardly generalized to weighted paths.

### 1.2.2 Partition function

We work in the coordinate system in which the NILP start at $p_i=(-i,i)$ and end at $q_j=(j,j)$ for $i,j=0,1,\cdots,n$. Notice that a trivial path is placed at $p_0=q_0=(0,0)$ to prevent other paths to go there, which is forbidden by the shape of the domain. Because of this trick, there is no constraint to take into account when enumerating the number of paths from $p_i$ to $q_j$, which is hence given by

$$A_{i,j}=\sum_{p=0}^{\min(i,j)} w^p \binom{i+j-p}{p,\,i-p,\,j-p}, \tag{1.44}$$

where we assigned a weight $w$ to the elementary step $(2,0)$. In [66], the trivial path trick is not used and the authors have to deal with Schröder paths that are constrained not to cross a horizontal line. Any tiling of the Aztec diamond contains the same number of horizontal dimers of the two colours. This is easily shown by noticing that this tiling can be constructed from the "all-horizontal" NILP configuration via a sequence of flips. Hence, setting $w=\alpha^2$ is equivalent to weighing all horizontal dimers by $\alpha$. Before proceeding let us compute the bivariate generating function $f_A(x,y) \equiv \sum_{i,j=0}^{\infty} A_{i,j}\, x^i y^j$ of $A_{i,j}$:

$$f_A(x,y) = \sum_{p=0}^{\infty} \sum_{i,j=p}^{\infty} w^p \binom{i+j-p}{p, i-p, j-p} x^i y^j = \frac{1}{1-x-y-wxy}. \tag{1.45}$$

By the LGV lemma, the partition function $\mathrm{AD}_n$ is the determinant of the matrix $A = (A_{i,j})_{i,j=0,\cdots,n}$. Krattenthaler extensively discusses in [75,76] techniques to evaluate (i.e. provide a simple enough formula) a variety of determinants. One of them relies on the decomposition of the matrix into the product of a lower triangular matrix $L$ and an upper triangular matrix $U$ and is called an *$LU$ decomposition*. If the principal minors of the matrix are non-vanishing–which is the case of $A$, as seen using the LGV lemma–this decomposition is unique [77] when the elements on the diagonal of $L$ are set to 1. Once the matrix $U$ is found, the determinant of $A$ is nothing but the product of the diagonal elements of $U$. This technique turns out to be particularly helpful for our purpose as it also provides a natural way of computing refined partition functions [22,23,25,27]. As is often the case when computing determinants, a first step is to guess the result by performing LU decomposition for small $n$. In our case this allows to conjecture

$$\begin{aligned}
&U_{i,j} = (1+w)^i \binom{j}{i}, \quad U^{-1}_{i,j} = (-1)^{i-j} (1+w)^{-j} \binom{j}{i}, \qquad \text{for} \quad i \leqslant j,\\
&L_{i,j} = \binom{i}{j}, \quad L^{-1}_{i,j} = (-1)^{i-j} \binom{i}{j}, \qquad \text{for} \quad i \geqslant j,
\end{aligned} \tag{1.46}$$

with all other entries equal to zero. In order to prove these conjectures, it is convenient to use generating functions. For a series or a polynomial $f(x) = \sum_{k\geq 0} f_k\, x^k$, we denote by $f(x)|_{x^k} \equiv f_k$ the coefficient of $x^k$ in $f(x)$. Two lemmas will be particularly useful.

**Lemma 1.2.2.** *Let $f(x)$ and $g(x)$ be two generating functions. We have*

$$f(x)g(x)\big|_{x^k} = \sum_{l=0}^{k} f(x)\big|_{x^l}\, g(x)\big|_{x^{k-l}}. \tag{1.47}$$

**Lemma 1.2.3.** *The binomial coefficient can be represented in four different ways by generating series or polynomial*

$$\binom{n}{k} = (1+x)^n\big|_{x^k} = (1+x)^n\big|_{x^{n-k}} = \left.\frac{1}{(1-x)^{k+1}}\right|_{x^{n-k}} = \left.\frac{1}{(1-x)^{n-k+1}}\right|_{x^k}. \tag{1.48}$$

Both lemmas are straightforward to prove. Let us first show that $L^{-1}$ is indeed the inverse le $L$. We have

$$\begin{aligned}
(LL^{-1})_{i,j} &= \sum_{k=0}^{n} L_{i,k}L^{-1}_{k,j} = \delta_{i\geq j}\sum_{k=j}^{i}(-1)^{k-j}\binom{i}{k}\binom{k}{j} \\
&= \delta_{i\geq j}\sum_{k=0}^{i-j}(-1)^k\binom{i}{i-j-k}\binom{k+j}{j} \\
&= \delta_{i\geq j}\sum_{k=0}^{i-j}(1+x)^i\big|_{x^{i-j-k}}\ \left.\frac{1}{(1+x)^{j+1}}\right|_{x^k} \\
&= \delta_{i\geq j}\,(1+x)^{i-j-1}\big|_{x^{i-j}} = \delta_{i,j},
\end{aligned} \tag{1.49}$$

where $\delta_{i\geq 0} = 1$ if $i \geq j$ and $\delta_{i\geq 0} = 0$ otherwise. The proof for the inverse of $U$ is similar. To complete the proof of (1.46), it remains to show that $A = LU$, or equivalently $L^{-1}A = U$. Using the generating function (1.45), we have

$$\begin{aligned}
(L^{-1}A)_{i,j} &= \sum_{k=0}^{i}(-1)^{i-k}\binom{i}{k}A_{k,j} = \sum_{k=0}^{i}(1-x)^i\big|_{x^{i-k}}\ \left.\frac{1}{1-x-y-wxy}\right|_{x^k y^j} \\
&= \left.\frac{(1-x)^i}{1-x-y-wxy}\right|_{x^i y^j} = (1-x)^{i-1}\left.\left(\frac{1+wx}{1-x}\right)^j\right|_{x^i} \\
&= \delta_{i\leq j}\ \left.\frac{(1+wx)^j}{(1-x)^{j-i+1}}\right|_{x^i} = \delta_{i\leq j}\sum_{k=0}^{\infty}\binom{j-i+k}{k}x^k\sum_{p=0}^{j}\binom{j}{p}x^p w^p\Big|_{x^i} \\
&= \delta_{i\leq j}\binom{j}{i}\sum_{p=0}^{i}\binom{i}{p}w^p = \delta_{i\leq j}\,(1+w)^i\binom{j}{i} = U_{i,j}.
\end{aligned} \tag{1.50}$$

This concludes the proof of (1.46). The partition function is now easily evaluated

$$\mathrm{AD}_n = \det_{0\leq i,j\leq n}(A_{i,j}) = \prod_{i=0}^{n} U_{i,i} = (1+w)^{n(n+1)/2}. \tag{1.51}$$

This result indeed reproduces the one first proved in [42, 63]. It implies that the weighted number of arrangements per dimer is

$$\log(z) = \lim_{n\to\infty} \frac{1}{n(n+1)} \log(\mathrm{AD}_n) = \log(\sqrt{1+w}). \tag{1.52}$$

In the case $w = 1$ we get $z = \sqrt{2} = 1.4142135..$ which is less than the number of arrangements in the rectangular case where we found $z = 1.7916228...$ By a symmetry argument or using (1.52) one can prove that the average density of horizontal dimers is $\rho_h = 1/2$, hence the discrepancy between the two domains cannot be explained by a homogeneous preferred orientation. It is also not a consequence of a smooth variation of dimer density throughout the domain. It is rather the result of a *spatial phase separation* phenomenon in the Aztec diamond whereby in the overwhelming majority of tilings, some regions of the domain become *frozen* while others fluctuate and are sometimes called *temperate* regions. This tendency, already visible in Figure 1.3, strengthens as the domain gets large as seen on Figure 1.4 so that in the scaling limit, when $n \to \infty$ and distances are rescaled by $n$, the frozen and temperate regions almost surely have deterministic shapes. The interface between these two regions is called the *arctic curve*. In the case of the Aztec diamond, it was rigorously proved by Jockusch, Propp and Shor[4] [31], that in the scaling limit the interface approaches the *circle* that is inscribed into the domain. More precisely stated [31], if we fix $\epsilon > 0$, then for all sufficiently large $n$, all but an $\epsilon$ fraction of the domino tilings of the Aztec diamond of order $n$ will have a temperate region whose boundary stays uniformly within distance $\epsilon n$ of the inscribed circle. In the case $w \neq 1$, the arctic curve becomes an ellipse. In addition to this phenomenon, Figure 1.4 suggests that inside the disordered region, the density of dimers is not homogeneous. Not only does the probability of observing a horizontal dimer depend on the rescaled coordinates, but it also drastically varies depending on the parity of the lattice coordinates. An exact formula was derived by Elkies, Cohn and Propp in [7] and re-derived by Allegra in [56] using the transfer matrix formalism.

[4]Peter Shor is well known for the eponymous quantum algorithm of prime number decomposition.

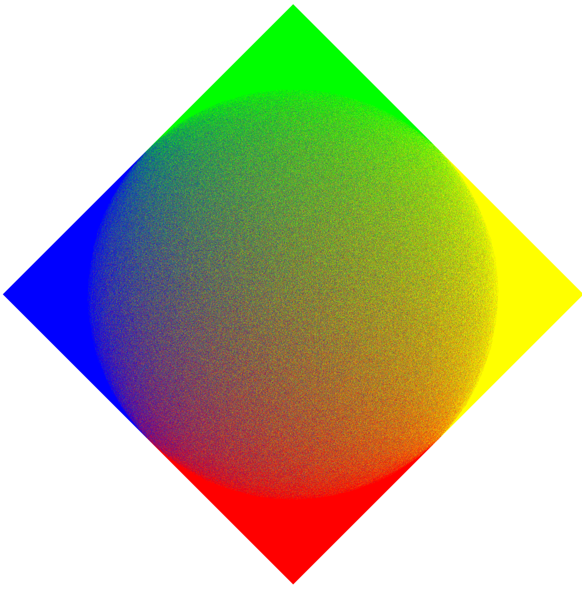


Figure 1.4: Tiling of the Aztec diamond of order $n = 10000$ sampled from the uniform distribution using the Shuffling algorithm.

Let us now give some heuristics for the arctic phenomenon. The boundary is such that imposing the orientation of a dimer on it might force the orientation of $O(n)$ other dimers, restricting the number of such configurations. For example, it is straightforward to show that the probability that the upper corner of the Aztec diamond of order $n$ is occupied by a vertical dimer is $2^{n(n-1)/2}/2^{n(n+1)/2} = 1/2^n$ and hence vanishes exponentially with increasing $n$. This kind of argument fails to explain the freezing observed deeper into the bulk. To understand this astonishing effect it is more appropriate to use the NILP representation. The arctic phenomenon is the result of two conflicting tendencies of NILP—and more generally of non-crossing paths—*in the scaling limit*. On the one hand, the number of microscopic configurations of a single path is maximal around the straight line that joins its starting point and its ending point (these considerations are made more quantitative in Chapter 5). On the other hand, the precise shape of a given path also constrains the other ones, due to their non-crossing nature. Hence, still in the scaling limit, the most likely trajectory for a path maximises the number of lattice configurations that are possible around this path, while letting enough space for the other paths to fluctuate. The influence of non-trivial weights is usually to deform arctic curves but may also lead to a degenerate situation, see Chapter 4. For definiteness, let us illustrate these considerations for the uppermost path. The closer the path gets to the upper boundary, the less are the constraints put on the other $n-1$ paths, at the price of greatly reducing the number of microscopic configurations for this path. Conversely, if it were to follow a straight (horizontal) trajectory this would maximize the number of compatible microscopic configurations, but would also maximize the constraints put on the other paths. The shape that is almost surely observed is an intermediary between these extremes. The abrupt transition occurring between the frozen and temperate region coincides with the outermost path(s) in some NILP description. Since the composition in the four types of dimers of a given portion of path depends on its slope, we may understand the variation of the density of the different kinds of dimers in the temperate region as the consequence of the variation of the slope of the paths in this region[5].

This heuristics explains the appearance of the arctic curve but does not provide a way of computing its precise shape. Despite its apparent simplicity, the tools used to answer this question are sometimes quite sophisticated [6, 9, 18, 35, 78–81]. The main objective of this thesis is to show how to use the so called *tangent method* to compute arctic curves using refined partition functions that are minimal modifications of partition functions where the boundary conditions are slightly perturbed. In the case of tiling models, once the LU decomposition of the LGV matrix is known, the 1-refined partition functions can be systematically computed [22, 23, 25, 27]. For vertex models, for which the tools mentioned above cannot typically be used (for general values of the statistical weights) the tangent method remains applicable and refined partition functions may be deduced from *inhomogeneous partition functions*, as explained in Chapter 4.

[5]The density of dimer of [7] can be used to compute the trajectory of any of the paths.

# Chapter 2

# Arctic curves in tiling models

This chapter uses material from [24] written with Philippe Ruelle and [17] written with Philippe Di Francesco and Emmanuel Guitter

## 2.1 Rectangoloid domain

The principal aim of this thesis is the understanding and the application of the tangent method introduced in [20] to compute arctic curves. We start this endeavour with the study of a particular tiling problem on a rectangoloid domain (see the definition hereafter) for which the analysis is particularly simple. The rectangoloid domain allows for much freedom in its definition as its scaling limit is described by a (nearly) *arbitrary real function* $\alpha(u)$, $u \in [0, 1]$. The existence of exact formulae despite the extensive freedom in the choice of $\alpha(u)$ is quite an unusual feature in the realm of tiling models. Indeed, just as the examples treated in the previous chapter, the majority of tractable tiling problems studied in the literature only have a handful of free parameters in their definitions, see e.g. [9, 22, 40, 42–47]).

In [23], the tangent method was used to derive *some portions* of the arctic curve of the rectangoloid tiling problem. Our contribution is the completion of their derivation for the remaining portions. It should be noted that the shape of the arctic curve was already determined fully rigorously in [82] for a particular $\alpha(u)$ ("freezing intervals" only) and then in [78,83] for a general $\alpha(u)$. The analysis in [78,82,83] goes beyond the computation of the arctic curve, at the cost of having to deal with advanced tools that render the computation of the arctic curve quite lengthy. In contrast, the tangent method derivation involves elementary tools and short computations, but does not provide anything else than the shape of the arctic curve itself. This computation can also be seen as a benchmark for the tangent method, which has been fully rigorously proved to work only in one particular case [21].

### 2.1.1 Definition of the model

We first define the model in terms of NILP contained in a quadrant, see Figure 2.1. The $n$ paths are labelled by an integer $i$ ranging from 1 to $n$. The $i$-th path starts from integer position $O_i = (a_i, 0)$ on the positive horizontal axis and ends on the vertical axis at $E_i = (0, i)$. An extra path, reduced to a single site, is conventionally added at the origin ($a_0 = 0$). Notice that the $a_i$'s are thought of as being *fixed* but can be *chosen arbitrarily*. Their distribution is described in the scaling limit by the function $\alpha(u)$ introduced above. The paths, between their starting and ending points, can only make two kinds of elementary steps, either north $(0, 1)$ or west $(-1, 0)$ and are constrained to be non-intersecting. This last condition implies that the $a_i$'s form a strictly increasing sequence, $0 < a_1 < a_2 < \ldots < a_n$. The collection of $n$ paths is therefore entirely contained in the rectangle of size $a_n \times n$. We are interested in NILP chosen at random amongst all of the allowed configurations. We do not consider position-dependent weights for the elementary steps, but the following discussion can be generalized to the area-weighted case [25] or even more generally to (area weighted) domino tilings of Aztec rectangles [27]. Since the number of elementary steps of each type is entirely determined by the set of $a_i$'s, we assign the same weight to each step, without loss of generality.

Each NILP configuration is associated with a tiling of a rectangoloid domain, as shown on Figure 2.1. It is composed of three types of tiles, called U, R and F, as they respectively correspond to the Upper, Right and Front faces of a cuboid drawn in perspective. The bijection is constructed by assigning U-tiles to west steps and R-tiles to north steps. The domain to be tiled is a rectangoloid, with triangular inclusions called defects on the lower boundary at $O_i$ and a sawtooth pattern on the left edge. The remaining spaces of the rectangoloid is filled by F-tiles. One can interpret the tiling as three-dimensional piles of cuboids[1] that are subject to some constraints dictated by the shape of the rectangoloid[2]. In those terms one may interpret the bijection with NILP as the intersection between cuboids and planes that are parallel to the sheet.

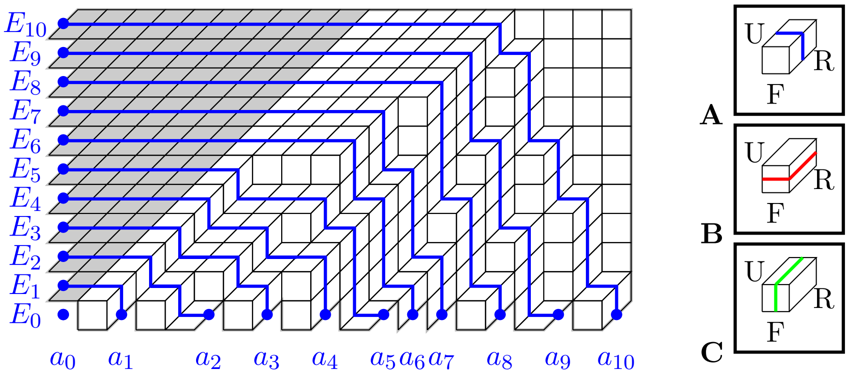



Figure 2.1: Particular configuration of NILP for the model under consideration and the associated tiling on a rectangoloid domain. Tiles drawn in grey are deterministically frozen. Insets **A**, **B** and **C** show the rule to translate a tiling in terms of NILP of type 1, 2 and 3, respectively.

We defined tilings on the rectangoloid domain in terms of a specific NILP description, that we call NILP of "type 1". The model admits two other equivalent definitions, in terms of NILP of "type 2" and of "type 3", as shown on Figure 2.1. An essential feature of this triple path description is that there is always one in terms of which a region regularly covered by one type of tile is *devoid of any path*. This observation is crucial when applying the tangent method and will be used throughout this thesis. In addition to being characterized by different elementary steps,

[1]Similarly, domino tilings can be seen as piles of Levitov blocks [84], see [85] for a nice illustration in the case of the Aztec diamond.

[2]Some completion is required around the defects.

the three types of paths also have different starting and ending points, see [23] and below for more details.

Due to the tightly packed arrangement of ending points $E_i$, there is a region that is *deterministically tiled* by a regular (frozen) pattern of U-tiles. More interestingly, *statistically frozen* regions also appear in the vast majority of tilings, in a similar way as in the Aztec diamond case. As already suggested by Figure 2.1, two frozen regions tend to form in the upper left[3] and upper right corners of the rectangle, each covered by only one type of tile, U or F. For a generic distribution of starting points $a_i$, no other frozen region appears. In the scaling limit, an arctic curve separates these two frozen regions from the entropic region. Its precise shape has been completely and explicitly determined in [23]; it only depends on the function $\alpha(u)$ that characterizes the distribution of the points $a_i$, in the scaling limit.

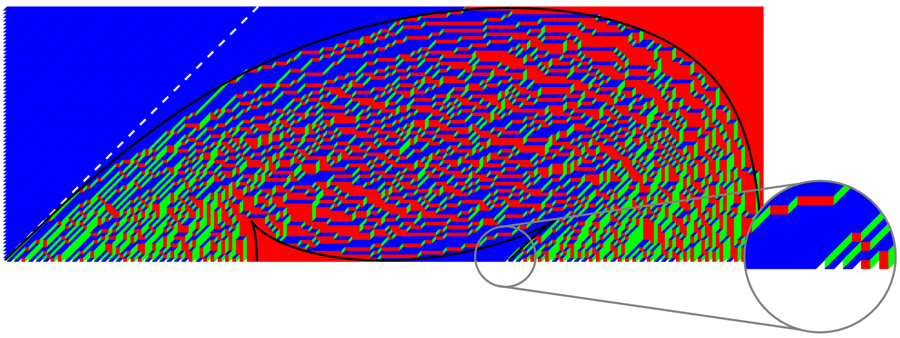


Figure 2.2: Random tiling of a rectangoloid domain with a freezing interval of first kind (macroscopic gap), with $n = 69$. The intervals on the left and on the right of the defect-free gap have the same length and contain one defect every second site (see the zoom). U-tiles are drawn in blue, F-tiles in red and R-tiles in green. The tiling was sampled using a generalization of the "domino-shuffling" exact sampling algorithm [86, 87] that allows for vanishing weights. The black curve is drawn using the parametric equation of (2.37). The white dashed line delimits the deterministically frozen region.

For non-generic distributions of $a_i$, more frozen regions may appear. In particular, this is the case when this distribution presents a *freezing interval*. These intervals can be of two types. In the tiling picture, either

[3]This statistically frozen region on the upper left corner extends the deterministically frozen one.

the interval is totally flat (no starting points) or it has a sawtooth pattern (starting points with maximal density). In both cases, a frozen region forms above the interval [23], as shown in Figures 2.2 and 2.3. This region is itself separated from the entropic bulk region by a new portion of arctic curve. Its parametric representation has been conjectured in [23] and derived via the tangent method in only one case, namely when the macroscopic interval has the sawtooth pattern and lies at one of the edges of the lower boundary.

Notice also that the qualitative features of a given portion of arctic curve may be understood in terms of the appropriate NILP description using the same heuristics as for the Aztec diamond.

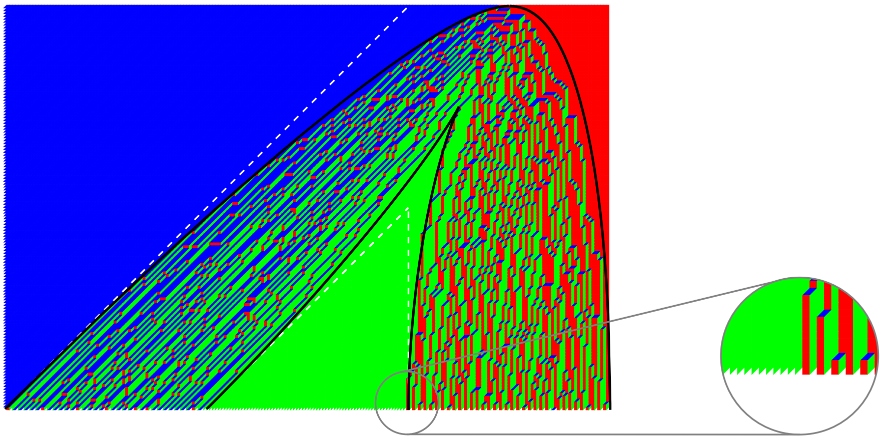


Figure 2.3: Random tiling of a rectangoloid domain with a freezing interval of second kind (sawtooth pattern) with $n = 140$. Two other intervals are adjacent to it and have one defect every two sites (see the zoom). The tiling was sampled using a Markov chain algorithm. The black curve is obtained from the parametric equation (2.37). The white dashed lines delimit the deterministically frozen regions.

### 2.1.2 Tangent method

The applicability conditions of the tangent method are not precisely known, but it is expected to be valid in models which can be formulated in terms of non-intersecting or osculating (allowed to touch) lattice paths. We consider a precise portion of arctic curve and the path description in terms of which the adjacent frozen region is empty. If we zoom on this portion of arctic curve, we find that it is composed of an accumulation of a large number of paths that fluctuate, in a way that is studied in Chapter 3. Let us call $x_0$ and $x_1$ their common starting and ending points in the rescaled domain. On the lattice, the outermost path would start from site $p_0$ and go to site $p_1$, in the scaling neighbourhood of $x_0$ and $x_1$ respectively.

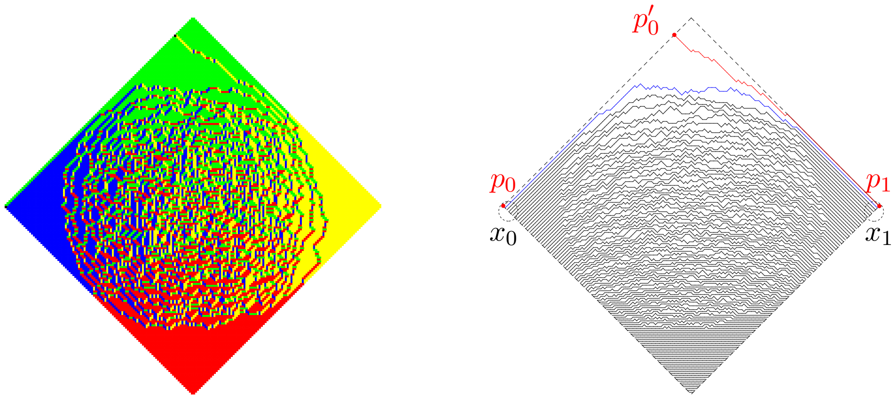



Figure 2.4: Tangent method principle. *Left*: Tiling of an Aztec diamond with two monomers (in black). *Right*: Corresponding path configuration, with the modified path in red and the one directly below it in blue.

The tangent method requires to slightly modify the model in such a way that the starting (or ending) point of the outermost path is moved from $p_0$ to another point $p'_0$ located on the boundary at a macroscopic distance from $p_0$, adjacently to the (empty) frozen region, see Figure 2.4. The specific way this change is implemented depends on the model under consideration. In one of the best-known cases, namely the Aztec diamond, one can achieve this by inserting two monomers, as shown in Figure 2.4-left. As a result of this change, the outermost path, now starting from $p'_0$ but still tied to its ending point $p_1$, will first traverse the frozen region, then approach and merge in the compact cluster of paths, and eventually reach its endpoint $p_1$ along with the rest of the cluster.

The tangent method relies on two assumptions, both supposedly valid in the scaling limit.

(i) It makes the working hypothesis that the change of starting point of the outermost path, from $p_0$ to $p_0'$, has so *little influence on the other paths* that the arctic curve itself is not affected. If this were not the case, the method would predict a different arctic curve, as compared to other methods.

(ii) It postulates that the outermost path is almost surely a straight line traversing the frozen region from $p_0'$ until it meets the arctic curve *tangentially*. After the meeting point, the path coincides with the arctic curve for some time. What happens next depends on the model under consideration, see Figure 2.4-right.

The efficiency of the method relies on the *tangency assumption* (ii). By varying the starting point $p_0'$, one obtains a one-parameter family of straight lines tangent to the arctic curve, which is then simply recovered as its envelope. We assume that each portion of arctic curve is convex, which is to be expected heuristically. If it were not the case, the tangent method would yield the convex hull of the arctic curve, which would have a straight portion. If the tangent method does not yield any straight portion we can then retrospectively justify this assumption.

Clearly, if the arctic curve is known, it is trivial to compute the family of tangential lines, each of which arrives at a given point $p_0'$ of the boundary. The tangent method aims at solving the inverse problem: given a point $p_0'$, find the slope of the line starting from $p_0'$ and hitting the arctic curve tangentially. This is usually done—as in [20–23, 25–29]—by resorting to a lattice point $q_0$ *outside the domain* and by considering the path from $q_0$ to $p_1$ which (by construction) enters into the original domain through a (random) boundary point. By adding trivial deterministically frozen extensions to the domain (see e.g. Figure 5.10) one sees that (i) and (ii) should also hold for this external starting point $q_0$. In the scaling limit and given $q_0$, the paths are straight lines with probability 1, entering the domain at some boundary point and hit the arctic curve tangentially. Since the probability distribution over the point of entry concentrates on a unique point in the scaling limit, call it $p_0'$, it can be computed on the lattice as the entry point of maximal probability. In this way, the point $p_0'$ can be obtained as a function of $q_0$ and allows to compute the slope of the line that will eventually meet the arctic curve tangentially. Varying $q_0$ yields the required family of tangent lines.

As pointed out in [20], the hypothesis (i) is hard to formalize in general, let alone to prove. The assumption (ii) is easier to formulate in general. In Chapter 5, we study the consequences of hypothesis (i) on the partition function, which enable us to use the tangent method *without extending the domain*. We also provide strong arguments supporting the tangency property (ii).

Notice that the general strategy exposed above must often be slightly adapted to the model under consideration. For instance, in the model of $q$-weighted lattice paths studied in [25,27], the straight lines are replaced by other known curves (see e.g. Figure 5.4). In this thesis, several variations of the tangent method are presented. In the case at hand, namely tilings of the rectangoloid domain, the application of the tangent method also presents slight adaptations, as we shall now explain.

### 2.1.3 Tangent method derivation for freezing boundaries

A natural way to tackle the determination of the arctic curve from a lattice computation would be to evaluate bulk expectation values to detect the frozen regions. However, it is usually a much simpler task to use the tangent method. The reason is that computing the most likely entry point only requires the knowledge of boundary one-point functions which are usually easier to evaluate than their bulk counterparts. For example, boundary one-point functions of the six-vertex model with domain wall boundary conditions for general weights are known [88] while the computation of its bulk expectation values could be completed for particular weights at the so-called free-fermion point [56] but is still ongoing for general weights [89].

Nevertheless, computing a boundary one-point function and its asymptotics requires some effort. In the model at hand however, this endeavour is greatly simplified by the existence of a simple formula for the partition function $Z(\{a_i\})$ with arbitrary $a_i$'s called the Gelfand-Tsetlin formula [8]. The trick to fully exploit this formula is to extend the domain "from below".

**Freezing boundary of first kind: macroscopic flat interval.** We first consider the case where the defect distribution contains a macroscopic flat interval, corresponding to a macroscopic gap in the distribution of the starting points $O_i = (a_i, 0)$ for the NILP of type 1 (see Figure 2.5). In other words, there is an integer $q < n$ such that $a_{q+1} - a_q = m$, with $m$ proportional to $n$. We assume that the density of starting points to the left of $a_q$ and to the right of $a_{q+1}$ is strictly greater than 0. Above this flat interval two disjoint statistically frozen regions appear, respectively composed of F-tiles, and U-tiles (see Figure 2.2). We apply the tangent method separately for the corresponding F-portion and U-portion of the arctic curve. The F-portion can be directly investigated using NILP of the first kind, while the U-type is more conveniently studied with the NILP of the second kind.

**F-portion.** To apply the tangent method, we extend the domain by a semi-infinite strip as indicated on Figure 2.5. We move the point $O_{q+1}$ to $O'(r) = (a_{q+1}, -r)$. We consider the ratio of the partition function of this modified problem to the one of the original problem. It can be written as a sum over the entry point $O(\ell) = (a_q + \ell, 0)$:

$$\frac{Z(\{a_i\}_{i\neq q+1}, O'(r))}{Z(\{a_i\})} = \sum_{\ell=1}^{m} Y_{r,\ell}\, \frac{Z(\{a_i\}_{i\neq q+1}, O(\ell))}{Z(\{a_i\})} = \sum_{\ell=1}^{m} Y_{r,\ell}\, H_\ell, \quad (2.1)$$

where $Y_{r,\ell}$ counts the number of configurations of the path between $O'(r)$ and $O(\ell)$. We also introduced $Z(\{a_i\}_{i\neq q+1}, O(\ell)) = Z(\{a'_i\})$ that counts the number of configurations in the original domain, but with slightly different settings: the starting points are $\{a'_i\}$ with $a'_i = a_i$ for $i \neq q+1$ and $a'_{q+1} = a_q + \ell$. Note that for $r$ fixed, $\ell$ is random. The tangent method enjoins us to compute the most likely value of $\ell/n$ as a function of $r/n$, in the scaling limit. This can either be achieved by computing the probability of observing $\ell/n$ and maximizing it, or slightly more conveniently by performing a saddle point analysis on (2.1).

We expect partition functions such as $Z(\{a_i\}_{i\neq q+1}, O(\ell))$ to grow exponentially with $n^2$, with various sub-dominant corrections. The asymptotic behaviour of $Z(\{a_i\})$ should be similar, up to sub-dominant contributions that appear in the asymptotics of $H_\ell$. We will explicitly evaluate $\log(H_\ell)$ and find that it is at most of order $n$.

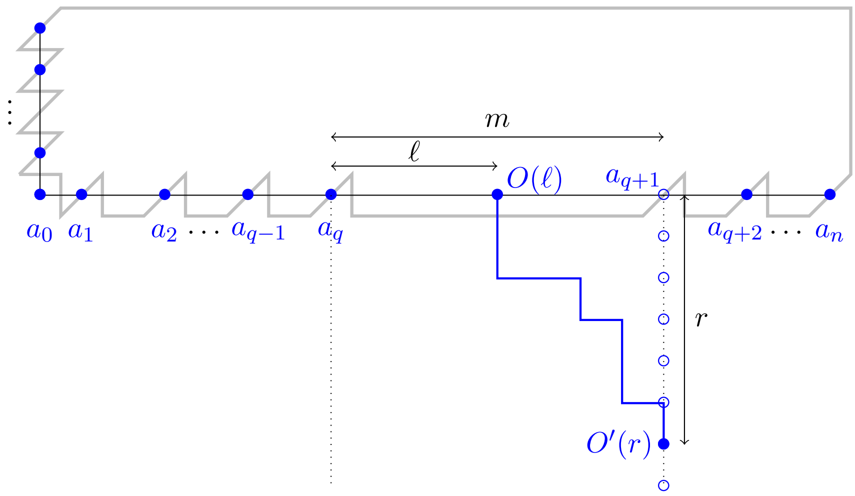



Figure 2.5: Extension of the domain by a semi-infinite strip to apply the tangent method using NILP of first type. The thick grey line represents the boundary of the original domain (without the extension). The open dots ($\circ$) are the possible new starting positions of $O'(r)$ with $r = 0, 1, \cdots$.

To proceed, we first compute the quantities appearing on the right-hand side of (2.1). In the strip, the path has $m - \ell$ west-oriented elementary steps and $r$ north-oriented ones, where the last step leading to $O(\ell)$ has to be north-oriented. We therefore have

$$Y_{r,\ell} = \binom{m - \ell + r - 1}{m - \ell}. \tag{2.2}$$

The boundary one-point function $H_\ell$ can be easily determined by using the Gelfand-Tsetlin formula [8]

$$Z(\{a_i\}) = \prod_{i=1}^{n} \prod_{s=0}^{i-1} \frac{a_i - a_s}{i - s}. \tag{2.3}$$

This formula can be proven in several ways. The most systematic way is—arguably—to use the LGV lemma and an LU decomposition, as done in [23]. Another proof uses the bijection of the tilings with semi-strict Gelfand patterns and proceeds by induction on $n$ as in [8]. In both cases, one has to *guess the result first* and then prove that it is correct, see [75, 76] for an extensive report of various useful techniques. The guesswork involves evaluating LGV determinants for several sizes and several choices of parameters. This step is a real bottleneck in the study of tiling models since the LGV determinant itself is often not suited to asymptotic analysis and it is *in general* very hard to guess a closed form formula, i.e. a product or quotient of "nice" factors [75]. A pleasant property of rectangoloid tilings is that when one fixes $n$ and the $a_i$'s, the prime factor decomposition of the LGV determinant is only composed of *small prime numbers*. This property—which is usually destroyed by even the slightest modification of the domain—suggests the existence of a "product" formula such as (2.3) and immensely simplifies the guesswork[4].

There are several interesting particular cases of (2.3), let us mention two of them. When $a_i = i$ for $i = 1, \cdots, q$ and $a_i = m+i$ for $i = q+1, \cdots, n$ with $q$ and $m$ positive integers, the tiling problem is equivalent—upon removal of deterministically frozen regions and a shear—to lozenge tiling of a hexagon. In this case (2.3) reduces to the celebrated MacMahon formula [91]. When $a_i = 2i$ for $i = 1, \cdots, n$ the corresponding tiling problem is called the Novak half-hexagon [92] and (2.3) gives $Z(\{a_i\}) = 2^{n(n+1)/2}$. These tilings are hence equinumerous with domino tilings of the Aztec diamond, yet no natural bijection between the two is known [22, 92], although interesting similarities exist [92].

Computing $H_\ell$ is usually the most involved step in a tangent method derivation, but the Gelfand-Tsetlin formula makes it straightforward in the present case,

$$\begin{aligned} H_\ell = \frac{Z(\{a'_i\})}{Z(\{a_i\})} &= \prod_{i=1}^{n}\prod_{s=0}^{i-1}\left(\frac{a'_i - a'_s}{a_i - a_s}\right) \\ &= \exp\left\{\sum_{i=0}^{q}\log\left(1+\frac{m-\ell}{a_i - a_{q+1}}\right) + \sum_{i=q+2}^{n}\log\left(1+\frac{m-\ell}{a_i - a_{q+1}}\right)\right\}. \end{aligned} \tag{2.4}$$

We now consider the rescaled domain where all lengths are rescaled by $n$. We introduce the rescaled variables

$$\ell = n\xi, \quad q = nu_1, \quad m = n\delta, \quad i = nu, \quad r = nz, \quad a_i = n\alpha\left(\frac{i}{n}\right), \tag{2.5}$$

[4] Instances of simple sum formulae enumerating tilings also exist (e.g. [90]) but are more difficult to guess. As will be explained in Section 5.4, for tiling models whose LGV matrix have an explicit LU decomposition, it is possible to compute refined partition functions, which are then typically expressed as sums [22, 23].

where the function $\alpha(u)$ characterises the defect distribution. We limit ourselves to asymptotic defect distributions such that $\alpha : [0, 1] \to \mathbb{R}^+$ is a (fixed) piecewise differentiable increasing function. Notice also that $\alpha'(u) \geq 1$ whenever it exists due to the condition $a_{i+1} - a_i \geq 1$. Since $a_{q+1} = a_q + m$, in the $n \to \infty$ limit, the function $\alpha(u)$ has a discontinuity at $u_1$ where it jumps by an amount $\delta$. Thus, we set $\alpha(u_1) = \lim_{u \to u_1^-} \alpha(u)$ while $\lim_{u \to u_1^+} \alpha(u) = \alpha(u_1) + \delta$.

We introduce the notation $A_n \approx B_n$, meaning that $\lim_{n\to\infty} \frac{1}{n} \log(A_n) = \lim_{n\to\infty} \frac{1}{n} \log(B_n)$. The symbol $\approx$ will keep the same meaning throughout this thesis, as opposed to other notations which are usually quite short-lived. We have $Y_{r,\ell} \approx \mathrm{e}^{nS_0}$ and $H_\ell \approx \mathrm{e}^{nS_1}$ with

$$S_0 = (\delta - \xi + z) \log(\delta - \xi + z) - (\delta - \xi) \log(\delta - \xi) - z \log z \tag{2.6}$$

and

$$S_1 = \int_0^1 \mathrm{d}u \, \log \left( 1 + \frac{\delta - \xi}{\alpha(u) - (\alpha(u_1) + \delta)} \right). \tag{2.7}$$

We perform the saddle point analysis on

$$\sum_{\ell=1}^{m} Y_{r,\ell} \, H_\ell \approx \int_0^\delta \mathrm{d}\xi \; \mathrm{e}^{n[S_0 + S_1]}. \tag{2.8}$$

Let us denote by $\xi^*$ the solution to the saddle point equation $\frac{\partial}{\partial \xi}(S_0 + S_1) = 0$. We could choose to parametrize the family of tangent lines by $z$, viewing $\xi^* = \xi^*(z)$ as a function of $z$, but it turns out to be more convenient to use the intercept $t \equiv \alpha(u_1) + \xi^*$ of the tangent line with the $X-$axis, thereby making $z = z(t)$ a function of $t$. The arctic curve is then parametrized by $t$ as well. Defining

$$x(t) \equiv \exp \left\{ - \int_0^1 \frac{\mathrm{d}u}{t - \alpha(u)} \right\}, \tag{2.9}$$

the saddle point condition yields

$$\frac{\alpha(u_1) + \delta - t}{\alpha(u_1) + \delta - t + z(t)} = x(t). \tag{2.10}$$

By definition the parameter $t$ must be contained in the interval $[\alpha(u_1), \alpha(u_1)+\delta] = [\alpha(u_1^-), \alpha(u_1^+)]$, but not all values in this interval correspond to real positive values of $z$.

Let us write

$$I(t) \equiv -\log x(t) = \int_0^{u_1^-} \frac{\mathrm{d}u}{t-\alpha(u)} - \int_{u_1^+}^1 \frac{\mathrm{d}u}{\alpha(u)-t}. \tag{2.11}$$

When $t$ is in the interior of $[\alpha(u_1), \alpha(u_1)+\delta]$, it is not in the image of the function $\alpha(u)$, so that $I(t)$ is a well-defined, continuous and decreasing function. Moreover, $I(t)$ diverges to $+\infty$ when $t$ goes to the lower bound $\alpha(u_1^-)$, and to $-\infty$ when $t$ goes to the upper bound $\alpha(u_1^+)$. Therefore, on the closed interval, $I(t)$ takes all values from $-\infty$ to $+\infty$, and $x(t)$ itself is a continuous increasing function, taking all positive values from 0 to $+\infty$. However, from (2.10), $z$ being real and positive implies that $x(t)$ is between 0 and 1. Therefore, the proper range of $t$ to consider is the interval $[\alpha(u_1), t_1]$ where $t_1 < \alpha(u_1) + \delta$ is defined by $x(t_1) = 1$. We then have $z(\alpha(u_1)) = +\infty$ and $z(t_1) = 0$, respectively corresponding to a vertical tangent and a horizontal tangent. We will indeed check that $t_1$ corresponds to the transition point from the F-frozen region to the U-frozen one.

Imposing in the rescaled domain that the tangent line associated with the parameter $t$ passes through the points $(t, 0)$ and $(\alpha(u_1) + \delta, -z(t))$ leads to the equation

$$x(t)Y + (1 - x(t))(X - t) = 0, \quad t \in [\alpha(u_1), t_1], \tag{2.12}$$

which is exactly the equation (3.9) of [23].

**U-portion.** We now consider the U-portion. To use the tangent method, we consider an equivalent description, in terms of NILP of second kind (see Figure 2.6). We extend the domain in a similar fashion as for the F-case by gluing a semi-infinite tilted strip and by moving $\tilde{O}_{n-q-1}$ to $\tilde{O}'(r) = (\tilde{a}_{n-q} + 1 - r, -r)$. The computation is quite similar to the F-type case. The boundary one-point function can be computed by using NILP of first kind and is again given by (2.4), while the number of configurations of the path (of second kind) between $\tilde{O}'(r)$ and $\tilde{O}(\ell) = (\tilde{a}_{n-q} + 1 + \ell, 0)$ is

$$Y_{r,\ell} = \binom{r+\ell-1}{\ell}, \tag{2.13}$$

where we imposed that the last step is north-east-oriented. Using the same notation for the rescaled variables, we find that asymptotically $Y_{r,\ell} \approx \mathrm{e}^{nS_0}$ with

$$S_0 = (z+\xi)\log(z+\xi) - \xi\log\xi - z\log z. \tag{2.14}$$

From the saddle point analysis, we find that

$$\frac{t-\alpha(u_1)+z(t)}{t-\alpha(u_1)} = x(t), \tag{2.15}$$

for the same continuous increasing function $x(t)$ given in (2.9). From (2.15), $z(t)$ being real and positive implies that $x(t)$ takes values from 1 to $\infty$. Therefore, the proper range of $t$ to consider is $[t_1, \alpha(u_1)+\delta]$ with $x(t_1) = 1$.

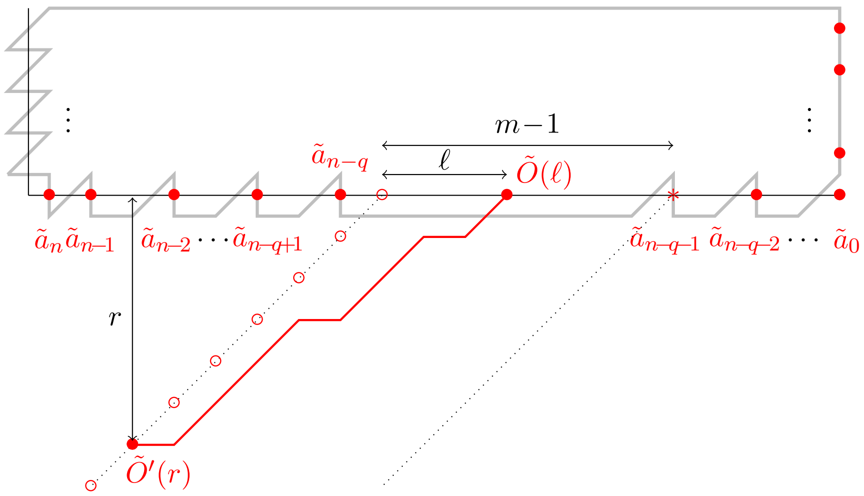



Figure 2.6: Extension of the domain with a semi-infinite strip to apply the tangent method using NILP of second kind. The starting points $\tilde{O}_i = (\tilde{a}_i, 0)$ and the ending points $\tilde{E}_i = (a_n + 1/2, i)$, are different for these paths, as represented (see [23] for more details). The star $(*)$ represents the position of the starting point that we move to $\tilde{O}'(r)$.

By imposing that the tangent line of parameter $t$ passes through $(t, 0)$ and $(\alpha(u_1) - z(t), -z(t))$, we find

$$x(t)Y + (1-x(t))(X-t) = 0, \quad t \in [t_1, \alpha(u_1)+\delta], \tag{2.16}$$

which is the same expression as in (2.12). The proof is hence complete for freezing boundaries of first kind, since the envelope of this family of lines has the same parametric representation as in [23], but with $t \in [\alpha(u_1), \alpha(u_1) + \delta]$.

**Freezing boundary of second kind: macroscopic sawtooth pattern.** We now turn to cases where a macroscopic portion of the lower boundary has a sawtooth pattern. In terms of NILP of first kind, it corresponds to an interval of indices $I = \{q, q+1, \cdots, q+m-1\}$ for which the starting points $O_i = (a_i, 0)$ are tightly arranged: $a_{i+1} - a_i = 1$ for $i \in I$. We assume that the density to the left of $a_q$ and to the right of $a_{q+m}$ is strictly less than 1 over a macroscopic distance. We further suppose that there is no starting point directly adjacent to this region, namely, $a_q - a_{q-1} > 1$ and $a_{q+m} - a_{q+m-1} > 1$. If the freezing interval is not directly adjacent to the left or right corner of the domain, this condition is simply a convenient way to define the interval. Again, by macroscopic portion, we mean that $m$ scales like $n$. In such a situation, there is a *deterministically frozen* triangular region and also a *statistically frozen* region around it. The whole frozen region is covered by a regular arrangement of R-tiles. We study the associated arctic curve using the tangent method with the third kind of NILP for which the region of interest is empty.

We extend the domain with a semi-infinite tilted strip (see Figure 2.7). We move $\hat{O}_p$ to $\hat{O}'(r) = (b_p - r, -r - 1/2)$. The same extension can be done if the freezing interval touches the left edge of the bottom boundary ($q = 0$). If it touches the right edge ($q + m = n$), we take for $\hat{O}_p$ the starting point of an ad-hoc trivial vertical path of third type added on the right of the domain. As before, we consider the quantity

$$\frac{Z(\{b_i\}_{i\neq p}, \hat{O}'(r))}{Z(\{b_i\})} = \sum_{\ell=0}^{m} Y_{r,\ell} \, \frac{Z(\{b_i\}_{i\neq p}, \hat{O}(\ell))}{Z(\{b_i\})} = \sum_{\ell=0}^{m} Y_{r,\ell} \, H_\ell, \tag{2.17}$$

where $Y_{r,\ell}$ counts the number of configurations of the path of third type between $\hat{O}'(r)$ and $\hat{O}(\ell) = (b_{p-1} + 1 + \ell, -1/2)$ and $Z(\{b_i\}_{i\neq p}, \hat{O}(\ell))$ counts the number of configurations in the original domain, but with a slightly modified bottom boundary. Since the path cannot go to the left, we have $Y_{r,\ell} = 0$ when $\ell < m - r$. When $\ell \geqslant m - r$ we find that

$$Y_{r,\ell} = \binom{r}{r + \ell - m}. \tag{2.18}$$

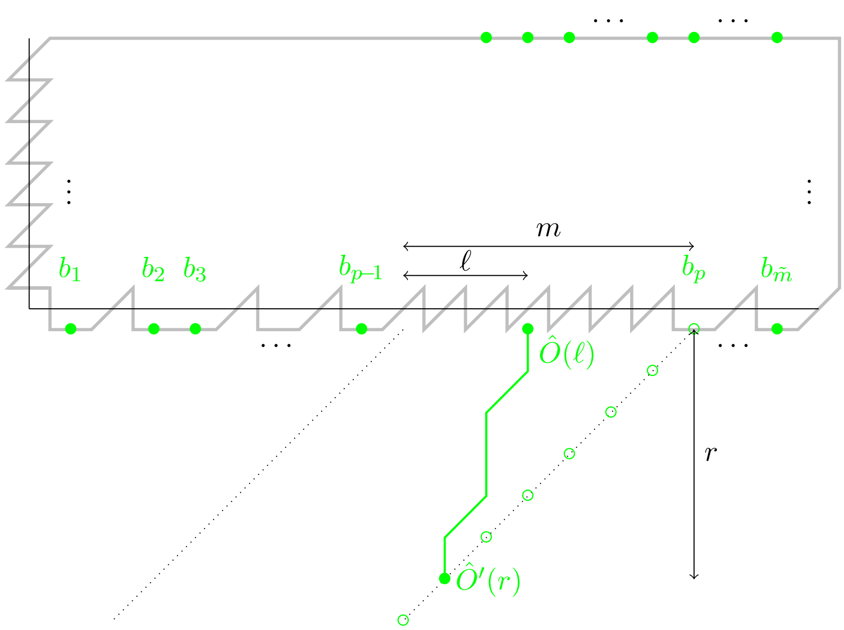



Figure 2.7: Extension of the domain to apply the tangent method, using NILP of third kind. The original starting points are $\hat{O}_i = (b_i, -1/2)$ for $i = 1, 2, \cdots, \tilde{m}$ with $\tilde{m} = a_n - n$ and the ending points are at $\hat{E}_i = (n+i, n+1/2)$. The parameter $p$ gives the position of the freezing interval in this NILP description, but we do not need to know its precise relation to $q$ (see again [23] for more details).

Instead of working with the third kind of NILP (an explicit formula is known [23] for $Z(\{b_i\})$), we express $H_\ell$ in terms of NILP of first kind,

$$H_\ell = \frac{Z(\{a_i'\})}{Z(\{a_i\})}, \tag{2.19}$$

with the following new starting points,

$$\begin{aligned} a_i' &= a_i, \qquad \text{for} \quad i \notin \{q+\ell+1, q+\ell+2, \cdots, q+m\}, \\ a_i' &= a_i + 1, \quad \text{for} \quad i \in \{q+\ell+1, q+\ell+2, \cdots, q+m\}. \end{aligned} \tag{2.20}$$

Put more simply, it means that all the points from $a_{q+\ell+1}$ to $a_{q+m}$ are shifted by one unit to the right. Using again the Gelfand-Tsetlin formula, we find that

$$\begin{aligned} H_\ell &= \prod_{i=1}^{n} \prod_{s=0}^{i-1} \left( \frac{a'_i - a'_s}{a_i - a_s} \right) \\ &= \prod_{i=q+\ell+1}^{q+m} \prod_{s=0}^{q+\ell} \left( \frac{a_i - a_s + 1}{a_i - a_s} \right) \prod_{i=q+m+1}^{n} \prod_{s=q+\ell+1}^{q+m} \left( \frac{a_i - a_s - 1}{a_i - a_s} \right) \\ &= \exp\left[ \sum_{i=q+\ell+1}^{q+m} \sum_{s=0}^{q+\ell} \log\left( \frac{a_i - a_s + 1}{a_i - a_s} \right) + \sum_{i=q+m+1}^{n} \sum_{s=q+\ell+1}^{q+m} \log\left( \frac{a_i - a_s - 1}{a_i - a_s} \right) \right], \end{aligned} \tag{2.21}$$

where we expressed this quantity in terms of the original $(a_i)_i$. We used the fact that the numerator $a'_i - a'_s$ can only differ from the denominator $a_i - a_s$ by an amount $+1$ , $-1$ or 0.

We introduce the rescaled variables as follows,

$$\begin{gathered} \ell = n\xi, \quad r = nz, \quad q = nu_1, \quad m = nu_2, \quad i = nu, \\ s = nv, \quad a_i = n\alpha\left(\frac{i}{n}\right), \end{gathered} \tag{2.22}$$

with the condition

$$\xi \geqslant u_2 - z. \tag{2.23}$$

In the limit $n \to \infty$ the function $\alpha(u)$ will be such that $\alpha'(u) = 1$ for $u \in [u_1, u_1 + u_2]$, so that $\alpha(u_1 + \xi) = \alpha(u_1) + \xi$ for any $\xi$ in $[0, u_2]$. We also note that the condition $a_q - a_{q-1} > 1$ implies that the left derivative of $\alpha$ at $u_1$ is strictly greater than 1, and likewise for the right derivative at $u_1 + u_2$,

$$\begin{aligned} &\lim_{u \to u_1^-} \alpha'(u) > 1, && \lim_{u \to u_1^+} \alpha'(u) = 1, \\ &\lim_{u \to (u_1+u_2)^-} \alpha'(u) = 1, && \lim_{u \to (u_1+u_2)^+} \alpha'(u) > 1. \end{aligned} \tag{2.24}$$

In the large $n$ limit, we have $Y_{r,\ell} \approx \mathrm{e}^{nS_0}$ and $H_\ell \approx \mathrm{e}^{nS_1}$ with

$$\begin{aligned} S_0 &= z\log z - (z+\xi-u_2)\log(z+\xi-u_2) - (u_2-\xi)\log(u_2-\xi), \\ S_1 &= \int_{u_1+\xi+\epsilon}^{u_1+u_2} \mathrm{d}u \int_0^{u_1+\xi} \mathrm{d}v \frac{1}{\alpha(u)-\alpha(v)} - \int_{u_1+u_2}^{1} \mathrm{d}u \int_{u_1+\xi}^{u_1+u_2-\epsilon} \mathrm{d}v \frac{1}{\alpha(u)-\alpha(v)}, \end{aligned} \tag{2.25}$$

where $\epsilon = 1/n$. The $\epsilon$ regularization can be read from the values taken by $i$ and $s$ in (2.21). A careful analysis of the asymptotics of (2.21) requires to treat separately the terms in which $i$ is close to $s$. This can be done by using the property $a_s = a_q + (s-q)$ for $s \in I$ and by assuming that the density of $a_i$'s is constant (and strictly less than 1) on a macroscopic portion to the right of $a_{q+m}$. This computation leads exactly to (2.25). We apply the steepest descent method on

$$\sum_{\ell=m-r}^{m} Y_{r,\ell}\, H_\ell \approx \int_{u_2-z}^{u_2} \mathrm{d}\xi\, \mathrm{e}^{n(S_0+S_1)}. \tag{2.26}$$

The derivatives of the actions $S_0$ and $S_1$ with respect to $\xi$ read

$$\begin{aligned} \frac{\partial S_0}{\partial \xi} &= \log\left(\frac{u_2-\xi}{z+\xi-u_2}\right), \\ \frac{\partial S_1}{\partial \xi} &= \int_0^{u_1+\xi-\epsilon} \frac{\mathrm{d}u}{\alpha(u)-\alpha(u_1+\xi)} + \int_{u_1+\xi+\epsilon}^{1} \frac{\mathrm{d}u}{\alpha(u)-\alpha(u_1+\xi)} \\ &= \text{p.v.} \int_0^1 \frac{\mathrm{d}u}{\alpha(u)-\alpha(u_1+\xi)}. \end{aligned} \tag{2.27}$$

The stationarity condition gives $\xi^*$, corresponding to the entry point of the tangent:

$$\frac{u_2-\xi^*}{z+\xi^*-u_2} \exp\left\{-\text{p.v.} \int_0^1 \frac{\mathrm{d}u}{\alpha(u_1+\xi^*)-\alpha(u)}\right\} = 1. \tag{2.28}$$

Let us again introduce $t = \alpha(u_1+\xi^*) = \alpha(u_1)+\xi^*$ the intercept of the tangent line with the $X-$axis. We also define

$$x(t) \equiv -\exp\left\{-\text{p.v.} \int_0^1 \frac{\mathrm{d}u}{t-\alpha(u)}\right\}, \tag{2.29}$$

where we introduce the minus sign in the definition of $x(t)$ for this range of $t$ in order to match the parametric form of the arctic curve given in [23]; it therefore implies $x(t) \leqslant 0$. The above condition then becomes

$$z(t) = \big(\alpha(u_1+u_2)-t\big)\big(1-x(t)\big). \tag{2.30}$$

The negativity of $x(t)$ ensures that the condition (2.23) is satisfied.

By definition $t$ must be in $[\alpha(u_1), \alpha(u_1+u_2)]$, but the whole interval may not be covered as we move $z$ from 0 to $+\infty$. Let us split the integral appearing in $x(t)$ as follows

$$I(t) = \int_0^{u_1} \frac{\mathrm{d}u}{t-\alpha(u)} + \int_{u_1}^{u_1+u_2} \frac{\mathrm{d}u}{t-\alpha(u)} + \int_{u_1+u_2}^{1} \frac{\mathrm{d}u}{t-\alpha(u)}, \tag{2.31}$$

with the prescription that the singularity of the integrand at $u = \alpha^{-1}(t)$ must be approached symmetrically from the left and from the right.

When $t$ is in the open interval $]\alpha(u_1), \alpha(u_1+u_2)[$, the singularity is in the second integral, for which the prescription on how to approach the singularity returns the principal value of the integral. Using the linearity of $\alpha(u)$ in the interval $[u_1, u_1+u_2]$, we find

$$\begin{aligned} \text{p.v.} \int_{u_1}^{u_1+u_2} \frac{\mathrm{d}u}{t-\alpha(u)} &= \text{p.v.} \int_{u_1}^{u_1+u_2} \frac{\mathrm{d}u}{t-(\alpha(u_1)+(u-u_1))} \\ &= \log\left(\frac{t-\alpha(u_1)}{\alpha(u_1+u_2)-t}\right), \end{aligned} \tag{2.32}$$

which takes all real values when $t$ varies over the open interval. The other two integrals are finite for $t$ in the open interval, but diverge as $t$ approaches the lower and upper bounds of the interval.

When $t$ gets close to the lower bound $\alpha(u_1)$, the integrands of the first two integrals in (2.31) are singular, respectively at $u_1^-$ and $u_1^+$. Expanding around these two points yields

$$\frac{1}{\alpha(u_1)-\alpha(u)} \sim \frac{1}{\alpha'(u_1^\pm)} \frac{1}{u-u_1}, \qquad u \sim u_1^\pm. \tag{2.33}$$

Applying the prescription to stay at a distance $\epsilon$ from the singularity, we see that the first two integrals both contain a logarithmically diverging term, equal to $-\log \epsilon/\alpha'(u_1^-)$ for the first integral, $\log \epsilon/\alpha'(u_1^+)$ for the second one. As $\alpha'(u_1^-) > 1$ and $\alpha'(u_1^+) = 1$, the two divergences do not cancel out. As $\epsilon$ goes to 0, their sum diverges to $\lim_{\epsilon\to 0}(1 - 1/\alpha'(u_1^-)) \log \epsilon = -\infty$. As a consequence, $I(\alpha(u_1))$ and $x(\alpha(u_1))$ both diverge to $-\infty$. A similar analysis for the upper bound $t \to \alpha(u_1+u_2)$ of the interval shows that $I(t)$ diverges to $+\infty$, and $x(t)$ tends to 0. By continuity the function $I(t)$ takes all values from $-\infty$ to $+\infty$ when $t$ is varied over the closed interval $[\alpha(u_1), \alpha(u_1+u_2)]$, and correspondingly $x(t)$ takes all values from $-\infty$ to 0.

It follows that $z(t)$ takes all values from $+\infty$ to 0 when $t$ runs over $[\alpha(u_1), \alpha(u_1+u_2)]$, equivalently the intercept $\xi^*$ covers the full interval $[0, u_2]$. In the rescaled domain, the equation of the tangent line of parameter $t$ passing through the points $(0, t)$ and $(\alpha(u_1)+u_2-z(t), -z(t))$ is then

$$x(t)Y + \big(1-x(t)\big)(X-t) = 0, \quad t \in [\alpha(u_1), \alpha(u_1+u_2)]. \tag{2.34}$$

From the limits given above, the tangent has a slope 1 for $t = \alpha(u_1)$, and is vertical for $t = \alpha(u_1 + u_2)$ (see Figure 2.3).

**Final results.** In all cases, we conclude that

$$x(t)Y + (1 - x(t))(X - t) = 0, \qquad t \in I_{\text{freez}}, \tag{2.35}$$

for $I_{\text{freez}}$ an interval of the real line where there is a freezing boundary of one of the two kinds discussed above. The envelope of this family of lines is constituted of all the portions of arctic curve, except for the ones touching the upper boundary. These remaining portions of the arctic curve that we have not analysed so far were derived in [23] by extending the domain above the lower edge. Alternatively, these portions, that will correspond to a range $t \in [\alpha(1), \infty[$ and $t \in ]-\infty, 0]$, can also be accessed by extending the domain from below, in the following way. For $t \in [\alpha(1), \infty[$ we consider NILP of first kind and attach on the right of the original domain a semi-infinite horizontal strip whose upper horizontal edge is aligned with that of the original domain and of height $(1 + a)n$ with $a > 0$ arbitrary. We then move $O_n$ to the lower horizontal edge of this extension. For $t \in ]-\infty, 0]$, we consider NILP of second kind and use a similar semi-infinite horizontal strip, this time attached on the left, also making $\tilde{O}_n$ start from the lower edge of the strip. In both cases, we recover $x(t)Y + (1 - x(t))(X - t) = 0$ by using the same kind of arguments as above. The complete family of tangent lines for the tiling problem of Figure 2.2 is represented on Figure 2.8.

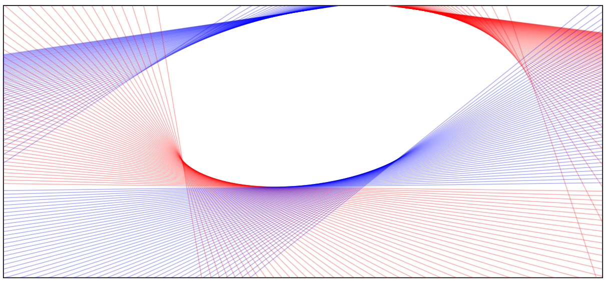


Figure 2.8: Family of tangent lines for the defect distribution of Figure 2.2. Blue lines correspond to frozen regions of U-type and red lines to F-type.

Retrieving the envelope of a family of tangent lines precisely corresponds to an inverse Legendre transform that can indeed be used to find the equation $X(Y)$ for different portions of the arctic curve [23]. More generally, given a family of curves parametrized by $t$ described by $F(X, Y, t) = 0$ with $F$ sufficiently smooth, its envelope is constituted of those points $(X, Y)$ satisfying $F(X, Y, t) = 0$ and $\partial_t F(X, Y, t) = 0$, see e.g. [93]. In the case at hand it leads to

$$\begin{cases} x(t)Y + (1 - x(t))(X - t) = 0, \\ x'(t)Y - x'(t)(X - t) - (1 - x(t)) = 0. \end{cases} \tag{2.36}$$

Solving the system above leads to the equation for the arctic curve:

**Theorem 2.1.1.** *The arctic curve of the rectangoloid whose asymptotic defect distribution is described by $\alpha(u)$, as predicted by the tangent method, is given by the following parametric equation*

$$\begin{aligned} X(t) &= t - \frac{x(t)\,[1 - x(t)]}{x'(t)}, \\ Y(t) &= \frac{[1 - x(t)]^2}{x'(t)}, \end{aligned} \tag{2.37}$$

*with $t \in ]-\infty, 0] \cup I_{freez} \cup [\alpha(1), \infty[$. When $t$ belongs to a freezing interval of second kind we have $x(t) \equiv -\exp\left\{-p.v. \int_0^1 \frac{\mathrm{d}u}{t-\alpha(u)}\right\}$ and in all other cases we have $x(t) \equiv \exp\left\{-\int_0^1 \frac{\mathrm{d}u}{t-\alpha(u)}\right\}$.*

**Discussion.** A few comments are in order. Let us first briefly describe the solution (2.37)—an in-depth analysis can be found in [23] and [78]. In the two particular cases mentioned above, namely the lozenge tiling of the hexagon and of the Novak half-hexagon, the arctic curves are respectively found to be an ellipse and a parabola, in agreement with known results [8, 92]. When freezing intervals are present, the arctic curve is a bit more intriguing and presents cusps, see Figures 2.2 and 2.3. Perhaps more strikingly, the arctic curve is analytic, and we notice retrospectively that all portions can be obtained from a single one by *analytical continuation*. This appears to be a property of free-fermionic models, but there is no completely general proof supporting this observation to the best of our knowledge. What is certain however is that this property does not hold for non-free-fermionic models for which a lack of analyticity may appear at the contact point with the boundary of the domain [17, 94]. It has actually been proven [95] that for a large class of boundary conditions (polygonal boundary, simply or multiply connected) the arctic curve of lozenge tiling models is *algebraic* (i.e. can be written as $P(X, Y) = 0$ with $P$ a polynomial). It should be noted however that for some choices of $\alpha(u)$ the rectangoloid domain does not fall in this class and leads to transcendental curves, see [23] or [78].

## 2.2 Quarter turn symmetric holey Aztec domino tiling

In this section, we apply the tangent method—as originally done by Di Francesco and Guitter in [17]—to yet another *tiling problem*, first studied in [59]. We find it interesting in several regards. Firstly, it has a combination of *fixed and periodic* boundary conditions, a situation for which not much is known regarding the arctic phenomenon, to the best of our knowledge. Secondly, a new subtlety in the application of the tangent method arises. Thirdly, its enumeration *does not factorize* into small prime integers. Quite interestingly however, despite the lack of closed form formula for the partition function, an asymptotic analysis can be performed. This is achieved by using another remarkable property of this model, namely the mysterious relation [17, 59] it has with the twenty-vertex model with domain wall boundary conditions of type 1. In [17], we treated these models conjointly. In this thesis however, since it is otherwise independent, the twenty-vertex model is discussed separately in Chapter 4. In this section, the aforementioned relation is expressed a bit indirectly in terms of the six-vertex model with domain wall boundary conditions, which is itself naturally related to the twenty-vertex model as explained in Chapter 4.

The tiling model in question is precisely defined in [59] and is called the quarter turn symmetric holey Aztec domino tiling (QTHADT) model. Loosely speaking, QTHADT are domino tilings of a $2n \times 2n$ domain of Aztec-like shape with a cross-shaped hole at its core that are *invariant under a* $90°$ *rotation* around the center of the cross as shown on Figure 2.9.

By symmetry, any configuration of the QTHADT model can be entirely described by its restriction to some well-chosen subdomain. A natural

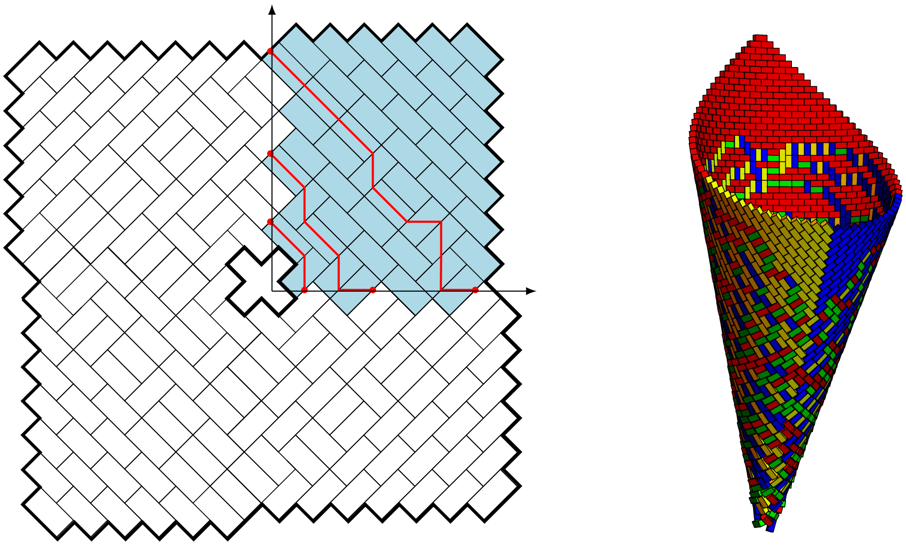


Figure 2.9: *Left*: Sample configuration QTHADT along with the associated NILP configuration in the fundamental domain (in blue). *Right*: Equivalent description of the model as a tiling of a cone with a hole at its apex.

choice is the upper right quadrant, which we call the fundamental domain. The quarter turn symmetry imposes periodicity along two edges of the fundamental domain, so that configurations of our model can be seen as a tiling of a cone as seen in Figure 2.9-right. This figure also shows that configurations in the fundamental domain may be described in terms of NILP whose definition is a bit different than those previously met. The main differences are that the number of paths $p$ can vary ($p = 0, 1, \cdots, n$, see Figure 3.1 for an example of $p = 0$ and $p = 1$) and that the starting and ending points of a given path must be placed symmetrically on the horizontal and vertical axis—modulo a unit shift—due to the periodic boundary conditions. In the reference system of Figure 2.9-left the starting points are at $(0, 1 + i_k)$ while the ending points are at $(i_k, 0)$ for $k = 1, 2, \cdots, p$ and $i_k = 1, 2, \cdots, n-1$. The elementary steps are $\{(1, 0), (1, -1), (0, -1)\}$ and the first step can only be $(1, 0)$ or $(1, -1)$. The appearance of an arctic curve may be understood with a similar heuristics as for the Aztec diamond case, the only difference is that the extremities (i.e. starting and ending points) are not fixed. Entropic considerations indicate that these extremities should be distributed sparsely near the cross and *densely near the boundary of the domain*.

We slightly generalize the model by introducing an extra weight $\gamma > 0$ per diagonal step, which amounts in the tiling language to assign a weight $\gamma$ to a particular type of tile (in green in Figure 2.9-right). In the limit $\gamma \to 0$ a QTHADT reduces to a lozenge tiling discussed in [96] that are in bijection with descending plane partitions (DPP) (see Figure 2.12).

### 2.2.1 Boundary one-point function

**Partition function.** It is often useful—and the case at hand is no exception—to compute the partition function before trying to evaluate boundary one-point functions. For completeness we reproduce (and slightly generalize) the computation done in [59]. A LGV determinant enumerates NILP with *fixed* starting and ending points. We start by fixing the $i_k$'s and later on we sum over all the possible starting and ending point positions. Let $M_{i,j}$ be the number of paths (with the required restrictions) from $(0, j+1)$ to $(i, 0)$. The offset in the indices is introduced for convenience and is related to the unit shift mentioned above. The number of NILP with fixed $i_k$'s is the determinant of the submatrix obtained by keeping the rows and columns indexed by $i_1, \cdots, i_p$ noted $|M|_{i_1,\cdots,i_p}^{i_1,\cdots,i_p}$. The partition function of the QTHADT model is hence

$$A_n(\gamma) = \sum_{p=0}^{n} \sum_{0\leq i_1<i_2<\cdots<i_p\leq n-1} |M|_{i_1,\cdots,i_p}^{i_1,\cdots,i_p} = \det_{0\leq i,j\leq n-1} (\delta_{i,j} + M_{i,j}), \quad (2.38)$$

where we used the von Koch formula[5], see e.g. [97]. The matrix elements $A_{i,j} \equiv \delta_{i,j} + M_{i,j}$ can be obtained from the generating function $F_A(u, v)$ given by

$$F_A(u, v) = \sum_{i,j\geqslant 0} A_{i,j} u^i v^j = \frac{1}{1-uv} + \frac{(1+\gamma)u}{(1-u)(1-u-v-\gamma uv)} \ . \quad (2.39)$$

This equation is deduced from the generating function $F_M(u, v) \equiv \sum_{i,j\geq 0} M_{i,j} u^i v^j$. One can evaluate $F_M(u, v)$ entirely in the language of

[5]This identity may be proven by Taylor expansion of the polynomial $\det_{0\leq i,j\leq n-1}(\delta_{i,j} + zM_{i,j})$.

generating function (using recurrence relations for $M_{i,j}$ is also tractable). There are two possibilities for the first few steps after the starting point. Either a path has $k \geq 1$ horizontal steps (generated by $\frac{u}{1-u}$) followed by a vertical step (generated by $v$), or it has $k \geq 0$ horizontal steps (generated by $\frac{1}{1-u}$) followed by a diagonal step (generated by $\gamma uv$). Notice that we took into account our restriction that there can be no $(0,-1)$ step directly after the starting point. After these few initial steps we have in both cases an unconstrained path generated by $\frac{1}{1-u-v-\gamma uv}$. Combining all these observations we conclude that

$$F_M(u,v) = \frac{1}{v}\left(\frac{u}{1-u}v + \frac{1}{1-u}\gamma uv\right)\frac{1}{1-u-v-\gamma uv}, \tag{2.40}$$

where the prefactor $1/v$ accounts for the unit shift in the definition of $M_{i,j}$. This concludes the proof of (2.39). Using this result one gets [59]

$$A_n(\gamma = 1) = 1, 3, 23, 433, 19705, 2151843, \cdots . \tag{2.41}$$

As already mentioned, the prime factors that appear in the decomposition of these numbers are not small (e.g. $2151843 = 3 \cdot 17 \cdot 42193$). Hence, there is not much hope of further simplifying $\det(A_{i,j})$, which is often a first step to find a simple formula for boundary one-point function [22, 23, 25, 27]. In our particular case, such a simple formula is not necessary for the asymptotic analysis.

**Refined partition function.** The path enumeration may be *refined* as follows: we introduce an extra multiplicative weight $\tau$ per horizontal step along the row of maximal height. This notion of refinement is closely related to the one we already discussed, the connection being the extraction of coefficients of the refined enumeration, which is a polynomial in $\tau$. For paths that start at position $(0,n)$ the refinement changes the weight as follows: paths are obtained either by combining at least one first horizontal step (an arbitrary $k \geq 1$ number of them, generated by $\frac{\tau u}{1-\tau u}$) followed by a vertical step (generated by $v$), or by combining an arbitrary number $k \geq 0$ of horizontal steps (generated by $\frac{1}{1-\tau u}$) followed by a diagonal step (generated by $\gamma uv$), both followed by a path starting at height $n-1$ (generated by $\frac{1}{1-u-v-\gamma uv}$). The net result is a change $A \to A(\tau) = \mathbb{I} + M(\tau)$ where $M(\tau)$ differs from $M$ in its $n$-th column only, now generated by

$$\sum_{i=0}^{\infty} M_{i,n-1}(\tau)u^i = \frac{1}{v}\left(\frac{\tau u}{1-\tau u}v + \frac{1}{1-\tau u}\gamma uv\right)\frac{1}{1-u-v-\gamma uv}\bigg|_{v^{n-1}}$$
$$= \frac{(\tau+\gamma)u}{(1-\tau u)(1-u-v-\gamma uv)}\bigg|_{v^{n-1}}. \quad (2.42)$$

The *refined partition function* is given by $\det_{0\le i,j\le n-1}(A_{i,j}(\tau))$. We conclude from the previous discussion that the complete generating function for $A(\tau)$ is

$$F_{A(\tau)}(u,v)$$
$$= \frac{1}{1-uv} + \frac{(1+\gamma)u}{(1-u)(1-u-v-\gamma uv)} + \left\{\frac{\tau+\gamma}{1-\tau u} - \frac{1+\gamma}{1-u}\right\}\frac{(1+\gamma u)^{n-1}}{(1-u)^n}uv^{n-1}$$
$$= \frac{1}{1-uv} + \frac{(1+\gamma)u}{(1-u)(1-u-v-\gamma uv)} + \frac{(\tau-1)u}{(1-u)(1-\tau u)}\left(\frac{1+\gamma u}{1-u}\right)^n v^{n-1}. \quad (2.43)$$

The tangent method requires to compute the asymptotics of $\frac{\det(A(\tau))}{\det(A)}$. Deducing it from (2.43) is quite a challenging task, there is no general recipe for a given generating function. Quite conveniently, this quantity is actually linked to the boundary one-point function of the six-vertex model with domain wall boundary conditions, whose asymptotics is known [14, 15, 98].

**Connection with the six-vertex model.** The precise definitions of the six-vertex model and of the associated weights are not central for the argument and will be given in Chapter 4. The refined partition function for the six-vertex model with domain wall boundary conditions, with vertex weights $a$, $b$, $c$ and an extra weight $\sigma$ per horizontal step in the top row (in the osculating path formulation) is $\det(B(\sigma))$, where the matrix $B(\sigma) = (B_{i,j}(\sigma))_{0\le i,j\le n-1}$ is generated by

$$F_{B(\sigma)}(u,v) = \frac{1}{1-uv} + \frac{xu}{(1-u)(1-yu-v-(x-y)uv)}$$
$$+ \frac{(\sigma-1)xu}{(1-u)(1-(y+x(\sigma-1))u)}\left(\frac{1+(x-y)u}{1-yu}\right)^n v^{n-1}, \quad (2.44)$$

with

$$x = \left(\frac{b}{a}\right)^2, \qquad y = \left(\frac{c}{a}\right)^2. \tag{2.45}$$

The proof of (2.44) is given in [99] and a similar computation is also done in [59]. Comparing with the expression found in the previous section, we are led to identify $y = 1$ and $x = 1+\gamma$. In the usual parametrization $a = \sin(\lambda+\eta)$, $b = \sin(\lambda-\eta)$ and $c = \sin(2\eta)$, this corresponds to taking $\lambda = \pi - 3\eta$, so that $a = c$ and $b/a = 2\cos(2\eta)$ and further imposing

$$\gamma = 1 + 2\cos(4\eta). \tag{2.46}$$

The generating functions $F_{A(\tau)}(u,v)$ and $F_{B(\sigma)}(u,v)$ are therefore identified upon taking

$$\tau = y + x(\sigma - 1) = 1 + (1+\gamma)(\sigma - 1) = 1 + 4\cos^2(2\eta)(\sigma - 1). \tag{2.47}$$

Notice that if we specialize this result to $\sigma = 1$ (implying $\tau = 1$), we conclude that the (non-refined) partition function of these two models also coincide.

### 2.2.2 Tangent method for the QTHADT

There is a subtlety in the application of the tangent method, arising from the non-fixed nature of the paths extremities. Configurations without a starting point at $(0,n)$ exist. For example, using the tiling on Figure 2.9-left, a single flip of dominoes—only one of which is in the fundamental domain—allows to obtain such a configuration. Those configurations are however almost surely not observed in the $n \to \infty$ limit. This is well understood heuristically, but the argument can be made more precise. The probability of not having a starting point at $(0,n)$ is found to be $P(\text{no path at }(0,n)) = \frac{A_{n-1}}{A_n} \approx \left(\frac{4}{3}\right)^{-2n}$, where we used a result from [59]. Hence, this event is exponentially suppressed. To apply the tangent method we move the starting point of the outermost path to $(0, n+M)$, as indicated on Figure 2.10.

The entry point in the original domain is at $(L,n)$. We wish to compute, in the scaling limit, the most likely entry point at $\ell \equiv L/n$ as a function of $m \equiv M/n$. In a similar way as in the rectangoloid case, we consider

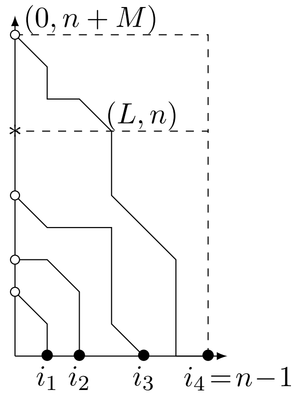



Figure 2.10: Extension of the domain to apply the tangent method. The starting point, assumed to be located at $(0, n)$, is moved to $(0, n+M)$. The entry point of the path is at $(L, n)$.

the sum

$$\sum_{L=0}^{n} \frac{A_n(\tau)|_{\tau^L}}{A_n} Y_{n,L}, \tag{2.48}$$

where $Y_{n,L}$ enumerates paths in the extension. Thanks to the results of the previous subsection we have $\frac{A_n(\tau)|_{\tau^L}}{A_n} = h_n(\sigma(\tau))|_{\tau^L}$, where $h_n(\sigma)$ is the generating function of the one-point function of the six-vertex model with domain wall boundary conditions, defined in [20], with the weights given in the previous subsection. The asymptotics of the problem is captured by the function

$$f(\sigma) \equiv \lim_{n\to\infty} \frac{1}{n} \log(h_n(\sigma)). \tag{2.49}$$

An in depth analysis of $f(\sigma)$ is provided in [15] using results of [98], see also [14] for an alternative computation. It is explicitly given in [15] when the argument is parametrized in a specific way. Particularizing their result for the present weights we get

$$\begin{aligned} f(\sigma(\xi)) &= \log\left(\frac{\sin(2\alpha\eta)\sin(\alpha\xi)\sin(4\eta-\xi)}{\alpha\sin(4\eta)\sin(\xi)\sin(\alpha(2\eta-\xi))}\right), \\ \sigma &= \sigma(\xi) = 2\cos(2\eta)\,\frac{\sin(4\eta-\xi)}{\sin(2\eta-\xi)}, \end{aligned} \tag{2.50}$$

with $\alpha = \frac{\pi}{\pi-2\eta}$, leading to

$$\tau(\sigma(\xi)) = \frac{\sin(2\eta+\xi)}{\sin(2\eta-\xi)} . \tag{2.51}$$

This defines implicitly $f$ as a function of $\tau$. We are now in a position to perform the saddle point analysis

$$\sum_{L=0}^{n} \frac{A_n(\tau)|_{\tau^L}}{A_n} Y_{n,L} \approx \int_0^1 \mathrm{d}\ell \int_0^{\min(m,\ell)} \mathrm{d}\phi \oint_0 \frac{\mathrm{d}\tau}{2\pi i}\, \mathrm{e}^{n(S_0(\ell,\tau)+S_1(\ell,m,\phi))}. \tag{2.52}$$

We used the Cauchy integral formula and some basic combinatorics to evaluate the asymptotics of $Y_{n,n\ell}$, and in so doing made an integral over the (rescaled) number of diagonal steps in the extension $\phi$ appear. We may get the most likely value of $\ell$ for fixed $m$ by extremizing the appropriate action $S_0(\ell,\tau) + S_1(\ell,m,\phi)$, over $\ell, \tau, \phi$, where

$$\begin{aligned} S_0(\ell,\tau) &= f(\xi(\tau)) - \ell\log(\tau), \\ S_1(\ell,m,\phi) &= \mathcal{L}(\ell+m-\phi) - \mathcal{L}(\ell-\phi) - \mathcal{L}(m-\phi) - \mathcal{L}(\phi) + \phi\log(\gamma), \end{aligned} \tag{2.53}$$

with $\mathcal{L}(x) \equiv x\log(x)$. A natural way to proceed would be to solve the saddle point equations, to eliminate $\xi$ in order to find $\tau(\ell)$ and $\phi(\ell)$ and deduce the equation of the family of lines parametrized by $\ell$. It should be noted however that the parametrization bears no particular significance, so that we are free to use the most convenient one. This trick is extensively used in the application of the tangent method [17, 20, 26, 29]. In our case we choose the parameter $\xi$, this gives

$$\ell = \tau\frac{df}{d\tau} = \frac{\tau(\xi)}{\tau'(\xi)} f'(\xi) \equiv \rho(\xi), \quad \gamma\frac{(\ell-\phi)(m-\phi)}{\phi\,(\ell+m-\phi)} = 1, \quad \frac{\ell+m-\phi}{\ell-\phi} = \tau \tag{2.54}$$

and finally

$$\frac{\ell}{m} = 1 + \frac{(1+\gamma)\tau(\xi)}{(\tau(\xi)-1)(\tau(\xi)+\gamma)} \equiv \beta(\xi). \tag{2.55}$$

The equation of the line through the endpoint $(0, 1+m)$ and the entry point $(\ell, 1)$ is $\frac{\ell}{m}(y-1) + x - \ell = 0$. Hence, the equation for the family of tangent lines, now parametrized by $\xi$, reads $\beta(\xi)(y-1) + x - \rho(\xi) = 0$. We end up with the following parametrization of the arctic curve:

**Theorem 2.2.1.** *The central portion of arctic curve of the QTHADT model in its fundamental domain, as predicted by the tangent method, is given by*

$$x(\xi) = \rho(\xi) - \beta(\xi)\frac{\rho'(\xi)}{\beta'(\xi)}, \quad y(\xi) = 1 + \frac{\rho'(\xi)}{\beta'(\xi)}\ , \qquad \xi \in [0, 2\eta]\ , \tag{2.56}$$

*where*

$$\rho(\xi)=\left[\alpha\Big(\cot(\alpha\xi)+\cot(\alpha(2\eta-\xi)\Big)-\cot(\xi)-\cot(4\eta-\xi)\right]\frac{\sin(2\eta+\xi)\sin(2\eta-\xi)}{\sin(4\eta)},$$
$$\beta(\xi)=\frac{\sin(2\eta+\xi)\sin(2\eta-\xi)}{\sin(\xi)\sin(4\eta-\xi)}\ . \tag{2.57}$$

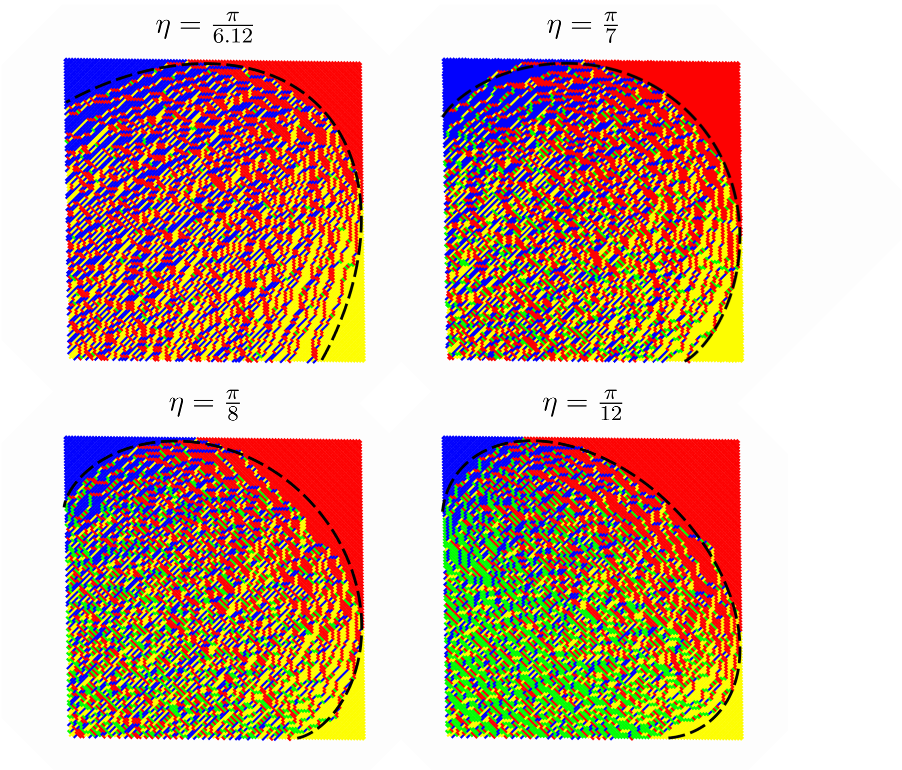



Figure 2.11: Typical configurations of the QTHADT for $n=100$ for several values of $\gamma=1+2\cos(4\eta)$. We show only the fundamental domain.

The range of $\xi$ in (2.56) is deduced from the constraint that $m\in[0,\infty]$, itself dictated by the geometry of our extension. It is easily checked that $x(\xi)=y(2\eta-\xi)$ hence the arctic curve is symmetric under $x\leftrightarrow y$, as expected. The range $\xi\in[0,2\eta]$ only describes a portion of the arctic curve from $(z,1)$ to $(1,z)$ with $z=x(0)=y(2\eta)$. We *expect* that the above expression extends to the range $\xi\in[\max(-2\eta,2\eta-\pi/2),\min(4\eta,\pi/2)]$, leading to two additional portions from $(0,z')$ to $(z,1)$ and from $(1,z)$ to $(z',0)$, with $z'=x(\min(4\eta,\pi/2))=y(\max(-2\eta,2\eta-\pi/2))$.

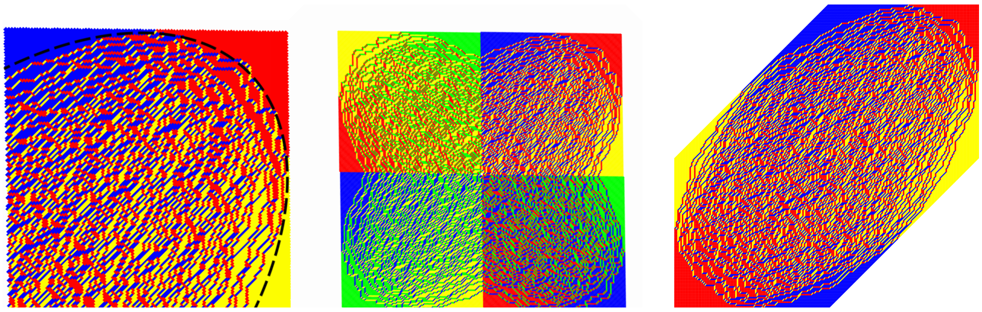


Figure 2.12: Typical configuration of the QTHADT for $n = 100$ in the limit $\gamma \to 0$, represented in three different ways. *Left*: Fundamental domain, with a plot of the predicted arctic curve. *Center*: 4-fold symmetric domino tiling on the whole quasi-square domain. *Right*: Rotationally symmetric rhombus tilings of a holey hexagon. Notice that a shear transformation is applied to the domain, so that the rotational symmetry is less apparent.

In Figure 2.11, we compare the prediction of Theorem 2.2.1 with numerics and find a good agreement. This completes the description of the arctic curve in the fundamental domain. The full arctic curve is obtained by iteratively applying $90°$ rotations to the fundamental domain arctic curve. Note that, in general (for $\gamma \neq 1$), these new copies differ from the analytic continuation of the fundamental domain copy.

The rotationally symmetric rhombus tilings of a holey hexagon—itself in bijection [96] with DPP—are a special case of QTHADT. For this rhombus tiling model, the fundamental domain has a rhombic shape drawn on the triangular lattice. Two extra copies of this domain obtained by two $120°$ rotations around a central triangular hole of size $2 \times 2 \times 2$ are added so as to form a quasi-regular hexagon of shape $n \times (n+2) \times n \times (n+2) \times n \times (n+2)$. The tiles are elementary rhombi covering two adjacent triangles and we demand that the tiling configurations be symmetric under $120°$ rotation. The version of this model shown in Figure 2.12 is obtained by applying a shear transformation, that turns the fundamental domain into a square. The tiling problem then has a NILP formulation which corresponds precisely to our setting but *without diagonal steps*, which can be imposed by fixing $\gamma = 0$ (equivalently $\eta = \pi/6$).

# Chapter 3

# Simulations

This Chapter is based on (unpublished) joint work with Jean-François de Kemmeter and Philippe Ruelle

In the previous chapters, multiple figures show large random tilings. A curious reader might wonder how to generate a tiling, or more generally a configuration $\mathcal{C}$ of a two-dimensional statistical model, according to a prescribed probability distribution $\pi(\mathcal{C})$. There are several ways of approaching this problem. By definition, a valid strategy is to construct the collection of *all* configurations (of a given domain), compute the associated probabilities and *sample* a configuration according to these probabilities. However, such an algorithm quickly becomes impractical since the number of configurations typically grows as $\mathrm{e}^{cn^2}$ (with $c > 0$ and $n$ the linear size), so that brute-force enumeration is already impossible for modest values of $n$. For example, there are more tilings of the Aztec diamonds of order $n = 25$ than protons in the observable universe.

## 3.1 Markov-chain Monte Carlo algorithms

In this section, we sketch some aspects of another approach, which is particularly versatile and also constitutes a vast field of study (see e.g. [100]). The idea is to start from an *easily constructed configuration* $\mathcal{C}_0$, like for example a tiling with only horizontal dominoes. A large number $N$ of random modifications called steps—each outputting a valid configuration—are then performed $\mathcal{C}_0 \to \mathcal{C}_1 \to \cdots \to \mathcal{C}_N$. Each of these steps $\mathcal{C} \to \mathcal{C}'$ only depends on the current configuration $\mathcal{C}$: the transition probability $p(\mathcal{C} \to \mathcal{C}')$ is independent of previous configurations and of the number of steps taken before, hence this whole process constitutes a *Markov chain* with a finite state space (we only consider models with a finite number of configurations). The probability of observing a configuration $\mathcal{C}$ after this procedure is $P_N(\mathcal{C}_0 \to \mathcal{C}) = \sum_{\mathcal{C}'} P_{N-1}(\mathcal{C}_0 \to \mathcal{C}')p(\mathcal{C}' \to \mathcal{C})$, with $P_0(\mathcal{C}_0 \to \mathcal{C}) = \delta_{\mathcal{C},\mathcal{C}_0}$. The goal is to design a Markov chain in such a way that after enough steps $N$ the initial configuration $\mathcal{C}_0$ is forgotten and $P_N(\mathcal{C}_0 \to \mathcal{C}) \simeq \pi(\mathcal{C})$.

An important property shared by the Markov chains that are relevant in our numerical simulations is ergodicity. A finite state Markov chain is ergodic if $\exists k \in \mathbb{N}$ such that $\forall\, \mathcal{C}_0, \mathcal{C}$ we have $P_k(\mathcal{C}_0 \to \mathcal{C}) > 0$. Notice that $k$ is the same for all $\mathcal{C}_0$ and $\mathcal{C}$. For finite Markov chains, this property is verified if every state can be reached from any other by a finite number of steps and if the probability of remaining in the same state at any time step is non-vanishing. We have the following theorem

**Theorem 3.1.1.** *(Doeblin [101]) A finite state ergodic Markov chain with transition probabilities $p(\mathcal{C}' \to \mathcal{C})$ has a unique distribution $P(\mathcal{C})$ that satisfies*

$$P(\mathcal{C}) = \sum_{\mathcal{C}'} P(\mathcal{C}')p(\mathcal{C}' \to \mathcal{C}) \tag{3.1}$$

*called its stationary distribution. We also have $\forall\, \mathcal{C}_0$*

$$\lim_{N\to\infty} P_N(\mathcal{C}_0 \to \mathcal{C}) = P(\mathcal{C}). \tag{3.2}$$

Hence, ergodicity suffices to ensure that the Markov chain indeed forgets the initial configuration $\mathcal{C}_0$ and that the probability to observe a given configuration $\mathcal{C}$ converges to a unique distribution $P(\mathcal{C})$, when the number of steps goes to infinity.

It remains to devise the transition probabilities in such a way that $P(\mathcal{C}) = \pi(\mathcal{C})$. Using the property $\sum_{\mathcal{C}'} p(\mathcal{C} \to \mathcal{C}') = 1$ Equation (3.1) is equivalent to $\sum_{\mathcal{C}'\neq\mathcal{C}} P(\mathcal{C})p(\mathcal{C} \to \mathcal{C}') = \sum_{\mathcal{C}'\neq\mathcal{C}} P(\mathcal{C}')p(\mathcal{C}' \to \mathcal{C})$, which is called the global balance condition (GBC). A stronger but easier to check condition ensuring that the GBC holds is the detailed balance condition (DBC), namely

$$P(\mathcal{C})p(\mathcal{C} \to \mathcal{C}') = P(\mathcal{C}')p(\mathcal{C}' \to \mathcal{C}). \tag{3.3}$$

If we can find $p(\mathcal{C} \to \mathcal{C}')$ such that the DBC is satisfied for $P(\mathcal{C}) = \pi(\mathcal{C})$, then Theorem 3.1.1 ensures that the Markov chain converges to the prescribed distribution $P(\mathcal{C}) = \pi(\mathcal{C})$, which is hoped to be well approximated by $P_N(\mathcal{C}_0 \to \mathcal{C})$ for large $N$.

The design of an ergodic Markov chain that satisfies the DBC (or only the GBC) must be adapted to the model under consideration, but some fairly general techniques exist [100]. We now show how to devise a Markov chain that generates QTHADT with a general weight $\gamma$. In other words, we want to generate tilings with a probability $\pi(\mathcal{C}) = \frac{w(\mathcal{C})}{A_n(\gamma)}$, where $w(\mathcal{C})$ is $\gamma$ to the power of the number of $(1, -1)$ elementary steps of $\mathcal{C}$ in its NILP description (see Figure 2.9). During each step of the Markov chain, only *local* modifications are performed, as they involve moves of pairs of neighbouring dominoes. These moves are of three types. The first possibility is to flips $2 \times 2$ blocks of aligned dominoes ( $\to$ or $\to$ ), strictly *inside the fundamental domain*. In terms of NILP these moves (locally) modify the shape of a path, but do not modify its extremities. The second kind of move is a flip of a $2\times2$ block that overlaps with the periodic boundaries of the fundamental domain. In the NILP picture, these moves simultaneously shift a pair of extremities a unit distance toward the center or away from it. Hence, performing only those two kinds of move does not lead to ergodicity since only tilings with the same number—previously noted $p$—of NILP as the initial configuration $\mathcal{C}_0$ can be reached. The missing ingredient is a third kind of move that we call “cross-moves”. They create or annihilate pairs of starting and ending points near the central cross (see Figure 3.1).

We first show exactly how those moves are performed in the case $\gamma = 1$. Every step of the Markov chain goes as follows: select uniformly at random a position $(i, j)$ in the fundamental domain, on the lattice on which the corners of the dominoes are located. If the $2 \times 2$ block whose upper vertex $\bullet$ at $(i, j)$ is entirely in the bulk of the fundamental domain, or on the periodic boundary, then perform a flip $\leftrightarrow$ if possible.

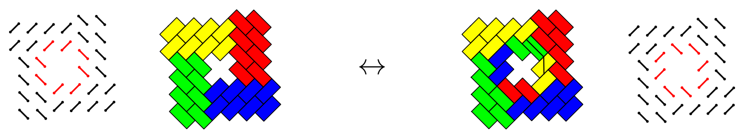


Figure 3.1: *Left*: "All-horizontal" configuration for $n = 3$, in the perfect matching picture and the tiling picture. *Right*: Configuration for $n = 3$ having a single path in its fundamental domain. In both cases NILP of the fundamental domain are drawn. The two configurations are related to one another by cross-moves, where dimers are shifted along the cycle surrounding the central cross, which is best seen in the perfect matching description.

If $(i, j)$ is adjacent to the cross, perform a cross-move if possible. If the selected site does not allow any move, *the configuration for this step remains unchanged*. This—perhaps counter-intuitive—way of dealing with "forbidden moves" is ubiquitous in Markov-chain simulations [100]. One can check that the Markov chain we just defined is ergodic, since any configuration of QTHADT can be attained by a finite sequence of steps and because the probability of letting the current configuration unchanged is always non-vanishing. Moreover, the transition probability is symmetric $p(\mathcal{C} \to \mathcal{C}') = p(\mathcal{C}' \to \mathcal{C})$ and hence the DBC is satisfied for the uniform distribution $\pi(\mathcal{C}) = \frac{1}{A_n(\gamma=1)}$.

The Markov chain for general $\gamma$ is nearly identical. The only difference is that if the selected site $(i, j)$ allows for a move $\mathcal{C} \to \mathcal{C}'$, then this move is performed with a probability equal to $\min\left(1, \frac{\pi(\mathcal{C}')}{\pi(\mathcal{C})}\right)$, called the *Metropolis acceptance probability* [102]. The DBC is satisfied for $\pi(\mathcal{C}) = \frac{w(\mathcal{C})}{A_n(\gamma)}$. Notice that in the Metropolis acceptance probability a lot of terms of the ratio $\frac{\pi(\mathcal{C}')}{\pi(\mathcal{C})}$ typically cancel—including the partition function $A_n(\gamma)$—which leads to simple expressions only involving the local change of weights which the move $\mathcal{C} \to \mathcal{C}'$ entails.

In actual implementations of Markov chains, it is always a good practice to consider a small domain and run the algorithm a large number of times to check that the correct distribution is obtained. Figure 3.2 shows a typical verification of our implementation. Another important issue is the choice of a number of steps $N$ such that the distribution is close enough to the stationary one. Upper bounds on the mixing time can be computed for some Markov chains [103], but may be difficult to get

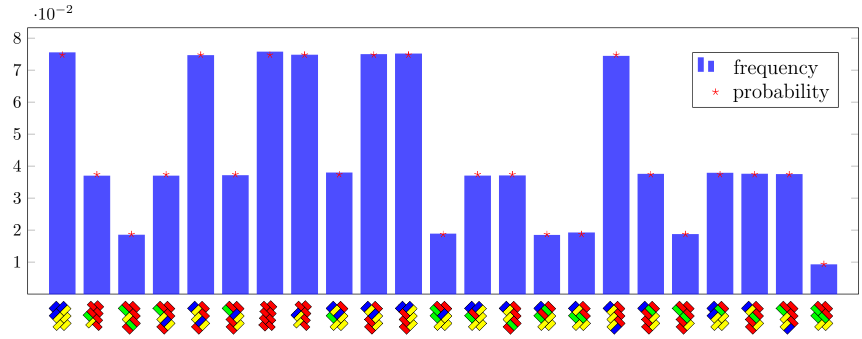



Figure 3.2: Observed frequencies of QTHADT for $n = 3$ and $\gamma = 1/2$, compared with the corresponding theoretical probabilities (measure on 230000 configurations).

and may not be useful in practice. Heuristic techniques such as the study of the auto-correlation time are sometimes used [104, 105]. In our case, we stop the simulation when the arctic curve is stabilized, as in [106]. Figures 2.11 and 2.12 were generated using this criterion. Such heuristics might be problematic in cases where the Markov chain exhibits a metastable behaviour [107].

The above construction of a Markov chain using local moves is quite versatile and may be easily extended to a wide variety of two-dimensional statistical models [100, 105, 106]. Several such Markov chains are used in this thesis. Most of the details of their implementations are spared to the reader, except for the one generating twenty-vertex model configurations, which was not discussed in the literature, see Chapter 4. We now briefly stress some limitations of local moves Markov chains and present several alternative algorithms. One issue is their painfully slow convergence in models that are critical, which is the case of the dimer model (a phenomenon called critical slowing down). Cluster algorithms alleviate this problem by performing moves on large clusters of local degrees of freedom [108–111]. A more practical way to boost the speed of simulations is to take advantage of the high parallelization capabilities of modern computers in order to perform local moves in parallel [112]. Another issue is the finite stopping time $N$. A very nice algorithm due to Propp and Wilson called *coupling from the past* almost surely stops after a finite time and is guaranteed[1] to output the stationary distribution [107]. The *shuffling algorithm*—again due to Propp and Kuperberg—combines the advantages of all these alternative algorithms at once [63].

## 3.2 Shuffling algorithm and fluctuations

There exists yet another kind of algorithm to generate tilings, which appears counter-intuitive at first glance. In the Markov chain approach, the starting point is a *trivially constructed configuration* of a domain whose brute force sampling cannot be achieved in practice. The idea of this other approach is to first consider a *trivial domain* whose configurations can easily be enumerated and that can hence be directly sampled. One then constructs, according to some—not entirely deterministic—rules, a configuration of a *larger domain* on which a direct enumeration is slightly more complicated. This procedure is repeated until one obtains a configuration of the desired domain. The rules must be designed in such a way that at the end of this construction, a valid configuration is output with the desired distribution. The shuffling algorithm discussed in Appendix A is an instance of this approach applied to Aztec diamonds. In this case, the trivial domain is an order $n = 1$ Aztec diamond having only two tilings, and at every step of the algorithm—that involves domino creation—a tiling of order $n$ is turned into one of order $n + 1$. A shuffling algorithm for the Novak half-hexagon exists [92] and a generalization of the shuffling algorithm [86] allows to generate tilings on any domain that can be embedded in an Aztec diamond.

We now know how to generate a tiling with a prescribed probability distribution. One can go further and generate a large amount of tilings in order to estimate various aspects of their statistics. In this section, we focus on domino tilings of the Aztec diamond (AD) in the large linear size limit. The associated statistics are accessible analytically and have been extensively studied [7, 56, 85, 113, 114]. Conveniently, the domino

[1] A practical requirement for an efficient implementation is for the set of configurations to admit a partial order (noted $\leq$) that is preserved by the moves [100, 107]. Such a partial order is easily found in models that have a height function description [107, 112].

shuffling algorithm introduced above provides an extremely efficient and reliable way to generate such tilings with the uniform distribution (see Appendix A for a proof). The numerical results presented in this section are computed using a collection of 100 000 tilings of the (uniform) Aztec diamond of order $n = 500$, generated using the shuffling algorithm. We optimized the implementation so that a tiling with $n = 500$ is produced in $\sim 6$ seconds on a modern laptop. This is orders of magnitude faster than a non-optimized Markov chain. For reference, an optimized coupling from the past algorithm runs for $\sim 900$ seconds on a Nvidia Tesla P100 GPU to generate a similar configuration [112]. To the best of our knowledge there is no numerical study of the fluctuations in the Aztec diamond, except brief ones in [112, 115].

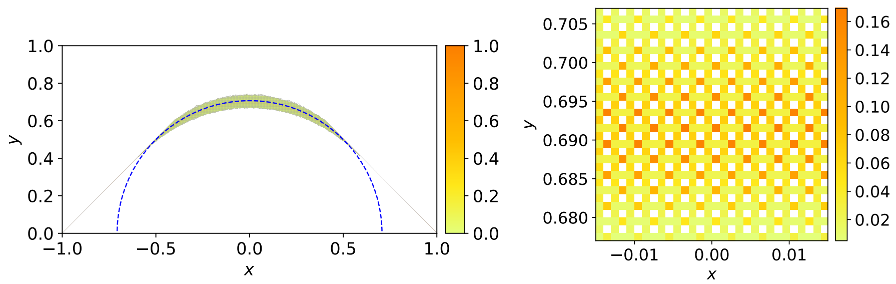



Figure 3.3: Vertical passage probability density of the uppermost path, where only integer and half-integer lattice coordinates where considered. *Left*: Rescaled domain picture. *Right*: Zoom around $(0, \frac{1}{\sqrt{2}})$ exhibiting lattice effects.

Despite the immutable appearance of tilings of large Aztec diamonds like the one of Figure 1.4, there is still randomness hidden within it. One way to uncover it is to *zoom on the rescaled Aztec diamond*. There are several distinct regions on which to zoom that may—and do—exhibit different behaviours, like the frozen region, the temperate region or most interestingly, the vicinity of the arctic curve. In this section, we focus on the latter case. More specifically we study NILP in the neighbourhood of the highest point on the arctic circle $(0, \frac{1}{\sqrt{2}})$. Figure 3.3 focuses on the uppermost path and shows its vertical passage probability density. The left picture shows that the density distribution is spread around the arctic circle, in the rescaled domain. This finite size effect disappears as $n$ grows so that one has to zoom in an $n$-dependent way—as done on the right picture—to actually see the fluctuations around the mean when $n \to \infty$. We will discuss shortly the precise zoom to perform in order to obtain a non-trivial result.

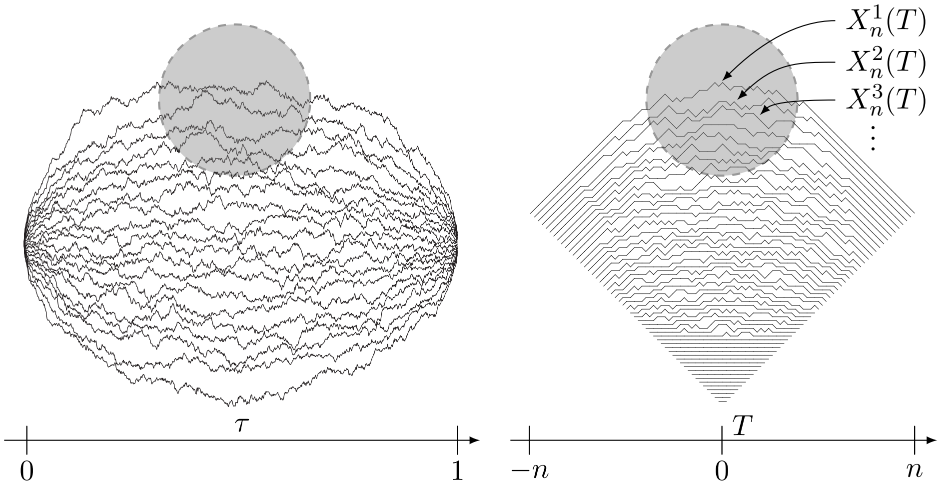



Figure 3.4: *Left*: Non-intersecting Brownian motions with the same starting and ending points, respectively at $\tau = 0$ and $\tau = 1$. Adapted from [116]. *Right*: Tiling of an Aztec diamond, in its NILP description. The uppermost paths, seen as functions of the abscissa $T \in [-n, n]$, are noted $X_n^i(T)$ with $i = 1, 2, \cdots$ indexing the paths from top to bottom. The reference frame is such that $X_n^1(\pm n) = -1/2$.

When $n \to \infty$ the NILP around $(0, \frac{1}{\sqrt{2}})$ locally resemble non-intersecting Brownian bridges around the maximum of their limit shape (see the grey circles in Figure 3.4). This apparently rough resemblance turns out to be quite deep as some aspects of the statistics of these two processes have been rigorously shown to be *exactly identical*, once appropriate rescalings (or zooms) are performed. More precisely, the *Airy process* describes the fluctuations of the (suitably rescaled) uppermost Brownian bridge mentioned above [117] and Johansson characterized the statistics of the uppermost path as follows:

**Theorem 3.2.1.** *(Johansson [113]) Let $X_n^1(T)$ be the vertical position of the uppermost path of a (uniform) AD at a given lattice abscissa $T$ (see Figure 3.4) and $\mathcal{A}_2(t)$ the Airy process. Then,*

$$\frac{X_n^1(2^{-1/6}n^{2/3}t) - \frac{n}{\sqrt{2}}}{2^{-5/6}n^{1/3}} \to \mathcal{A}_2(t) - t^2, \tag{3.4}$$

*as $n \to \infty$, in the sense of convergence of finite-dimensional distributions.*

In particular, this theorem implies that, for a fixed rescaled abscissa $t \in \mathbb{R}$ and for all $s \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_n^1(2^{-1/6}n^{2/3}t) - \frac{n}{\sqrt{2}}}{2^{-5/6}n^{1/3}} + t^2 \leq s\right) = \mathbb{P}(\mathcal{A}_2(t) \leq s) = F_2(s), \quad (3.5)$$

where $F_2(s)$ is the (GUE) Tracy-Widom distribution. This distribution also describes the (properly rescaled) largest eigenvalue of random matrices in the gaussian unitary ensemble (GUE) [118]. Its most distinctive features are its negative mean equal to $-1.771..$, its non-vanishing skewness equal to $0.224..$ and the asymmetric super-exponential decays of its left and right tails [118], see Figure 3.11.

In the above scaling, the shift by $-\frac{n}{\sqrt{2}}$ allows to measure the distance from the (putative) lattice position of the highest point of the arctic circle. The other details of the scaling might appear surprising at first glance. Let us give some heuristics to develop our intuition [38, 119]. Using the dimer density formula derived in [7] one can show that the density of paths on the $y-$axis around $(0, \frac{1}{\sqrt{2}})$ strictly vanishes from above and behaves like $\sqrt{\frac{1}{\sqrt{2}} - y}$ from below. Assuming that this square root behaviour also holds at the local scale and if $\lfloor \frac{n}{\sqrt{2}} \rfloor - Y_k$ denotes the vertical lattice position of the $k-$th uppermost path, then we expect

$$\int_0^{Y_k/n} \sqrt{y}\, \mathrm{d}y \simeq \frac{k}{n}, \quad (3.6)$$

which implies that $Y_k$ should be of order $n^{1/3}$. In the scaling limit, an infinite number of paths hence condensate to form the arctic curve. On the lattice, outermost paths should be separated by a distance of order $n^{1/3}$ and hence be typically far from each other. Since their mutual influence is limited, it is natural to expect a given path to behave roughly as if it were not surrounded by other paths. Let us approximate its behaviour by that of a single directed lattice path between fixed points separated by a distance $T$ on the lattice. For such a path, the standard deviation is of order $\sqrt{T}$, see e.g. Subsection 5.1.2. Since the vertical rescaling is already fixed to be of order $n^{1/3}$ we conclude that the horizontal rescaling should be of order $n^{2/3}$. Besides the vertical and horizontal rescalings, equation (3.5) also contains an additive $t^2$ term, which exactly compensates for the curvature of the paths around this point of the arctic curve.

Up to now we remained intentionally evasive on the precise definition of the Airy process and on the nature of the sophisticated mathematical tools needed in its study. Rather than elaborating on that ($\mathcal{A}_2(t)$ is defined in [113]), we collect some formulae from the mathematical literature and verify their validity. These results—most of which are gathered in [97, 120]—are often expressed in terms of a *Fredholm determinant* of some operator. It is the case for (3.5), where the GUE Tracy-Widom distribution is defined by the Fredholm determinant

$$F_2(s) = \det\left(\mathbb{I} - A_0|_{L^2[s,\infty[}\right), \tag{3.7}$$

where $A_0$ is a limiting case of the extended Airy operator $A_t$ whose kernel $A_t(x, y)$ is

$$A_t(x,y) = \begin{cases} \int_0^\infty \mathrm{e}^{-\xi t}\mathrm{Ai}(x+\xi)\mathrm{Ai}(y+\xi)\,\mathrm{d}\xi & \text{if } t \geq 0, \\ -\int_0^\infty \mathrm{e}^{\xi t}\mathrm{Ai}(x-\xi)\mathrm{Ai}(y-\xi)\,\mathrm{d}\xi & \text{if } t < 0, \end{cases} \tag{3.8}$$

with $\mathrm{Ai}(x)$ the Airy function of the first kind. The Fredholm determinant of an operator $K$ acting on functions $f \in L^2([s,\infty[)$ as $Kf(x) = \int_s^\infty K(x,y)f(y)\,\mathrm{d}y$ is defined[2] as

$$\det(\mathbb{I} - K|_{L^2[s,\infty[}) = 1 + \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} \int_s^\infty \cdots \int_s^\infty \det_{1\leq i,k\leq j}(K(x_i,x_k))\,\mathrm{d}x_1\cdots\mathrm{d}x_j, \tag{3.9}$$

with the condition that the restriction of $K$ to $[s,\infty[$ is a trace class operator, namely we have $\int_s^\infty K(x,x)\,\mathrm{d}x < \infty$. Definition (3.9) is impractical for numerical evaluations. The particular Fredholm determinant (3.7) can be expressed in terms of the Hastings-McLeod solution to the Painlevé II equation [118]. Numerical integration of this differential equation turns out to be quite involved [120, 121], the main problem being the sensitivity to the boundary conditions—which must be fixed asymptotically—in conjunction with machine precision. A much

[2] It could be seen as an extension of the von Koch formula used in (2.38) to the case of operators acting on infinite dimensional spaces, rather than matrices.

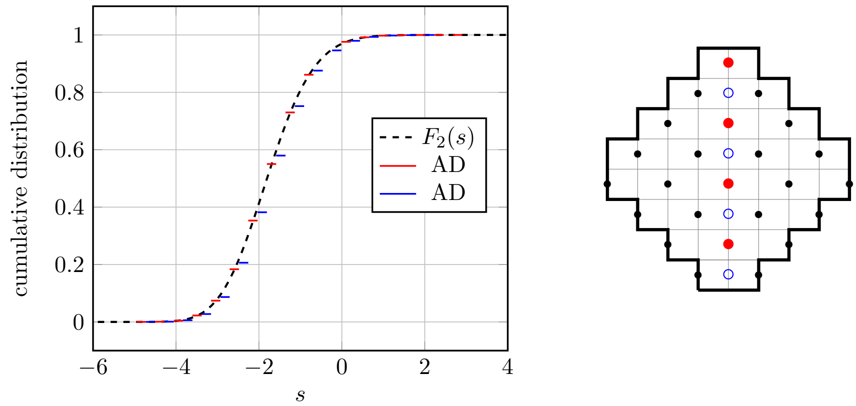
easier approach, which can be adapted to a wide variety of Fredholm determinants, is due to Bornemann [97,120]. The numerical evaluations of Fredholm determinants presented in this section is based on Bornemann's method.

### 3.2.1 Uppermost path

In this subsection, we verify numerically some consequences of Theorem 3.2.1, which concerns the statistics of the uppermost path alone. Figure 3.5 presents a verification of equation (3.5) at $T = 0$. The table in this figure compares some important moments of the measured and theoretical distributions, namely the mean, variance, skewness and excess kurtosis[3]. In addition to showing a good agreement with (3.5), this figure also highlights an unexpected dependence in the "parity" of the underlying lattice points, which is a *lattice effect*. The precise nature of these effects depends on the definition of the NILP, of which there are many. Lattice effects manifest themselves dramatically when considering probability densities, as seen on Figures 3.3 and 3.6. It should be stressed that probability density functions (PDF's) are not predicted by Theorem 3.2.1. It only describes joint cumulative distribution functions (CDF's) with finitely many variables, which are totally blind the aforementioned lattice effects.

Lattice effects are a consequence of the discrete nature of the elementary steps, which favours some passage points over others. A part of the explanation is that the access to certain points may be limited to one kind of elementary step, like for the open blue dots in Figure 3.5. On the other hand, the elementary step composition of a portion of path is not uniform and depends on its average slope. Assuming that lattice effects only modulate the PDF's by a multiplicative factor, a simple heuristic argument based on the analysis of the elementary step composition of a single lattice path with fixed extremities of equal height gives a quite convincing result, as seen on Figure 3.6. In the remainder of this section, with the exception of Figures 3.8 and 3.9, *we get rid of lattice effects*

[3] The kurtosis of the normal distribution is equal to 3. Excess kurtosis measures the departure from that value.




| | Mean | Variance | Skewness | Excess Kurtosis |
|---|---|---|---|---|
| | $\mu = \mathbb{E}[X]$ | $\sigma^2 = \mathbb{E}[(X-\mu)^2]$ | $\mathbb{E}[\left(\frac{X-\mu}{\sigma}\right)^3]$ | $\mathbb{E}[\left(\frac{X-\mu}{\sigma}\right)^4]-3$ |
| AD | $-1.676..$ | $0.819..$ | $0.215..$ | $0.0752..$ |
| $F_2'(s)$ | $-1.771..$ | $0.813..$ | $0.224..$ | $0.0935..$ |

Figure 3.5: Verification of (3.5) at $T=0$, with $F_2(s)$ computed from (3.7). The colours correspond to the parity of $X_n^1(T=0)$, see the picture on the right. The table shows some moments (with their respective definition for a generic random variable $X$) of the corresponding PDF's, after smoothing out lattice effects.

by counting the number of "hits" in "boxes" containing two or three neighbouring passage points.

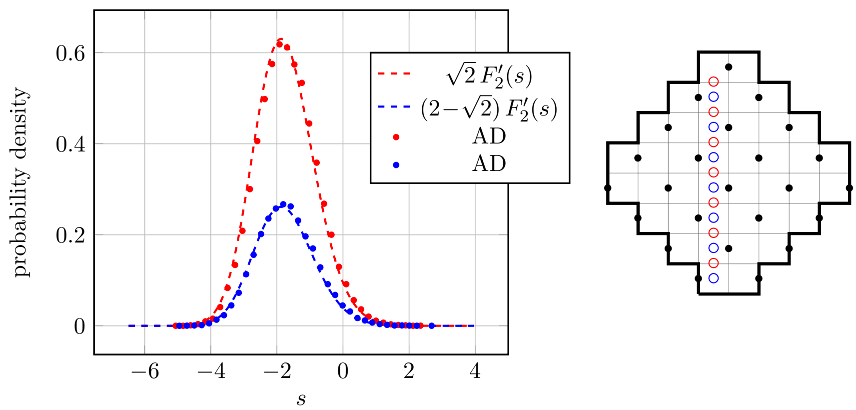



Figure 3.6: Probability density of the (rescaled) vertical coordinate of the uppermost path at $T=-1/2$. Colours correspond to the parity of $X_n^1(T=-1/2)$, see the right picture. Multiplying the GUE Tracy-Widom PDF by some factor computed heuristically gives a good agreement.

The Airy process is translationally invariant in the $t$ direction [117]. Nothing changes if we inspect the passage probability at another fixed $T$, apart from lattice effects which can be estimated using the heuristics sketched above. The passage probability at a single fixed abscissa is hence completely described. Theorem 3.2.1 is more general, as it characterizes a whole *process*. We now consider two points of the uppermost paths of different rescaled abscissa $t$. The joint CDF is given [120, 122] in terms of a slightly generalized Fredholm determinant of an operator with a matrix kernel

$$\mathbb{P}(\mathcal{A}_2(t)\leq s_1, \mathcal{A}_2(0)\leq s_2) = \det\left(\mathbb{I} - \begin{pmatrix} A_0 & A_t \\ A_{-t} & A_0 \end{pmatrix}\Bigg|_{L^2[s_1,\infty[\oplus L^2[s_2,\infty[}\right), \tag{3.10}$$

the definition of which can be found in [97], where a numerical evaluation method is also provided. On Figure 3.7 we compare the joint PDF of the rescaled version of $X_n^1(T)$ and the joint PDF of (3.10) for a given $t = 1/2$. We also evaluate the covariance for several values of $t$. The theoretical value of $\mathrm{cov}(t) \equiv \mathrm{cov}(\mathcal{A}_2(t), \mathcal{A}_2(0))$ given by

$$\begin{aligned}\mathrm{cov}(t) &= \mathbb{E}(\mathcal{A}_2(t)\mathcal{A}_2(0)) - \mathbb{E}(\mathcal{A}_2(t))\,\mathbb{E}(\mathcal{A}_2(0)) \\ &= \int_{\mathbb{R}^2} s_1 s_2 \frac{\partial^2 \mathbb{P}(\mathcal{A}_2(t)\leq s_1, \mathcal{A}_2(0)\leq s_2)}{\partial s_1 \partial s_2} \mathrm{d}s_1 \mathrm{d}s_2 - \left(\int_{\mathbb{R}} s \frac{dF_2(s)}{ds}\,\mathrm{d}s\right)^2\end{aligned} \tag{3.11}$$

is confronted to numerical results in Figure 3.8.

Theorem 3.2.1 yields the joint probability distributions when *finitely many* intermediate points $t_1, t_2, \cdots, t_k$ are considered. Let us finish our analysis of the statistics of the uppermost path by considering a global quantity that involves *all the values* of $t$. In [113] Johansson conjectured that his result could be extended to a functional limit theorem (as done in [123] for another model), which would then imply that

$$\lim_{n\to+\infty} \mathbb{P}\left(\max_{t\in\mathbb{R}} \left(\frac{X_n^1(2^{-1/6}n^{2/3}t) - \frac{n}{\sqrt{2}}}{2^{-3/2}n^{1/3}}\right) \leq m\right) = F_1(m), \tag{3.12}$$

where $F_1(m)$ is the gaussian orthogonal ensemble (GOE) Tracy-Widom distribution, that describes the maximal eigenvalue of random matrices in the GOE [118]. Notice that in (3.12) the parabola $-t^2$ is *not*

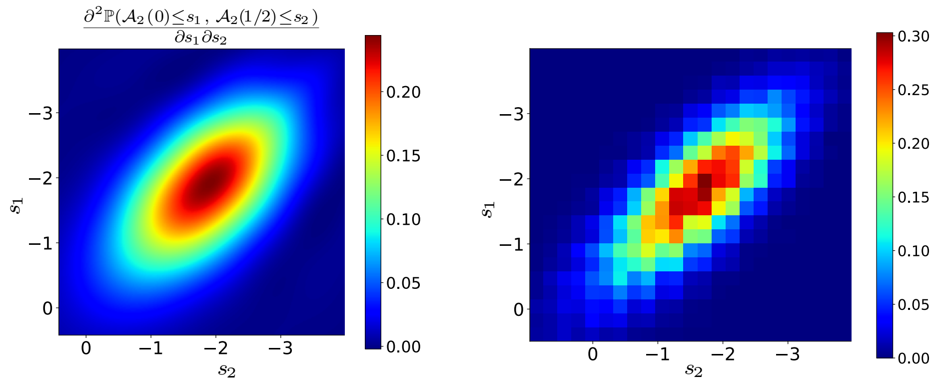



Figure 3.7: *Left*: Theoretical joint PDF of $\mathcal{A}_2(0)$ and $\mathcal{A}_2(t=1/2)$, computed from (3.10). *Right*: Corresponding joint PDF observed on our collection of AD.

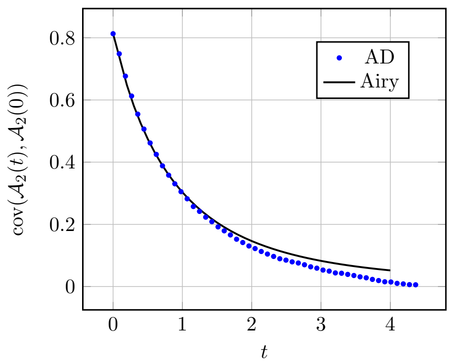


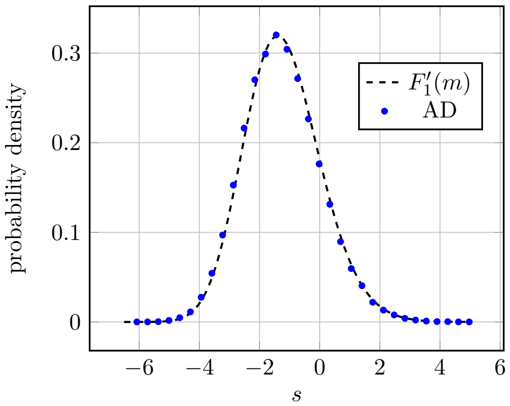

Figure 3.8: Covariance for the uppermost path at two different abscissae, computed from our collection of AD, compared with the theoretical result (3.11).

subtracted and notice also that the power of 2 in the denominator is different[4] than in (3.5). The GOE Tracy-Widom distribution $F_1(m)$ can be expressed [97] as

$$F_1(m) = \det\left(\mathbb{I} - V_A|_{L^2[m,+\infty[}\right), \quad V_A(x,y) = \frac{1}{2}\mathrm{Ai}\left(\frac{x+y}{2}\right), \tag{3.13}$$

where $V_A(x,y)$ is the kernel of $V_A$. Figure 3.9 presents a verification of (3.12), at the PDF level. As a consistency check, one can compare the means of the distributions of Figures 3.5 and 3.9 and indeed verify that $-1.771.. < -1.207...$

[4] It results from equation (1.14) of [123], where the right hand side should be $F_1(4^{1/3}\xi)$ as mentioned in [124].




| | Mean | Variance | Skewness | Excess Kurtosis |
|---|---|---|---|---|
| AD | −1.263.. | 1.619.. | 0.286.. | 0.148.. |
| $F_1'(m)$ | −1.207.. | 1.608.. | 0.293.. | 0.165.. |

Figure 3.9: Verification of (3.12), describing the maximum of the uppermost path. The table compares the moments of the GOE Tracy-Widom distribution $F_1(m)$ computed from (3.13) with those gotten from our collection of AD.

### 3.2.2 First few uppermost paths

It is natural to wonder about the behaviour of the paths directly below the uppermost one. It is widely believed that, after an identical rescaling as in (3.5), this collection of uppermost paths $X_n^i(T)$ (see Figure 3.4) converges in the sense of finite dimensional distributions to the *Airy line ensemble* $\mathcal{A}_2^i(t)$, $i = 1, 2, \cdots$. This ensemble also describes the uppermost paths of a system of non-intersecting Brownian bridges as shown in [117], where a definition of $\mathcal{A}_2^i(t)$ is also found. Using our conjecture and results from [120] we should have for $i = \mathcal{O}(1)$

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_n^i(2^{-1/6}n^{2/3}t) - \frac{n}{\sqrt{2}}}{2^{-5/6}n^{1/3}} + t^2 \leq s\right) = \mathbb{P}(\mathcal{A}_2^i(t) \leq s) = F_2^i(s), \quad (3.14)$$

where

$$F_2^i(s) = \sum_{j=0}^{i-1} \frac{(-1)^j}{j!} \left[\frac{\mathrm{d}^j}{\mathrm{d}z^j} \det\left(\mathbb{I} - zA_0|_{L^2[s,+\infty[}\right)\right]_{z=1}. \quad (3.15)$$

The distributions $F_2^i$ are exactly those of the $i$-th largest eigenvalue of GUE matrices, in particular we have $F_2^1(s) = F_2(s)$. Figure 3.10 shows

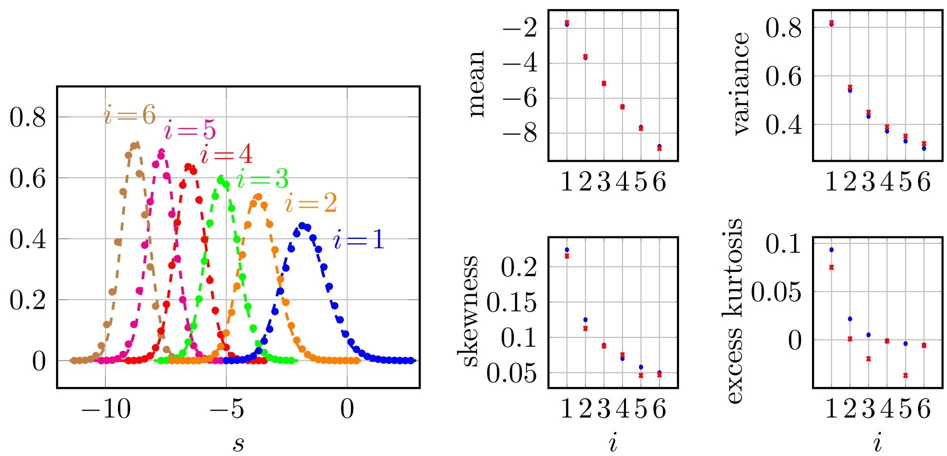



Figure 3.10: Verification of (3.14), describing the vertical position of the first few uppermost paths at $T = -1/2$. *Left*: Direct comparison between $(F_2^i)'(s)$ (dashed curves) computed using (3.15) and corresponding probability densities observed in the AD (dots). *Right*: The blue dots are moments of $(F_2^i)'(s)$ while the red crosses are moments computed from our collection of AD.

a comparison between the numerics on the AD and the prediction for $T = -1/2$, at the PDF's level. As $i$ increases, the PDF's of the first few uppermost paths have a decreasing mean and variance, as one may expect from the intuition that paths tend to "squeeze" other paths below them.

We might wonder if the distributions $F_2^i(s)$ are fundamentally different from one another or not. To highlight those differences we standardize the PDF's, i.e. translate and scale it to set $\mu = 0$ and $\sigma = 1$, for $i = 1$ and $i = 2$ in Figure 3.11. It appears that these distributions are quite similar, once the expected dilatation and shift are removed. However, they also differ, most notably on the asymptotic behaviour of their tails. Regarding the right tail, $(F_2^1)'(s)$ decreases less rapidly than $(F_2^2)'(s)$. This is to be expected since the second path is limited in its upward fluctuations by the uppermost path. At the left tail the exact opposite behaviour is observed. The mechanism behind this behaviour is less obvious, but it is consistent with the idea that the uppermost path is kept from going downward by more paths below it.

The Airy line ensemble is a process that does not only describe paths separately. It also encompasses the interaction between them. The joint

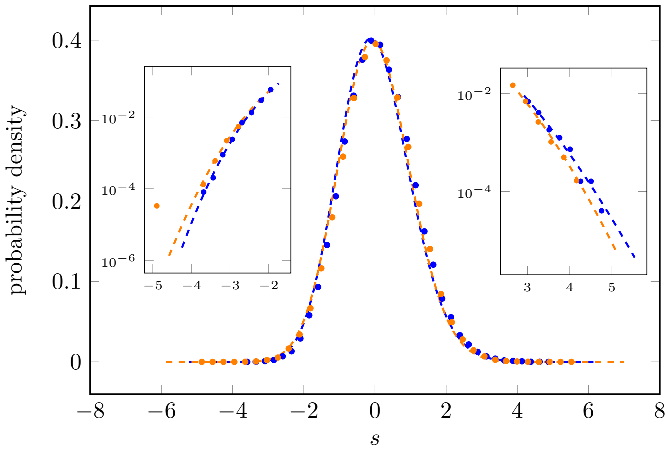



Figure 3.11: Standardized PDF's of the first two uppermost paths of Figure 3.10, $i = 1$ in blue and $i = 2$ in orange. These two PDF's most notably differ at their tails as seen on the insets (in log scales). Numerical evaluation of the tail behaviour requires a large collection of tilings, since they describe rare events.

CDF for the first two paths, at a common time $t$, is [97]

$$\mathbb{P}(\mathcal{A}_2^1(t) \le x, \mathcal{A}_2^2(t) \le y)$$
$$= \begin{cases} F_2(x) & \text{if } x \le y, \\ F_2(y) - \left[\frac{\partial}{\partial z}\det\left(\mathbb{I} - \begin{pmatrix} zA_0 & A_0 \\ zA_0 & A_0 \end{pmatrix}\Big|_{L^2[y,x]\oplus L^2[x,+\infty[}\right)\right]_{z=1} & \text{if } x > y. \end{cases} \tag{3.16}$$

As a last check of our conjecture, we highlight the correlation between the two uppermost paths by using (3.16) to study the sum of the rescaled vertical position of the first two paths at a fixed $t$, whose CDF should be [120]

$$\mathbb{P}(\mathcal{A}_2^1(t)+\mathcal{A}_2^2(t) \le s) = \int_{-\infty}^{+\infty} \frac{\partial}{\partial \xi}\mathbb{P}(\mathcal{A}_2^1(t) \le \xi, \mathcal{A}_2^2(t) \le s-x)\Big|_{\xi=x} \mathrm{d}x. \tag{3.17}$$

In Figure 3.12 this prediction is compared with the statistics observed numerically on the AD, at the PDF level. This PDF is more spread than the convolution of $(F_2^1)'$ with $(F_2^2)'$ which is the PDF we would get if the first two paths were independent. This can be viewed as a consequence of the non-intersection constraint.

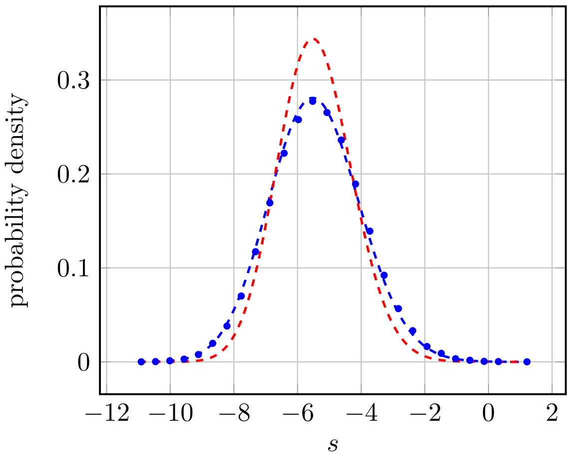



Figure 3.12: Study of the correlation between the two uppermost paths at $T = 0$. In blue, probability density of the sum of the vertical position of the two uppermost paths. In red, the convolution of the PDF's of the first two paths shows what would happen if the first two paths were independent. The dashed curve are computed from (3.16) and (3.17), the dots are AD results.

### 3.2.3 Concluding remarks

The precise shape of arctic curves strongly depends on the model in which they arise. For models that can be described in terms of NILP, the Airy line ensemble is believed [38, 117] to universally appear around generic points of the arctic curve when the density of NILP is low. This loosely defined notion of universality (see [125] for an introduction) is particularly well exemplified by the GUE Tracy-Widom distribution $F_2(s)$. This distribution has been found in a wide variety of seemingly unrelated situations such as the interface growth in liquid crystals [126] or in self assembling particles of various sizes [127], the zero-point fluctuations of the leftmost particle in a one-dimensional system of non-interacting fermions in a harmonic trap [119], the longest increasing subsequence of a large permutation taken uniformly at random [128] and the list goes on. We showed that some of these universal distributions can be accessed numerically by using the shuffling algorithm, which allows to gather sufficiently many large random tilings.

# Chapter 4

# Vertex models

This chapter is based on (unpublished) joint work with Jean-François de Kemmeter and Philippe Ruelle and on [17] written with Philippe Di Francesco and Emanuel Guitter

---

Beneath the apparent banality of an ice-cube are hidden fascinating thermodynamical properties. If we momentarily forget our daily experience, we might be surprised to find that ice floats on water, which contradicts the intuition that molecules (or atoms) are more tightly arranged in the solid state than in the liquid one. A more technical property is the non-vanishing entropy of ice at zero temperature [129], the so-called *Pauling's residual entropy*. These features are both intimately related to the structure of ice. Ice is composed of $H_2O$ molecules, held together by hydrogen bonds. The two hydrogen atoms of each molecule form a—more or less—fixed angle that constrains the possible arrangements of the molecules, resulting in the lower density of the solid phase. Ordinary frozen water $I_h$ can be seen as a three-dimensional hexagonal lattice having oxygen atoms at its vertices and hydrogen atoms on its edges, close to the oxygen atom to which it is covalently bonded and farther away from the neighbouring oxygen atom with which it forms a hydrogen bond. The *ice rule* imposes that there are exactly two hydrogen atoms close to each vertex. Residual entropy is a consequence of the large number of configurations allowed by the ice rule.

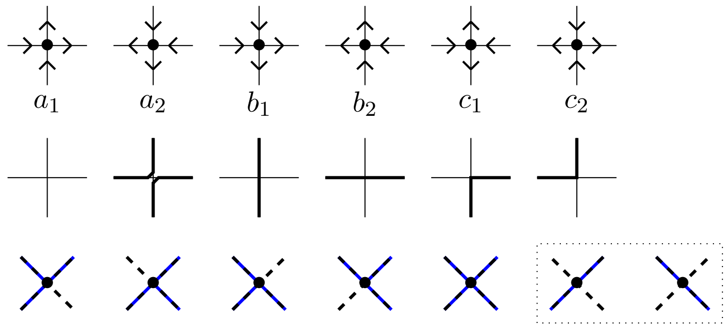



Figure 4.1: *Top*: The six possible arrow orientations around each vertex. *Middle*: Osculating paths representation of those six vertices. *Bottom*: Rules used in Figure 4.3 in the one-to-many correspondence between 6V configurations and domino tilings.

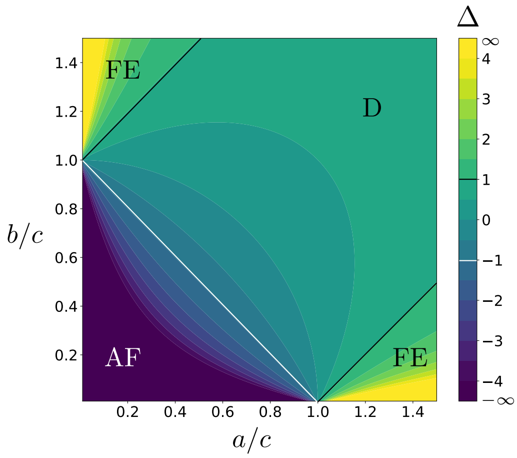
The six-vertex (6V) model was originally formulated as a very simplified version of this problem [130–132] that has the good taste of being *exactly solvable* [54]. It is defined on a two-dimensional square lattice. On each edge of this lattice lies an arrow that indicates to which oxygen atom the hydrogen atom on this edge is the closest. An arrow may also be seen as an electric dipole. As shown on Figure 4.1, the ice rule leads to $6 = \binom{4}{2}$ possible arrow configurations around a vertex, hence the name of the model. $H_2O$ molecules compressed between two layers of graphene have been observed[1] to form solid nanometre-thick sheets and to arrange in a square lattice [134]. In a model of ice, each of the six possible vertices would be assigned the same statistical weight. We consider here a more general situation where they carry weights denoted $a_1$, $a_2$, $b_1$, $b_2$, $c_1$ and $c_2$. The model then also describes ferroelectric and anti-ferroelectric materials [54]. The partition function is a weighted sum over configurations $\mathcal{C}$ allowed by the ice rule and the boundary conditions

$$Z = \sum_{\mathcal{C}} \prod_{i=1}^{2} a_i^{N_{a_i}} b_i^{N_{b_i}} c_i^{N_{c_i}}, \tag{4.1}$$

where $N_{a_i}$, $N_{b_i}$ and $N_{c_i}$ are the number of vertices of each type in configuration $\mathcal{C}$.

We restrict our attention to the subset of weights that are symmetric under arrow reversal, namely $a_1 = a_2 = a$, $b_1 = b_2 = b$ and $c_1 = c_2 = c$.

[1]The validity of this observation is still a matter of debate [133].




| order | per. | DWBC |
|---|---|---|
| D-FE | 1 | 2 |
| D-AF | $\infty$ | $\infty$ |

Figure 4.2: *Left*: Phase diagram of the 6V model, for periodic boundary conditions or DWBC. A common rescaling of all the vertices only changes the partition function by an overall factor, hence only the ratios $a/c$ and $b/c$ are relevant. We trace the contours of $\Delta$ in the range $[-4, 4]$. *Right*: Description of the phase transitions depending on the boundary conditions, showing results from [54, 98, 135].

For periodic boundary conditions, integrability techniques such as the coordinate Bethe ansatz allows to explicitly compute the free energy per site [54]. The computation shows that depending on the value of $\Delta \equiv \frac{a^2+b^2-c^2}{2ab}$, the model can be in three distinct phases, see Figure 4.2. When $\Delta > 1$ it is in the ferroelectric (FE) phase, when $-1 < \Delta < 1$ it is in the disordered (D) phase and when $\Delta < -1$ it is in the anti-ferroelectric (AF) phase. In the FE phase, the system is *frozen*: arrows on parallel edges all point in the same direction, in the thermodynamic limit. If $a > b + c$, either (almost) all vertices are of type $a_1$ or of type $a_2$. If $b > a + c$, they are either $b_1$ or $b_2$. As the name suggests, the D phase is disordered, but a peculiar one, since (connected) correlations decay with distance as an inverse power law, rather than exponentially. The correlation length is hence infinite and the model *critical* in this entire phase. In the AF phase, arrows tend to alternate in direction, but (unlike the FE case) become almost surely frozen only in the limit $\frac{a}{c} \to 0$ and $\frac{b}{c} \to 0$. In this limit, vertices alternate between $c_1$ and $c_2$. There are two possible arrangements. In the AF phase, correlations decay exponentially.

At this point the reader might have guessed that, as for tiling models, the 6V model is strongly influenced by boundary conditions. This property was not *a priori* obvious and was first substantiated in [135], where a dependence of the free energy per site on the boundary conditions was demonstrated. Some boundary conditions lead to an arctic phenomenon. The heuristics surrounding it may also be extended to the 6V model since its configurations can be described in terms of osculating paths, using the bijection shown on Figure 4.1. These paths differ from those previously discussed in that they are allowed to share a vertex at *kissing points* and in their weighting, which not only takes into account the nature of the elementary steps, but also the exact order in which they are taken.

The domain wall boundary conditions (DWBC) is presented on Figure 4.3. It was introduced in [136] to study scalar products in the XXZ Heisenberg quantum spin-chain, using the quantum inverse scattering method (QISM), see [137] for a review and Figure 6.10 therein for an explanation of the name "domain wall". By solving the recursion relation found by Korepin [136], Izergin proved [138, 139] what is now called the Izergin-Korepin formula for the partition function of the *inhomogeneous* 6V model with DWBC. The model is inhomogeneous in the sense that the possible weights of a vertex depend on its position $(i, j)$. The only constraint on this very general setting is that the $\Delta_{i,j}$—defined in the obvious way—must be independent of $(i, j)$.

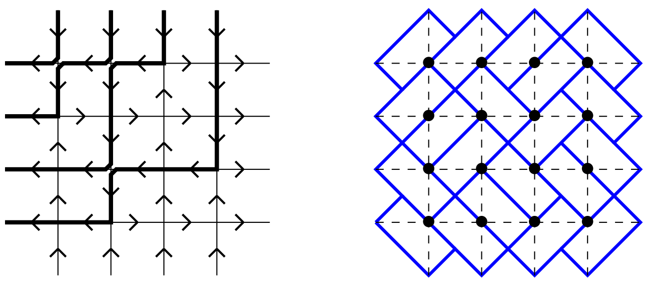


Figure 4.3: *Left*: Configuration of the 6V model with DWBC, along with its osculating paths description. *Right*: One of the corresponding possible tiling of the AD, obtained using the rules of Figure 4.1.

The connection with tiling models goes beyond the heuristics of the arctic phenomenon [98], as the 6V model with DWBC may be seen as an interacting version of the AD tiling model, see Figure 4.3. The correspondence is not one-to-one, due to the ambiguity in the $c_2$ vertex, but relates partition functions of the two models, when $a = b = 1$ and $c = \sqrt{2}$ ($\Delta = 0$). Maybe unsurprisingly, the arctic curve is found to be a circle at this special point of the phase diagram called the free-fermion point [140]. It is natural to wonder what happens for other weights. The phase diagrams look the same for periodic and domain wall boundary conditions (see Figure 4.2), the only difference being the order of the phase transition [98, 135], and the nature of the phases (see simulations in [106,109,111,141]). The FE phase is the least affected one as it is again almost surely frozen in the thermodynamic limit. The only difference is that now, if $a > b + c$, the north-west half of the domain is filled with $a_2$ vertices, the south-east half with $a_1$ vertices and there are $\sim n$ vertices of type $c_2$ at the junction of these regions, on a diagonal of the square domain. If $b > a + c$, vertices $b_1$ and $b_2$ each constitute approximately half of the vertices, the remainder being $c_1$ vertices located on the other diagonal. In the D phase, a central disordered region is separated from frozen regions by an arctic curve. Correlations in the disordered region decay algebraically. In the AF *phase*, three kinds of *regions* now coexist, that we may label by D, FE and AF as they resemble configurations in the corresponding phases, when periodic boundary conditions are applied. An arctic curve again separates an outer frozen FE region from a D region, but a third inner AF region also appears around the center of the domain. In the AF region, vertices tend to alternate between $c_1$ and $c_2$ and correlations decrease exponentially with distance. This AF region is separated from the D region by a star-shaped "antarctic" curve, see [111] for illustrations.

The arctic curve of this model was derived for fully general weights in [14, 15, 20]. The first successful approach [14, 15] was based on a non-local observable called the emptiness formation probability (EFP), whose asymptotic was extracted using the technical assumption that the roots of some saddle point equations condensate. The EFP technique yields an arctic curve whose shape is the solution of $F(x, y, z) = 0$ and $\partial_z F(x, y, z) = 0$, which exactly describes the envelope of a family of curves parametrized by $z$, which moreover turned out to be straight lines. This observation was then elevated to a universal geometric principle by the development of the tangent method [20], which provides an alternative derivation of the arctic curve. To the best of our knowledge, no analytical results are available for the antarctic curve[2].

## 4.1 Six-vertex model with pDWBC

There exist other boundary conditions (all resembling DWBC in some way) for which the 6V model exhibits an arctic curve [105]. The reflecting boundary conditions have been studied analytically in [29, 30, 142]. In this section, we study the 6V model on a rectangular $n \times s$ (with $s \leq n$) domain with partial domain wall boundary conditions (pDWBC). These boundary conditions are a mixture of fixed and free boundary conditions. In the case of pDWBC, arrows on the upper boundary of the rectangle are free, see Figure 4.4. Because of the ice rule, there are exactly $s$ down arrows on the upper boundary. For $s = n$, pDWBC exactly reproduce DWBC. The path description of the configurations with pDWBC involves $s$ osculating paths. They start from the $s$ edges on the left boundary and end at $s$ distinct edges on the upper boundary, whose position is not imposed. Heuristic considerations are quite similar to previous cases, in particular the heuristics for the free endpoints is analogous to that for the starting and ending points on the periodic boundary of QTHADT. The difference however, is the existence of kissing points that allows for a new kind of frozen region of high path density composed of $a_2$ vertices. An arctic phenomenon occurs in the D and AF phases [141]. In the FE case, the partition function was studied in detail in [143], but as for DWBC, with no arctic phenomenon.

For DWBC, one may impose that $a_1 = a_2$, $b_1 = b_2$ and $c_1 = c_2$, without loss of generality. Indeed, for fixed boundary conditions the number of horizontal and vertical steps is fixed, and so is the difference between the number of left-turns and right-turns. This leads to four *conservations laws*, namely it fixes $N_{a_1} + N_{a_2} + N_{b_1} + N_{b_2} + N_{c_1} + N_{c_2}$, $2N_{a_2} + 2N_{b_2} + N_{c_1} + N_{c_2}$, $2N_{a_2} + 2N_{b_1} + N_{c_1} + N_{c_2}$ and $N_{c_2} - N_{c_1}$, which enable one to

[2] The shape of an antarctic curve has been derived in [9] for the two-periodic AD, see also [80] for more complicated weights. In this context, the disordered region is called "liquid" and the central AF region is called "gaseous". It should be stressed that a liquid or gaseous region in the 6V model has nothing to do with water or steam.

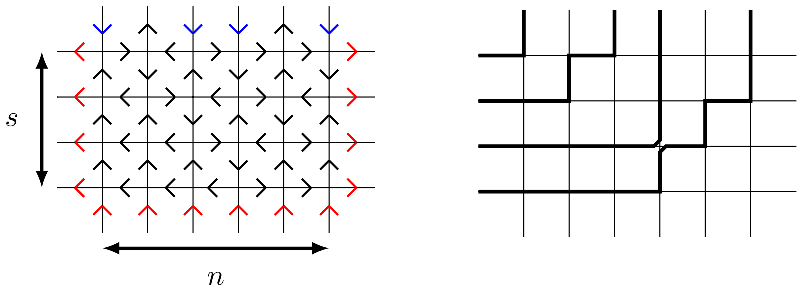



Figure 4.4: *Left*: Configuration of the 6V with pDWBC for $s = 4$ and $n = 6$. The arrows in red are fixed. The upper boundary remains free and must contain $s$ down arrows (in blue) due to the ice rule. *Right*: Osculating paths description of this configuration.

always impose the aforementioned conditions. In the pDWBC case, the number of horizontal steps is not fixed and imposing $a_1 = a_2$, $b_1 = b_2$ and $c_1 = c_2$ is a *restriction*. Asymmetric weights also lead to interesting arctic phenomena for other boundary conditions [16, 144]. For pDWBC it was shown numerically that this asymmetry deforms the arctic curve [105]. In this section, we use the tangent method to compute the arctic curve in the symmetric (free-fermion) case $a = b = 1$ and $c = \sqrt{2}$.

### 4.1.1 Family of tangent lines

The geometrical setting for the tangent method is identical to that of the DWBC case, see Figure 4.5. The two derivations differ in the only non-trivial part of the tangent method, namely the computation of the asymptotics of the boundary one-point function $H_{n,s}^{(k)}$, encompassed in a function called $r_\sigma(z)$. For completeness and to introduce the notations, we briefly recall the determination of the family of tangent lines $F(x, y, z)$ in terms of $r_\sigma(z)$, which is identical to that of [20].

The one-point function $H_{n,s}^{(k)}$ is the probability of observing the only type $c$ vertex of the bottom row in the $k$-th column (with $1 \leq k \leq n$), see Figure 4.4. To its left are $k-1$ type $b$ vertices and to its right $(n-k)$ type $a$ ones. In the osculating paths description, $H_{n,s}^{(k)}$ is the probability that the outermost path takes its first upward step at position $k$. It is

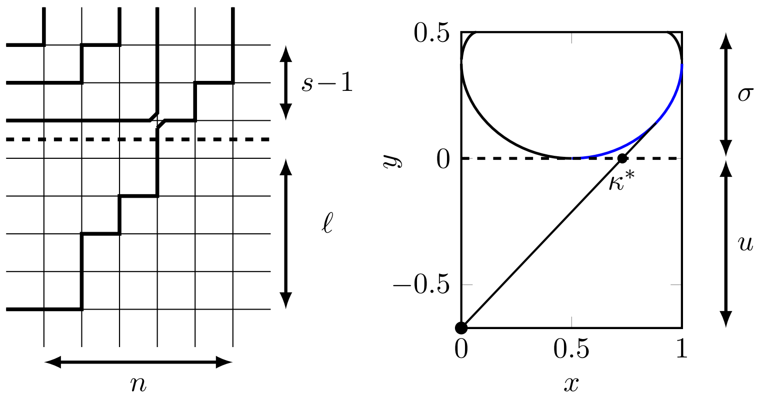



Figure 4.5: *Left*: Tangent method set-up, on the lattice. *Right*: Rescaled domain picture. The blue portion is the one accessible using this extension.

given by

$$H_{n,s}^{(k)} = b^{k-1}\, c\, a^{n-k} \frac{Z_{n,s-1}^{(k)}}{Z_{n,s}}, \tag{4.2}$$

where $Z_{n,s-1}^{(k)}$ is the partition function of a 6V model on a $n \times (s-1)$ grid with slightly modified pDWBC, in the sense that the bottom line has a down arrow at position $k$. A natural quantity to consider when dealing with *inhomogeneous* models is—as will become clear with (4.19)—the generating function of the one-point function

$$h_{n,s}(z) \equiv \sum_{k=1}^{n} H_{n,s}^{(k)}\, z^{k-1}. \tag{4.3}$$

All we need to know about its asymptotics is captured in

$$r_\sigma(z) \equiv \lim_{n\to+\infty} \frac{1}{n} z \frac{d}{dz} \log h_{n,s}(z), \tag{4.4}$$

where the aspect ratio $\sigma = \frac{s}{n}$ is kept fixed. Indeed, $r_\sigma(z)$ can be used to compute $S_\sigma(\zeta) \equiv \lim_{n\to\infty} \frac{1}{n} \log H_{n,s}^{(k)}$. The link is established by studying the asymptotics of (4.3),

$$h_{n,s}(z) \approx \int_0^1 \mathrm{d}\zeta\ \mathrm{e}^{n[S_\sigma(\zeta)+\zeta\log z]} \approx \mathrm{e}^{n[S_\sigma(\zeta^*)+\zeta^*\log z]}, \tag{4.5}$$

with $\zeta^*$ solution to the saddle point equation. We then have

$$\begin{cases} S'_\sigma(\zeta^*) = -\log z, \\ r_\sigma(z) = \zeta^*, \end{cases} \tag{4.6}$$

and hence $S_\sigma(\zeta^*)$ is related to $r_\sigma(z)$ through some kind of Legendre transform. The two relations of (4.6) may also be seen as being the inverse of each other. It will turn out to be sufficient to only have an explicit expression for $r_\sigma(z)$, computing $S_\sigma(\zeta)$ is not mandatory for the present purpose, even though such objects are at the heart of next chapter. Notice that in (4.5), we supposed $r_\sigma(z) \in [0,1]$, which can be verified when $z > 0$ using the explicit formula $r_\sigma(z)$ derived below.

To apply the tangent method we extend the domain as shown on Figure 4.5. Once again, the partition function of the extended domain $Z_n^{\text{ext}}$ can be expressed as a sum over the entry point $k$ of two independent partition functions. If we parametrize the weights as $t \equiv \frac{b}{a}$ and $\frac{c}{a} = \sqrt{1-2\Delta t + t^2}$, we get

$$\frac{Z_n^{\text{ext}}}{Z_{n,s}} = \sum_{k=1}^{n} \frac{1}{a^n} \frac{1}{t^{k-1}\sqrt{1-2\Delta t + t^2}} H_{n,s}^{(k)} Y_{k,\ell}, \tag{4.7}$$

where $Y_{k,\ell}$ is the partition function in the $\ell \times n$ extension that includes all vertices below the dashed line in Figure 4.5, with a prescribed exit point at column $k$. To evaluate $Y_{k,\ell}$ simply notice that in the extension $N_{c_2} = N_{c_1} + 1$. The enumeration for a fixed $m = N_{c_2}$ is straightforward, and we find

$$Y_{k,\ell} = \left(\frac{b}{a}\right)^{k+\ell-1} a^{n\ell} \sum_{m=0}^{k-1} \binom{k-1}{m} \binom{\ell-1}{m} \left(\frac{c}{b}\right)^{2m+1}. \tag{4.8}$$

Injecting into (4.7) we obtain

$$\frac{Z_n^{\text{ext}}}{Z_{n,s}} = a^{n(\ell-1)} t^{\ell-1} \sum_{k=1}^{n} H_{n,s}^{(k)} \sum_{m=0}^{k-1} \binom{k-1}{m} \binom{\ell-1}{m} \left(\frac{1-2\Delta t+t^2}{t^2}\right)^{m}. \tag{4.9}$$

The asymptotics is given by

$$\frac{Z_n^{\text{ext}}}{Z_{n,s}} \approx a^{n(\ell-1)} t^{\ell} \int_0^1 \mathrm{d}\kappa \int_0^\kappa \mathrm{d}\eta \, \mathrm{e}^{nS(\kappa,\eta,u)}, \tag{4.10}$$

where

$$\begin{aligned} S(\kappa,\eta,u) = S_\sigma(\kappa) + \mathcal{L}(\kappa) - \mathcal{L}(\eta) - \mathcal{L}(\kappa-\eta) + \mathcal{L}(u) - \mathcal{L}(\eta) \\ - \mathcal{L}(u-\eta) + \eta \log\left(\frac{t^2-2\Delta t+1}{t^2}\right), \end{aligned} \tag{4.11}$$

with $\mathcal{L}(x) \equiv x\log(x)$. To find the most likely entry point $\kappa^*$ in the rescaled domain, we perform a saddle point analysis. The saddle point equations are

$$S'_\sigma(\kappa^*)+\log(\kappa^*)-\log(\kappa^*-\eta^*) = 0, \tag{4.12}$$

$$-2\log(\eta^*)+\log(\kappa^*-\eta^*)+\log(u-\eta^*)+\log\left(\frac{t^2-2\Delta t+1}{t^2}\right) = 0. \tag{4.13}$$

Using (4.6), we deduce from (4.12) that

$$S'_\sigma(\kappa^*) = -\log\left(\frac{\kappa^*}{\kappa^*-\eta^*}\right) \quad \Rightarrow \quad z = \frac{\kappa^*}{\kappa^*-\eta^*}, \quad \kappa^*(z) = r_\sigma(z). \tag{4.14}$$

We now have all the ingredients—with the exception of a formula for $r_\sigma(z)$—to compute the family of tangent line. We should solve the implicit equation (4.14) to find $\eta^*(\kappa^*)$ and then inject this solution into (4.13) to find $\kappa^*(u)$. Instead of doing that, we choose to parametrize the family of tangent lines by $z$ instead of $u$. In those terms the family of lines is

$$F(x,y,z) = x - \frac{\kappa^*(u(z))}{u(z)}y - r_\sigma(z) = 0. \tag{4.15}$$

There remains to express $\frac{\kappa^*(u(z))}{u(z)}$ as an explicit function of $z$. Solving the quadratic equation (4.13) and using $z^{-1} = 1 - \frac{\eta^*}{\kappa^*}$, one gets after elementary manipulations

$$\frac{\kappa^*(u(z))}{u(z)} = \frac{z(t^2-2\Delta t+1)}{(z-1)(t^2 z-2\Delta t+1)}. \tag{4.16}$$

Hence, we conclude that

$$F(x,y,z) = x - \frac{z(t^2-2\Delta t+1)}{(z-1)(t^2 z-2\Delta t+1)}y - r_\sigma(z) = 0. \tag{4.17}$$

The remainder of this section is devoted to the computation of $r_\sigma(z)$, which heavily depends on the model at hand.

### 4.1.2 Almost homogeneous limit

We consider the vertically *inhomogeneous* 6V model with pDWBC where vertices on line $j$ (numbered from 1 to $s$ from top to bottom) are assigned a weight

$$a_j = 1, \quad b_j = t_j, \quad c_j = \sqrt{1-2\Delta t_j + t_j^2}, \tag{4.18}$$

in such a way that $\Delta_j = \Delta$. To retrieve the original partition function $Z_{n,s}$, it suffices to evaluate the *homogeneous* limit $t_j \to t$ of the inhomogeneous partition function $Z_{n,s}(t_1,\cdots,t_s)$. The inhomogeneous partition function carries more information since an *almost homogeneous* limit (where an inhomogeneity is kept in the last line) can be used to make contact with the one-point function

$$\frac{Z_{n,s}(t,\cdots,t,t_s)}{Z_{n,s}} = \frac{c_s}{c} \sum_{k=1}^{n} H_{n,s}^{(k)} \left(\frac{b_s}{b}\right)^{k-1} = \frac{c_s}{c}\, h_{n,s}(t_s/t). \tag{4.19}$$

From now on we work in the free-fermion case $\Delta = 0$. The authors of [145] have shown that

$$Z_{n,s}(t_1,\cdots,t_s) = \prod_{i=1}^{s} \frac{c_i}{1-t_i} \prod_{1\le i<j\le s} \frac{1+t_i t_j}{(1-t_i t_j)(t_j-t_i)} \det_{1\le i,j\le s} \left[t_j^{i-1} - t_j^{\,n+s-i}\right]. \tag{4.20}$$

We first evaluate the homogeneous limit where $t_j \to t = 1$. For the DWBC, the analogue of (4.20) is the Izergin-Korepin determinantal formula, for which the homogeneous limit can be taken stepwise, by evaluating the limit over one inhomogeneity at a time. In this case, zeros coming from Taylor expanding the determinant are exactly cancelled by poles contained in the prefactor [139]. What remains in the determinant are derivatives of some function[3]. In the case at hand, the homogeneous limit of formula (4.20) can be evaluated in a similar way, but only poles of odd order appear in the prefactor. The result of the homogeneous limit is

$$Z_{n,s}(1,\cdots,1) = -c^s 2^{s(s-1)/2} \left(\prod_{k=1}^{s} \frac{1}{(2k-1)!}\right) \det_{1\le i,j\le s}(A_{ij}), \tag{4.21}$$

were the entries

$$A_{ij} \equiv f_i^{(2j-1)}(1) = \frac{(i-1)!}{(i-1-(2j-1))!} - \frac{(n+s-i)!}{(n+s-i-(2j-1))!}, \tag{4.22}$$

are the derivatives of $f_i(t) \equiv t^{i-1} - t^{n+s-i}$. We take the (standard) convention that if the factorial of a negative number appears in a denominator, the corresponding term is set to zero. A rigorous derivation

[3] Another way of evaluating this limit, given in [99], is at the core of the proof of (2.44). In this approach, coefficients of generating functions naturally appear in the determinant.

of (4.21) is presented in Appendix B. The almost homogeneous limit is taken in a similar fashion and yields

$$Z_{n,s}(1,\cdots,1,t_s)=c^{s-1}c_s 2^{\frac{(s-1)(s-2)}{2}}\frac{(1+t_s)^{s-1}}{(1-t_s)^{2s-1}}\left(\prod_{k=1}^{s-1}\frac{1}{(2k-1)!}\right)\det_{1\le i,j\le s}(\tilde{A}_{ij}), \tag{4.23}$$

where the entries $\tilde{A}_{ij}$ only differ from $A_{ij}$ in the last column where $\tilde{A}_{is}=t_s^{i-1}-t_s{}^{n+s-i}$. We define $t_s=1+\xi$ and permute the lines of $A$ and $\tilde{A}$, therefore working with

$$B_{i,j}=A_{s+1-i,j} \quad , \quad \tilde{B}_{i,j}=\tilde{A}_{s+1-i,j}. \tag{4.24}$$

Putting (4.21) and (4.23) into (4.19) one gets

$$h_{n,s}(1+\xi)=2^{1-s}(2s-1)!\left(\frac{1}{\xi}\right)\left(\frac{2+\xi}{\xi^2}\right)^{s-1}\frac{\det\tilde{B}}{\det B}. \tag{4.25}$$

We now have to evaluate the asymptotics of a ratio of determinants that only differ in their last column. Such situations are typical in the application of the tangent method. In [15, 20], the asymptotics of a ratio of determinants is evaluated directly by using random matrix model techniques [98]. In the case at hand however, one can find the LU decomposition of $B$ and use techniques of [22] to find a finite size expression for $\frac{\det\tilde{B}}{\det B}$, whose asymptotics is then straightforward to evaluate.

### 4.1.3 LU decomposition

The general strategy is that if we can find a strictly lower triangular matrix $L$ and an upper triangular matrix $U$ such that $B=LU$, then since $\tilde{B}$ differs from $B$ only by its last column we have $\tilde{B}=L\tilde{U}$, with the same matrix $L$. The matrix $\tilde{U}$ differs from $U$ by its last column and can be computed from $\tilde{B}$ using $L^{-1}\tilde{B}=\tilde{U}$. Since the determinants $\det(B)$ and $\det(\tilde{B})$ are respectively $\prod_{i=1}^{s}U_{ii}$ and $\prod_{i=1}^{s}\tilde{U}_{ii}$, we obtain

$$\frac{\det\tilde{B}}{\det B}=\frac{\tilde{U}_{ss}}{U_{ss}}, \tag{4.26}$$

with

$$\tilde{U}_{ss}=\sum_{p=1}^{s}L^{-1}_{sp}\tilde{B}_{ps}=\sum_{p=1}^{s}L^{-1}_{sp}\left[(1+\xi)^{s-p}-(1+\xi)^{n+p-1}\right]. \tag{4.27}$$

There only remains to find the LU decomposition of $B$.

**Theorem 4.1.1.** *The matrices $L^{-1}$ and $U$ that appear in the LU decomposition of $B$ have matrix elements*

$$\begin{aligned} L^{-1}_{ij} &= (-1)^{i+j}\binom{i-1}{j-1}\prod_{l=j}^{i-1}\frac{n-s+i+l}{n-s+l} \quad \text{for } i \geq j, \\ U_{ii} &= -(i-1)!\prod_{l=i}^{2i-1}(n+l-s), \end{aligned} \qquad (4.28)$$

*in addition to being respectively lower triangular and upper triangular.*

The formulae (4.28) have been first guessed for small values of $n$ and $s$ by an explicit calculation of the LU decomposition, using a prime factor decomposition; they were subsequently proved to hold in general (the proof is in Appendix B). Using (4.27) we now have a (very explicit) finite size formula for $h_{n,s}(1+\xi)$. It takes the form of an alternating sum, which is quite typical in such problems [22]. Indeed, the alternating sign can ultimately be traced back to cofactors in the evaluation of $L^{-1}$. The asymptotics of alternating sums is not easy to evaluate. In [22] it is shown how to deal with this issue by massaging the alternating sum, with the help of generating functions, in a form that is suited for a straightforward asymptotic analysis. The technical details of this computation are again gathered in Appendix B. One finds that

$$\tilde{U}_{ss} = -\binom{n-1}{s-1}^{-1}\sum_{k=2s-1}^{n+s-1}\binom{n+s-1}{k}\binom{k-s}{s-1}\xi^k, \qquad (4.29)$$

so that (4.25) and (4.26) yield

$$h_{n,s}(1+\xi) = C\left(\frac{1}{\xi}\right)\left(\frac{2+\xi}{\xi^2}\right)^{s-1}\sum_{k=2s-1}^{n+s-1}\binom{n+s-1}{k}\binom{k-s}{s-1}\xi^k, \qquad (4.30)$$

with

$$C = \frac{2^{1-s}(2s-1)!\binom{n-1}{s-1}^{-1}}{(s-1)!\prod_{l=s}^{2s-1}(n+l-s)}. \qquad (4.31)$$

The precise value of $C$ does not matter, since it is independent on $\xi$ and disappears in the logarithmic derivative in (4.4). A similar observation shows that the only important input of the LU decomposition is actually $L^{-1}_{sj}$.

### 4.1.4 Asymptotics and arctic curve

There only remains to compute the asymptotics of (4.30). We must analyse the sum

$$\sum_{k=2s-1}^{n+s-1} \binom{n+s-1}{k} \binom{k-s}{s-1} \xi^k \approx \int_{2\sigma}^{1+\sigma} \mathrm{d}u \, \mathrm{e}^{nR(\sigma,u,\xi)}, \tag{4.32}$$

where

$$R(\sigma,u,\xi) = \mathcal{L}(1+\sigma) - \mathcal{L}(u) - \mathcal{L}(1+\sigma-u) + \mathcal{L}(u-\sigma) - \mathcal{L}(u-2\sigma) - \mathcal{L}(\sigma) + u \log \xi. \tag{4.33}$$

The saddle point equation

$$\frac{\partial R}{\partial u} = \log\left(\frac{\xi(1+\sigma-u)(\sigma-u)}{u(2\sigma-u)}\right) = 0. \tag{4.34}$$

yields two solutions

$$u_\pm(\sigma,\xi) = \frac{\xi + 2\sigma(1+\xi) \pm \sqrt{\xi^2 + 4\sigma^2(1+\xi)}}{2(1+\xi)}. \tag{4.35}$$

We must have $u \in [2\sigma, 1+\sigma]$. We notice that $\xi \in [0,+\infty[$, since $z = 1+\xi = \frac{\kappa^*}{\kappa^*-\eta^*}$ and $0 < \eta^* < \kappa^*$. Hence, we must exclude the case $u_-(\sigma,\xi) < \sigma$. We also check that $u_+(\sigma,\xi)$ spans the whole range $[2\sigma, 1+\sigma]$ when $\xi \in [0,\infty[$.

Using (4.33) and $\frac{\partial R}{\partial u}\big|_{u_+} = 0$ we obtain $\frac{\partial}{\partial \xi} R(\sigma,u_+,\xi) = \frac{u_+}{\xi}$. We hence have

$$r_\sigma(1+\xi) = (1+\xi)\left[\frac{-\sigma(4+\xi)}{\xi(2+\xi)} + \frac{u_+}{\xi}\right], \tag{4.36}$$

which upon substitution in (4.17) specialized at $t=1$ and $\Delta = 0$ finally leads to the equation

$$\begin{aligned} &F(x,y,\xi) \\ &= x - \frac{2(1+\xi)}{\xi(2+\xi)} y - (1+\xi)\left[\frac{-\sigma(4+\xi)}{\xi(2+\xi)} + \frac{\xi + 2\sigma(1+\xi) + \sqrt{\xi^2 + 4\sigma^2(1+\xi)}}{2\xi(1+\xi)}\right]. \end{aligned} \tag{4.37}$$

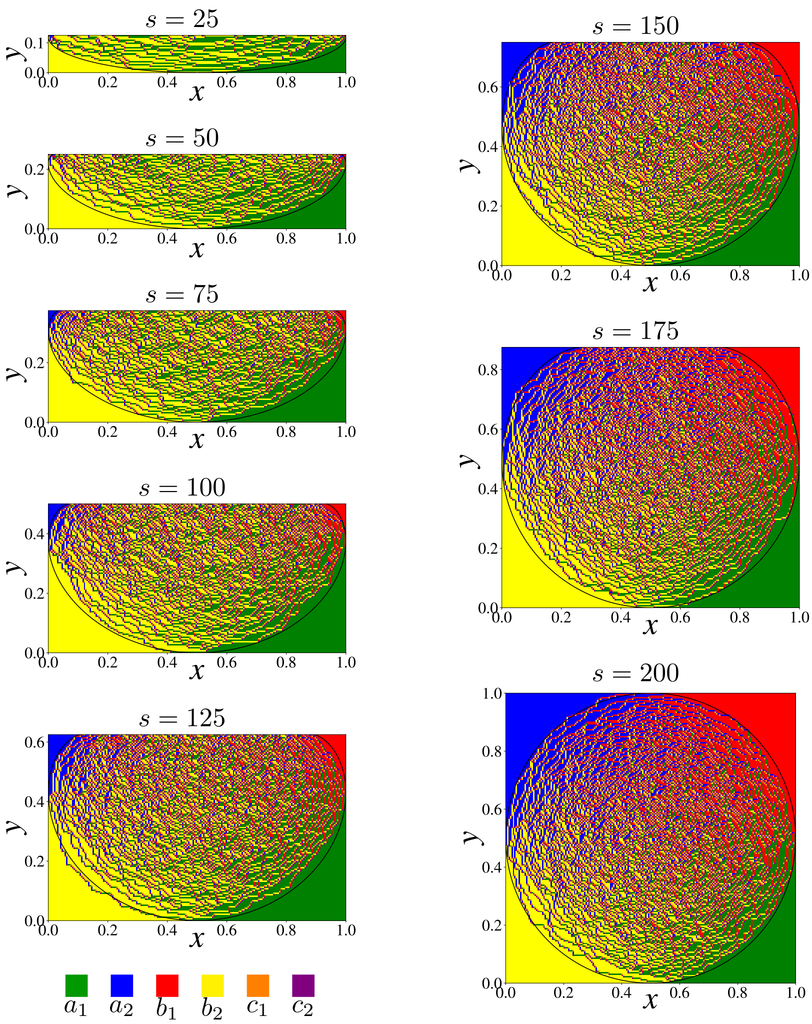



Figure 4.6: Configurations of the 6V model with pDWBC for $a=b=1$ and $c=\sqrt{2}$ ($\Delta=0$ and $t=1$) for $n=200$ and various values of $s$. Configurations were generated by a Metropolis algorithm (similar to that of [105]) that was run for $10000\times n\times s$ steps. The black curve is drawn from (4.38).

The arctic curve is the envelope and satisfies $F(x,y,\xi)=0$ and $\frac{\partial}{\partial\xi}F(x,y,\xi)=0$. Solving these equations leads to a parametrization of the arctic curve

$$
\begin{aligned}
x(\xi) &= \frac{2\sigma^2\xi(\xi+1)+(2+\xi(2+\xi))(\xi+\sqrt{\xi^2+4\sigma^2(1+\xi)})}{2(2+\xi(2+\xi))\sqrt{\xi^2+4\sigma^2(1+\xi)}},\\
y(\xi) &= \sigma\frac{-\sigma(2+\xi)^3+2(2+\xi(2+\xi))\sqrt{\xi^2+4\sigma^2(1+\xi)}}{2(2+\xi(2+\xi))\sqrt{\xi^2+4\sigma^2(1+\xi)}},
\end{aligned}
\tag{4.38}
$$

where $\xi \in [0, \infty[$. The curve (4.38) can be shown to be a degree six algebraic curve, whose exact expression is not very enlightening. Our computation is valid for the south-east portion of arctic curve. The south-west portion is deduced by symmetry. Since we are at the free-fermion point $\Delta = 0$ we expect that the analytic continuation of the arctic curve also describes the two remaining portions. Figure 4.6 shows that this conjecture is well supported by numerical simulations. When $\sigma = 1$ an arctic circle is recovered, which is expected since the boundary conditions then reduce to DWBC [140].

## 4.2 Twenty-vertex model with DWBC2

Amongst exactly solvable models, free-fermionic ones enjoy a particularly rich structure which enables the exact characterization of very subtle properties, such as the fluctuations studied in Chapter 3. Interacting models are notoriously harder to deal with, especially for generic integrable weights, and are at the heart of a very active field of research (see e.g. [146–150]). Derivations of arctic curves in such models are relatively rare [16, 20, 30, 151, 152]. The first *tour de force* was the computation of the arctic curve of the 6V model with DWBC in [14, 15, 20], which has in general *non-analyticities* at the point of contact with the boundary. The easiest version of the computation [20, 30] relies on the tangent method, the most technical part being the evaluation of the asymptotics of an almost homogeneous Izergin-Korepin determinant, which is most simply done in the D phase where a standard determinant identity boils it down to a differential equation that can be explicitly solved [14, 30]. In this section, we follow an idea of Baxter that allows to generalize results of the 6V model to an *ice-type model on the regular triangular lattice* whose *weights are restricted* in a very specific way [54, 153]. The idea is to exploit a correspondence between ice models on the triangular and on the Kagome lattice and then use the celebrated *Yang-Baxter equations* to *expel the extra lines* of the Kagome lattice so as to recover a 6V model, a procedure called *unravelling*.

Each edge of the triangular lattice is assigned an orientation on which a generalized version of the ice rule is imposed, namely *each vertex is incident to as many incoming as outgoing arrows*. This leads to the $20 = \binom{6}{3}$

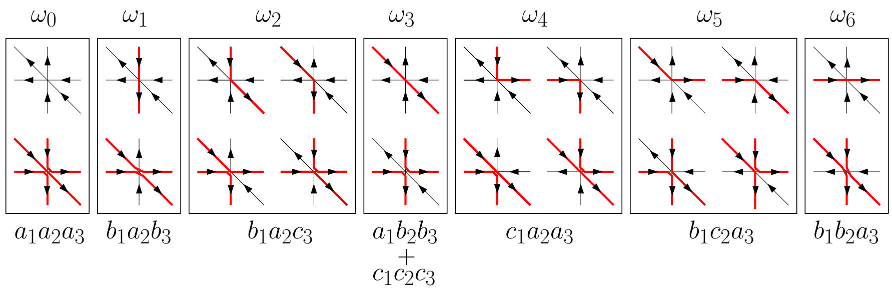



Figure 4.7: The twenty possible vertices, represented in the arrow language and in terms of osculating paths. South, south-east and east arrows are assigned a portion of path. The first row is related to the second by arrow reversal. We restrict our attention to very specific weights shown on the picture.

possible environments around a given vertex displayed in Figure 4.7. An alternative name of this model—that we shall now use—is hence the *twenty-vertex* (twenty-vertex (20V)) model. For convenience, we represent the triangular lattice as a square lattice supplemented with a second diagonal within each face. A configuration can equivalently be described by osculating paths—which can now have triple kissing points—using the bijection presented on Figure 4.7. In this section, we constantly oscillate between these two descriptions, as the unravelling is best described in the arrow picture, while the tangent method is most conveniently expressed in the osculating paths picture. We restrict the (rectified) triangular lattice to a square and fix the boundary conditions in a way that is reminiscent to DWBC. In [59], four natural generalizations were considered and named DWBC1,2,3,4. The first two are closely related (see Figure 4.8) and an arctic phenomenon only occurs in the first three, see Figure 4.9.

The different nature of the 20V osculating paths slightly modifies the heuristics of the arctic phenomenon. For DWBC3, paths start on the left boundary and end on the lower boundary, with diagonal edges on these boundaries remaining empty. This is quite similar to DWBC, up to an irrelevant change of path description. Quite expectedly, empty (E), all vertical (V) and all horizontal (H) regions also form for DWBC3. In the 20V case however, paths can have diagonal steps and enjoy more freedom than their 6V counterpart, resulting in the disappearance of the maximal density region that would have been in the lower left corner

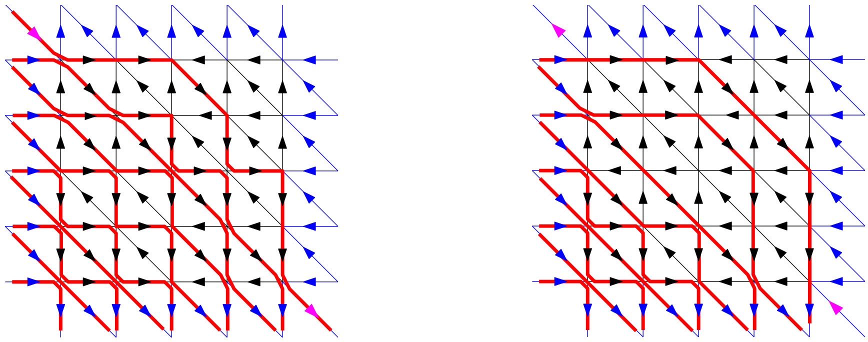


Figure 4.8: Configurations of the 20V model, in their arrow and osculating paths description. The blue arrows are fixed. *Left*: DWBC1. *Right*: DWBC2. These two boundary conditions differ by the magenta arrows.

otherwise. This extra freedom also explains the modest size of the V and H regions, which are barely visible in Figure 4.9-right. The heuristics becomes slightly more subtle when the left and bottom boundary become saturated with starting and ending points, i.e. when diagonal steps are also occupied. This is exactly the definition of DWBC1. In this case, we recover the expected E, H, V and also all horizontal, diagonal and vertical (HDV) regions, but two *new frozen regions with double kissing points* appear near the north-west and south-east corners. These regions are respectively the all horizontal and diagonal (HD) and all diagonal and vertical (DV) regions. These HD and DV regions are respectively separated from the adjacent H and V regions by an interface along the second diagonal. This peculiar behaviour is explained by the maximal density of starting points and the non-crossing constraint combined with the natural tendency of external paths to initially go as far as possible from other paths. Interestingly enough, the arctic curve typically has two cusps.

In this section, we mainly focus on the 20V model with DWBC2 defined on Figure 4.8 whose combinatorics is mysteriously related to that of QTHADT. The only difference between DWBC1 and DWBC2 is the orientations of the upper left and lower right diagonal edges that are opposite (see Figure 4.8). In the osculating paths language, DWBC1 have a (single) additional path which, according to the working hypothesis of the tangent method (number (i) in Subsection 2.1.2), does not influence the other paths in the scaling limit and should hence be irrelevant for

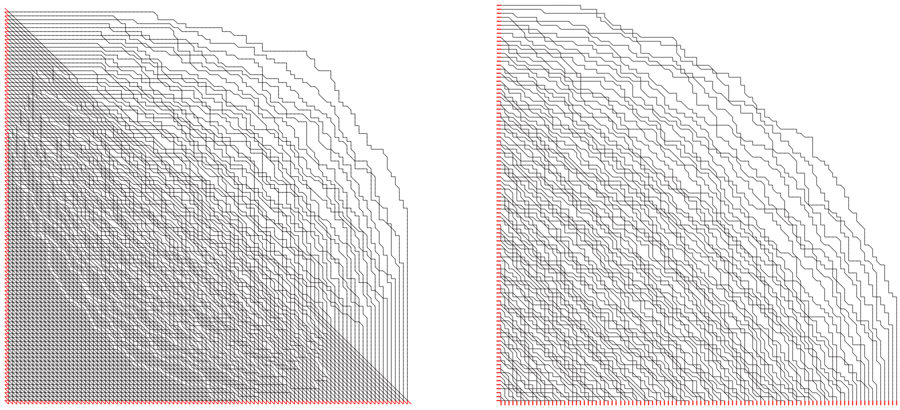


Figure 4.9: Two configurations of the 20V model with uniform weights, with DWBC1 (left) and DWBC3 (right). The portions of path imposed by the boundary conditions are drawn in red.

the arctic phenomenon. Configurations of the 20V model with DWBC1 are in one-to-one correspondence with those with DWBC2 by a simple 180° rotation, in the arrow language. In the osculating paths language, it amounts to performing both a 180° rotation and a complementation in which occupied and unoccupied edges are interchanged. For instance, the configurations depicted in Figure 4.8 are image of each other under this rotation.

### 4.2.1 Restriction on the weights

There are *a priori* twenty possible weights for the vertices. There are only seventeen independent parameters, since the fixed boundary imposes—as in the DWBC case—three conservation laws. However, the unravelling procedure requires the weights to be chosen in a certain way, shown in Figure 4.7. In addition to those restrictions, we further constrain the weights to make the computation as light as possible. The most general (to our knowledge) application of the tangent method to the 20V model with DWBC2 is very similar [17].

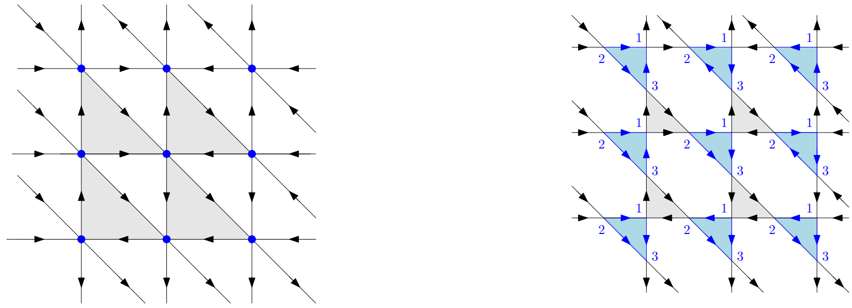


Figure 4.10: Transformation of a piece of triangular lattice into a piece of Kagome lattice. The Kagome lattice is naturally divided into three sub-lattices 1, 2 and 3 as indicated.

**Kagome lattice:** As noticed in [54], for appropriate vertex weights, the 20V model may be reformulated as an ice model on the Kagome[4] lattice obtained by slightly shifting upward each horizontal line. Some of the triangles (in light grey in Figure 4.10) of the original lattice are preserved during this procedure. Their incident edges keep their orientation. New triangles (in blue in Figure 4.10-right) appear. One can always find an orientation of their edges such that the ice rule is satisfied at each vertex of the Kagome lattice, as seen by a direct enumeration. Moreover, the maximum number of such orientations is two per newly formed triangle. When two orientations are possible, the choice is made independently for each of the newly formed triangle. Each 20V model configuration is thus in correspondence with a number of configurations on the Kagome lattice. This lattice is naturally split into three sub-lattices numbered 1, 2 and 3 (see Figure 4.10-right). The ice rule leads to six possibilities for each of the three types of tetravalent vertices, which are assigned weights that are *invariant under (global) arrow reversal*, see Figure 4.11-left. If we now go back and shift horizontal lines downward, the blue triangles of Figure 4.10-right can be identified with vertices of the triangular lattice as they become infinitesimally small. If we sum over the ambiguous orientations on blue triangles, we obtain a 20V model having now ten possible weights: $a_1a_2a_3$, $b_1a_2b_3$, $b_1a_2c_3$,

[4]Kagome is the Japanese name for this pattern, which sometimes appears on traditional woven bamboo baskets.

$c_1a_2a_3$, $b_1c_2a_3$, $b_1b_2a_3$, $a_1b_2c_3 + c_1c_2b_3$, $a_1b_2b_3 + c_1c_2c_3$, $c_1b_2b_3 + a_1c_2c_3$ and $c_1b_2c_3 + a_1c_2b_3$.

**Yang-Baxter equation:** Crucially, we demand that the Yang-Baxter equation be satisfied, namely *the same weights* are obtained for the 20V model if, in our equivalence, we shift the horizontal lines of the triangular lattice down instead of up. This leads to the identity depicted in Figure 4.11-right, and imposes the extra conditions

$$\begin{aligned} a_1b_2c_3 + c_1c_2b_3 =& b_1a_2c_3, \quad c_1b_2b_3 + a_1c_2c_3 = c_1a_2a_3, \\ & c_1b_2c_3 + a_1c_2b_3 = b_1c_2a_3. \end{aligned} \tag{4.39}$$

These conditions are crucial as they will allow us to unravel the 20V model configurations with DWBC2 into configurations of a single 6V model with standard DWBC (corresponding to the Kagome sub-lattice labelled by 1, see below). The set of the 20V model weights is then reduced to a list of seven values, with the dictionary depicted in Figure 4.7.

This construction leads to weights that are *invariant under* 180° *rotation* (in the arrow language) around any vertex, see Figure 4.7. Since this transformation interchanges DWBC1 and DWBC2, the partition functions of those two models are the same. Another consequence is that the arctic curve of the 20V model with DWBC1 is the image under 180° rotation of the arctic curve of the 20V model with DWBC2. Moreover, those boundary conditions only differ by the presence of an additional path in DWBC1 (see Figure 4.8), which is irrelevant for the arctic curve. Hence, since vertex weights are invariant under 180° rotation, we deduce that *the arctic curve itself is symmetric under* 180° *rotation*.

**Integrable weight parametrization:** We parametrize the weights of the Kagome lattice in such a way that (4.39) is satisfied, as done in [59]. Introducing the notations

$$A(z,w)=z-w, \quad B(z,w)=q^{-2}z-q^2w, \quad C(z,w)=(q^2-q^{-2})\sqrt{zw}, \tag{4.40}$$

we set

$$\begin{aligned} &a_1=\nu A(z,w), & &b_1=\nu B(z,w), & &c_1=\nu C(z,w), \\ &a_2=\nu A(qz,q^{-1}t), & &b_2=\nu B(qz,q^{-1}t)\ , & &c_2=\nu C(qz,q^{-1}t), \\ &a_3=\nu A(qt,q^{-1}w), & &b_3=\nu B(qt,q^{-1}w), & &c_3=\nu C(q\,t,q^{-1}w), \end{aligned} \tag{4.41}$$

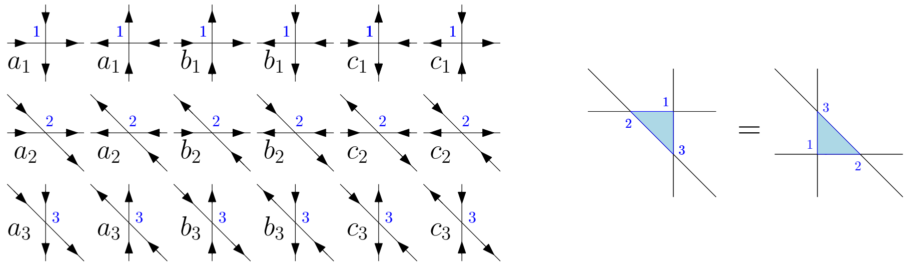



Figure 4.11: *Left*: Weights of the vertices on the three sub-lattices 1, 2 and 3 of the Kagome lattice. *Right*: Schematic representation of the Yang-Baxter equation. The ice-rule is imposed at each vertex. The equality must hold for any valid arrow configuration on the external edges, upon summation over the internal edge orientation.

with some arbitrary $q \in \mathbb{C}^*$, and arbitrary spectral parameters $z$, $w$ and $t$, respectively attached to the horizontal, vertical and diagonal direction. The way to solve (4.39) in order to obtain a solution like (4.41) is explained in [54]. One can verify explicitly that (4.41) satisfies (4.39). The normalization of the weights on each sub-lattices can be chosen independently. It does not affect the statistics of the model, since the number of vertices of each sub-lattices is fixed. It can hence be chosen arbitrarily. We take the same normalization $\nu$ for each sub-lattices and set $\nu = 1/(2\mathrm{i}t^{1/3})$ for future convenience.

The above homogeneous weights are members of a more general family of inhomogeneous (i.e. node-dependent) integrable weights

$$\begin{array}{lll}
\frac{a_1(i,j)}{\nu} = A(z_i, w_j), & \frac{b_1(i,j)}{\nu} = B(z_i, w_j), & \frac{c_1(i,j)}{\nu} = C(z_i, w_j), \\
\frac{a_2(i,k)}{\nu} = A(q\, z_i, \frac{t_k}{q}), & \frac{b_2(i,k)}{\nu} = B(q\, z_i, \frac{t_k}{q}), & \frac{c_2(i,k)}{\nu} = C(q\, z_i, \frac{t_k}{q}), \\
\frac{a_3(k,j)}{\nu} = A(q\, t_k, \frac{w_j}{q}), & \frac{b_3(k,j)}{\nu} = B(q\, t_k, \frac{w_j}{q}), & \frac{c_3(k,j)}{\nu} = C(q\, t_k, \frac{w_j}{q}),
\end{array} \tag{4.42}$$

where a different spectral parameter $z_i$, $w_j$ and $t_k$ is attached to each horizontal, vertical and diagonal line respectively (the normalization $\nu$ is again arbitrary). Here the horizontal lines are numbered by $i = 1, \ldots n$ from bottom to top, the vertical lines by $j = 1, 2, \ldots, n$ from left to right, and the diagonal lines by $k = 1, 2, \ldots, 2n-1$ from the lower left to the upper right corner. In this setting, the pair $(i, j)$, (resp. $(i, k)$ and $(j, k)$) therefore refers to the node of sub-lattice 1 (resp. 2 and 3) at the crossing of the $i$-th horizontal and $j$-th vertical lines (respectively of the $i$-th horizontal/$k$-th diagonal and of the $j$-th vertical/$k$-th diagonal lines). A crucial property of the above integrable weights is that they *satisfy the Yang-Baxter equation at each node for arbitrary spectral parameters* $z_i$, $w_j$ and $t_k$. This property is used in Appendix E.

**Disordered regime:** Returning to the situation with homogeneous spectral parameters $z$, $w$ and $t$, we choose[5]

$$q = \mathrm{e}^{\mathrm{i}\,\eta}, \quad z = \mathrm{e}^{\mathrm{i}\,(\eta+\lambda)}, \quad w = \mathrm{e}^{-\mathrm{i}\,(\eta+\lambda)}, \quad t = \mathrm{e}^{\mathrm{i}\,\mu}, \qquad (\eta, \lambda, \mu \in \mathbb{C}). \tag{4.43}$$

After all these restrictions only three independent parameter remain and the seven weights of Figure 4.7 are given by

$$\begin{aligned}
\omega_0 &= \sin(\lambda+\eta)\sin\Big(\frac{\lambda+3\eta+\mu}{2}\Big)\sin\Big(\frac{\lambda+3\eta-\mu}{2}\Big),\\
\omega_1 &= \sin(\lambda-\eta)\sin\Big(\frac{\lambda-\eta+\mu}{2}\Big)\sin\Big(\frac{\lambda+3\eta-\mu}{2}\Big),\\
\omega_2 &= \sin(2\eta)\sin(\lambda-\eta)\sin\Big(\frac{\lambda+3\eta-\mu}{2}\Big),\\
\omega_3 &= \sin(2\eta)^3+\sin(\lambda+\eta)\sin\Big(\frac{\lambda-\eta+\mu}{2}\Big)\sin\Big(\frac{\lambda-\eta-\mu}{2}\Big),\\
\omega_4 &= \sin(2\eta)\sin\Big(\frac{\lambda+3\eta+\mu}{2}\Big)\sin\Big(\frac{\lambda+3\eta-\mu}{2}\Big),\\
\omega_5 &= \sin(2\eta)\sin(\lambda-\eta)\sin\Big(\frac{\lambda+3\eta+\mu}{2}\Big),\\
\omega_6 &= \sin(\lambda-\eta)\sin\Big(\frac{\lambda+3\eta+\mu}{2}\Big)\sin\Big(\frac{\lambda-\eta-\mu}{2}\Big).
\end{aligned} \tag{4.44}$$

The weights considered in [153] are equivalent as they are merely another parametrization[6]. We finally restrict our choice to *real* values of the parameters $\eta$, $\lambda$ and $\mu$, which implies in particular that the three Kagome 6V models are in the D phase [54]. The range of these parameters is taken so as to ensure that all $\omega$'s are positive, namely

$$0 < \eta < \lambda < \pi - \eta, \quad \eta - \lambda < \mu < \lambda - \eta\ . \tag{4.45}$$

Note as a consequence that $\eta < \frac{\pi}{2}$ and $\lambda + 3\eta > \mu$.

[5]This choice corresponds to imposing $z\,w = 1$, which may be done without loss of generality since only the ratios of weights matter for the statistics of the model.

[6]With the notations of [153], the correspondence is $\lambda = \zeta + \pi - \theta + (\phi - \pi)/2$, $\mu = \phi - \zeta - 3\theta$ and $\eta = (\pi - \phi)/2$.

The 6V model on the Kagome sub-lattice labelled 1 has weights $(a_1, b_1, c_1) = (a, b, c)/t^{1/3}$, where we recognize the standard 6V model weight parametrization in the disordered phase

$$a = \sin(\lambda + \eta), \quad b = \sin(\lambda - \eta), \quad c = \sin(2\eta). \tag{4.46}$$

### 4.2.2 Unravelling

As explained in [59] and depicted in Figure 4.12, the Yang-Baxter equation allows to unravel the configurations of the 20V model with DWBC2 with weights (4.41), in its Kagome description. At the end of this procedure, the only remaining non-trivial part of a configuration is a single 6V model on sub-lattice 1. As a consequence, the partition function of the 20V model with DWBC2 (denoted $\mathcal{Z}_n$) with the weights (4.44) and that for the 6V model with weights $(a, b, c)$ of (4.46) above (denoted $Z_n = Z_n[a, b, c]$) are then directly proportional, namely:

**Theorem 4.2.1.** *(Di Francesco-Guitter [59]) The partition function of the 20V model with DWBC2 and of the 6V model with DWBC are related by*

$$\mathcal{Z}_n = \left(\frac{a_2 a_3}{t^{1/3}}\right)^{n^2} Z_n = \left(\sin\left(\frac{\lambda+3\eta-\mu}{2}\right)\sin\left(\frac{\lambda+3\eta+\mu}{2}\right)\right)^{n^2} Z_n, \tag{4.47}$$

*where their respective weights are fixed to* (4.44) *and* (4.46)*.*

In the rest of this section, quantities written with a calligraphic font refer to the 20V model with DWBC2, and their non-calligraphic counterpart to the 6V model with DWBC. Their respective weight dependence are sometimes made implicit, in which case they are given by (4.44) and (4.46), possibly particularized.

**Refined enumeration.** The use of the tangent method requires a refined enumeration of the 20V model configurations where we keep track of the position where the uppermost path hits the right boundary, i.e. the vertical line $j = n$ (see Figure 4.8-right or 4.13-left). As explained in Appendix E, this enumeration may be obtained by changing $w \to w\,\theta$ for the spectral parameter of the last column ($j = n$). Note that the

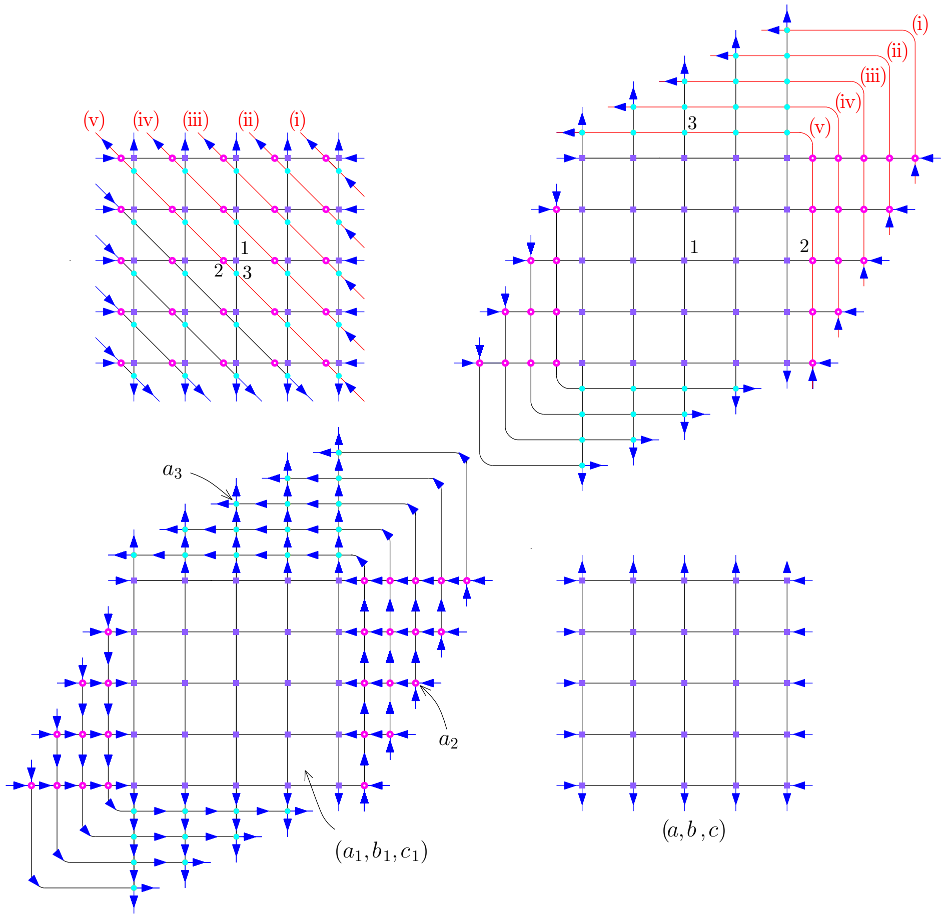



Figure 4.12: Unravelling procedure. *Upper left*: We consider the 20V model with DWBC2 in its Kagome description. *Upper right*: Diagonal lines are expelled outward from the central square grid using the Yang-Baxter equation. The red lines are moved to the north-east and then bent, starting with the outer one (i) and ending with the diagonal one (v), all that while retaining their order. During the translation triangles of the Kagome lattice are inverted, so one has to be careful in keeping track of the sublattice types. The other diagonals are expelled to the south-west in a similar fashion. This has the effect of isolating the vertices according their type. Vertices of sub-lattice 1 stay in the central square region while vertices of sub-lattices 2 and 3 are expelled, respectively to the left and right, and to the top and bottom. *Lower left*: The DWBC2 prescription and the ice rule fix the orientations of the expelled edges so that all the nodes of sub-lattice 2 (resp. 3) receive a weight $a_2$ (resp. $a_3$), and hence only contribute to the partition function by a global multiplicative factor $(a_2a_3)^{n^2}$. *Lower right*: The remaining central configuration region is a 6V model with DWBC and weights $(a_1, b_1, c_1)$ or equivalently $(a, b, c)$ of (4.46) up to a global factor $(1/t^{1/3})^{n^2}$.

uppermost path corresponds in this case to the path starting at the horizontal edge $(0, n) \to (1, n)$ and ending at the vertical edge $(n, 1) \to (n, 0)$. We introduce the refined partition function

$$\mathcal{Z}_n(\tau) = \sum_{L=1}^{n} \mathcal{Z}_{n,L}\, \tau^{L-1}, \tag{4.48}$$

where $\mathcal{Z}_{n,L}$ denotes the partition function of the 20V model configurations with DWBC2 for which the uppermost path hits the vertical line $j = n$ at position $(n, L)$. Note that the step just before the hitting point is either horizontal or diagonal. We call $\mathcal{Z}^{-}_{n,L}$ and $\mathcal{Z}^{\diagdown}_{n,L}$ the corresponding restricted refined partition functions, and introduce

$$\mathcal{Z}^{-}_n(\tau) = \sum_{L=1}^{n} \mathcal{Z}^{-}_{n,L}\, \tau^{L-1}, \qquad \mathcal{Z}^{\diagdown}_n(\tau) = \sum_{L=1}^{n} \mathcal{Z}^{\diagdown}_{n,L}\, \tau^{L-1}, \tag{4.49}$$

satisfying the obvious sum rule

$$\mathcal{Z}_n(\tau) = \mathcal{Z}^{-}_n(\tau) + \mathcal{Z}^{\diagdown}_n(\tau). \tag{4.50}$$

As for the 6V model with DWBC configurations, we consider a similar refinement as follows: as is well-known, configurations of the 6V model with DWBC are in bijection with configurations of osculating paths. Those are obtained as for the 20V model by drawing path steps along the edges oriented east or south and by connecting them uniquely into non-crossing but possibly kissing well-oriented paths (going from the west boundary to the south boundary of the square grid). We then denote by $Z_{n,L}$ the partition function for those configurations where the uppermost path (starting at the horizontal edge $(0, n) \to (1, n)$ and ending at the vertical edge $(n, 1) \to (n, 0)$) hits the line $j = n$ at position $(n, L)$. Note that this hitting point is necessarily preceded by a horizontal step. We finally set

$$Z_n(\sigma) = \sum_{L=1}^{n} Z_{n,L}\, \sigma^{L-1}. \tag{4.51}$$

Now that refined partition functions are properly defined, we may generalize Theorem 4.2.1 to this case:

**Theorem 4.2.2.** *The refined partition functions $\mathcal{Z}^{-}_n(\tau)$ and $\mathcal{Z}^{\diagdown}_n(\tau)$ of the 20V model with DWBC2 are related to the refined partition function*

$Z_n(\sigma)$ *of the 6V model with DWBC via*

$$\mathcal{Z}_n^{-}(\tau) + g(\sigma)\, \mathcal{Z}_n^{\diagdown}(\tau) = \left(\frac{a_2 a_3}{t^{1/3}}\right)^{n^2} Z_n(\sigma), \tag{4.52}$$

*where*

$$\begin{aligned} \tau &= \sigma \frac{\sigma \sin(\lambda-\eta) \sin\left(\frac{\lambda+3\eta-\mu}{2}\right) - \sin(\lambda+\eta) \sin\left(\frac{\lambda-\eta-\mu}{2}\right)}{\sigma \sin(\lambda-\eta) \sin\left(\frac{\lambda-\eta-\mu}{2}\right) - \sin(\lambda+\eta) \sin\left(\frac{\lambda-5\eta-\mu}{2}\right)} \; \frac{\sin\left(\frac{\lambda+3\eta+\mu}{2}\right)}{\sin\left(\frac{\lambda-\eta+\mu}{2}\right)}, \\ g(\sigma) &= \frac{\sigma \sin(2\eta) \sin\left(\frac{\lambda+3\eta+\mu}{2}\right)}{\sigma \sin(\lambda-\eta) \sin\left(\frac{\lambda-\eta-\mu}{2}\right) - \sin(\lambda+\eta) \sin\left(\frac{\lambda-5\eta-\mu}{2}\right)}. \end{aligned} \tag{4.53}$$

The proof is a refined version of the unravelling presented in Appendix E. Note that, for $\sigma = 1$, we have $\tau = 1$ and $g(\sigma) = 1$ as expected so as to recover (4.47).

**Asymptotics of one-point functions.** The refined *one-point functions* $\mathcal{H}_n(\tau)$ and $H_n(\sigma)$ are simply defined as normalized refined partition functions

$$\mathcal{H}_n(\tau) = \frac{\mathcal{Z}_n(\tau)}{\mathcal{Z}_n}, \qquad H_n(\sigma) = \frac{Z_n(\sigma)}{Z_n}. \tag{4.54}$$

In the limit of large $n$, the asymptotics for the one-point function $H_n(\sigma)$ of the 6V model with DWBC and the weights $(a, b, c)$ of (4.46) above is characterized by the function

$$f(\sigma) = \lim_{n\to\infty} \frac{1}{n} \log\left(H_n(\sigma)\right), \tag{4.55}$$

whose expression is known [14] and may be given in parametric form as

$$\begin{aligned} f(\sigma(\xi)) &= \log\left(\frac{\sin(\alpha(\lambda-\eta)) \sin(\xi+\lambda-\eta) \sin(\alpha\xi)}{\alpha \sin(\lambda-\eta) \sin(\alpha(\xi+\lambda-\eta)) \sin(\xi)}\right), \\ \sigma(\xi) &= \frac{\sin(\lambda+\eta) \sin(\xi+\lambda-\eta)}{\sin(\lambda-\eta) \sin(\xi+\lambda+\eta)}, \end{aligned} \tag{4.56}$$

where

$$\alpha = \frac{\pi}{\pi - 2\eta}. \tag{4.57}$$

Using the correspondence (4.53), the parametrization $\sigma(\xi)$ of $\sigma$ translates into one for $\tau$, namely

$$\tau(\xi)=\frac{\sin(\lambda+\eta)\sin\left(\frac{\lambda+3\eta+\mu}{2}\right)\sin(\xi+\lambda-\eta)\sin\left(\xi+\frac{\lambda-\eta+\mu}{2}\right)}{\sin(\lambda-\eta)\sin\left(\frac{\lambda-\eta+\mu}{2}\right)\sin(\xi+\lambda+\eta)\sin\left(\xi+\frac{\lambda+3\eta+\mu}{2}\right)}. \tag{4.58}$$

Using the fact that $g(\sigma)$ is bounded independently of $n$ and Theorems 4.2.1 and 4.2.2 we find that

$$\lim_{n\to\infty}\frac{1}{n}\log\left(\mathcal{H}_n\left(\tau(\xi)\right)\right)=f(\sigma(\xi)). \tag{4.59}$$

Let us now introduce for future use the function

$$r(\tau)=\tau\frac{d}{d\tau}\left(\lim_{n\to\infty}\frac{1}{n}\log\left(\mathcal{H}_n(\tau)\right)\right). \tag{4.60}$$

The function $r(\tau)$ is given in parametric form by the above parametrization (4.58) of $\tau$ and the following parametrization for $r(\tau)$:

$$\begin{aligned}r(\tau(\xi))&=\frac{\tau(\xi)}{\partial_\xi\tau(\xi)}\partial_\xi f(\sigma(\xi))\\&=\big(\cot(\xi+\lambda-\eta)-\cot(\xi)+\alpha\cot(\alpha\,\xi)-\alpha\cot(\alpha(\xi+\lambda-\eta))\big)\\&\times\frac{\sin(\xi+\lambda+\eta)\sin(\xi+\lambda-\eta)\sin\left(\xi+\frac{\lambda-\eta+\mu}{2}\right)\sin\left(\xi+\frac{\lambda+3\eta+\mu}{2}\right)}{\sin(2\eta)\left[\sin(\xi+\lambda+\eta)\sin(\xi+\lambda-\eta)+\sin\left(\xi+\frac{\lambda-\eta+\mu}{2}\right)\sin\left(\xi+\frac{\lambda+3\eta+\mu}{2}\right)\right]},\end{aligned} \tag{4.61}$$

with $\alpha$ as in (4.57).

**Remark 4.2.3.** *In the particular case $\mu=\lambda-5\eta$, the parametrization $\tau(\xi)$ of $\tau$ simplifies into:*

$$\tau(\xi)=\frac{\sin(\lambda+\eta)\sin(\xi+\lambda-3\eta)}{\sin(\lambda-3\eta)\sin(\xi+\lambda+\eta)} \tag{4.62}$$

*and that of $r(\tau)$ into:*

$$\begin{aligned}r(\tau(\xi))=&(\cot(\xi+\lambda-\eta)-\cot(\xi)+\alpha\cot(\alpha\,\xi)-\alpha\cot(\alpha(\xi+\lambda-\eta)))\\&\times\frac{\sin(\xi+\lambda+\eta)\sin(\xi+\lambda-3\eta)}{\sin(4\eta)}.\end{aligned} \tag{4.63}$$

### 4.2.3 Tangent method for the case $\mu = \lambda - 5\eta = 0$

We now turn to the explicit computation of the arctic curve of the 20V model with DWBC2. We restrict ourselves to the case

$$\mu = 0, \quad \lambda = 5\eta, \tag{4.64}$$

where the weights thus depend on a single parameter $\eta$. From their general expression (4.44), it is easily checked that, *up to a global normalization factor* $\sin(2\eta)\sin^2(4\eta)$, the weights for the various vertex environments are all equal to 1 except for the empty and full vertices which have a weight

$$\varpi_0 = \frac{\sin(6\eta)}{\sin(2\eta)} = 1 + 2\cos(4\eta). \tag{4.65}$$

We use the notation $\varpi_0$ to recall that a global normalization factor $\sin(2\eta)\sin^2(4\eta)$ has been factored out in all the weights, i.e. $\omega_0 = \sin(2\eta)\sin^2(4\eta) \times \varpi_0$ while $\omega_i = \sin(2\eta)\sin^2(4\eta) \times 1$ for $i = 1$ to 6. In particular, the weights are invariant under the transformation $x \leftrightarrow y$, which implies that the arctic curve is *symmetric under reflection along the first diagonal.* In the particular case $\eta = \pi/8$, all the weights are equal to 1. The main result of this section is the following:

**Theorem 4.2.4.** *The E and V portions of arctic curve for the 20V model with DWBC2 at $\mu = 0$ and $\lambda = 5\eta$, as predicted by the application of the tangent method (using the shear trick for the V portion), has the following parametric equation:*

$$\begin{aligned}
x_E(\xi) &= 1 + \frac{\partial_\xi R_E(\xi)}{\partial_\xi S_E(\xi)}, \quad y_E(\xi) = R_E(\xi) - S_E(\xi)\frac{\partial_\xi R_E(\xi)}{\partial_\xi S_E(\xi)}, \quad \xi \in [0, \pi - 6\eta],\\
x_V(\xi) &= 1 + \frac{\partial_\xi R_V(\xi)}{\partial_\xi S_V(\xi)}, \quad y_V(\xi) = R_V(\xi) - S_V(\xi)\frac{\partial_\xi R_V(\xi)}{\partial_\xi S_V(\xi)}, \quad \xi \in [-2\eta, 0].
\end{aligned} \tag{4.66}$$

*where*

$$\begin{aligned}
R_E(\xi) &= h(\xi,\eta)\,\frac{\sin(\xi + 6\eta)\sin(\xi + 2\eta)}{\sin(4\eta)},\\
S_E(\xi) &= \frac{\sin(\xi + 6\eta)\sin(\xi + 2\eta)}{\sin(\xi)\sin(\xi + 4\eta)},\\
R_V(\xi) &= h(\xi,\eta)\,\frac{\sin(\xi + 6\eta)\sin(\xi + 2\eta)}{\sin(4\eta)},\\
S_V(\xi) &= \frac{\sin(\xi + 6\eta)\sin(\xi + 2\eta)}{\sin(\xi)\sin(\xi + 4\eta)} - \frac{\sin(\xi + 6\eta)\sin(\xi + 2\eta)}{\sin(\xi - 2\eta)\sin(\xi + 2\eta)},
\end{aligned} \tag{4.67}$$

*with* $\alpha = \pi/(\pi - 2\eta)$ *and*

$$h(\xi, \eta) = \cot(\xi+4\eta)-\cot(\xi)+\alpha\cot(\alpha\,\xi)-\alpha\cot(\alpha(\xi+4\eta)). \qquad (4.68)$$

The E portion $(x_{\mathrm{E}}(\xi), y_{\mathrm{E}}(\xi))$ is symmetric under $x \leftrightarrow y$. The $180°$ rotation yields the HDV portion $(1-x_{\mathrm{E}}(\xi), 1-y_{\mathrm{E}}(\xi))$. The range of $\xi$ does not change when applying the symmetries and is still $\xi \in [0, \pi-6\eta]$ for the HDV portion. The V portion $(x_{\mathrm{V}}(\xi), y_{\mathrm{V}}(\xi))$ can be used to deduce three other portions. Applying to it the $x \leftrightarrow y$ symmetry, we obtain the H portion $(y_{\mathrm{V}}(\xi), x_{\mathrm{V}}(\xi))$. Applying the $180°$ rotation symmetry to the V and H portions respectively yields the HD portion $(1-x_{\mathrm{V}}(\xi), 1-y_{\mathrm{V}}(\xi))$ and the DV portion $(1-y_{\mathrm{V}}(\xi), 1-x_{\mathrm{V}}(\xi))$. Theorem 4.2.4 hence completely describes the arctic curve.

The application of the tangent method to compute the E portion is almost identical to that presented in Section 2.2—it actually leads to the exact same curve in the uniform case ($\gamma = 1$ and $\omega_0 = \cdots = \omega_6$). We will compute the V portion, whose derivation presents new subtleties, as opposed to the E portion which we will hence not consider (see [17] for details).

**Shear trick.** The V portion was originally computed by Di Francesco and Guitter in [17] using the tangent method via the *shear trick*. The core idea is to resort to a different family of osculating paths, designed in such a way that the V region is empty. The construction is described in Figure 4.13.

As depicted in Figure 4.14-left, the shear trick creates node environments which form some "inverted" 20V model, with a set of vertices identical to those of the original 20V model, up to a left-right symmetry. Note that the shear trick does *not* transform a given vertex into its left-right symmetric vertex: for instance, the empty vertex is mapped onto the vertex with exactly two occupied vertical edges and conversely. The inverted 20V model has modified weights inherited from its pre-image: the non-trivial weight $\varpi_0 = 1 + 2\cos(4\eta)$ is now attached to the vertex with two occupied vertical edges or its complementary vertex (left column of Figure 4.14-left).

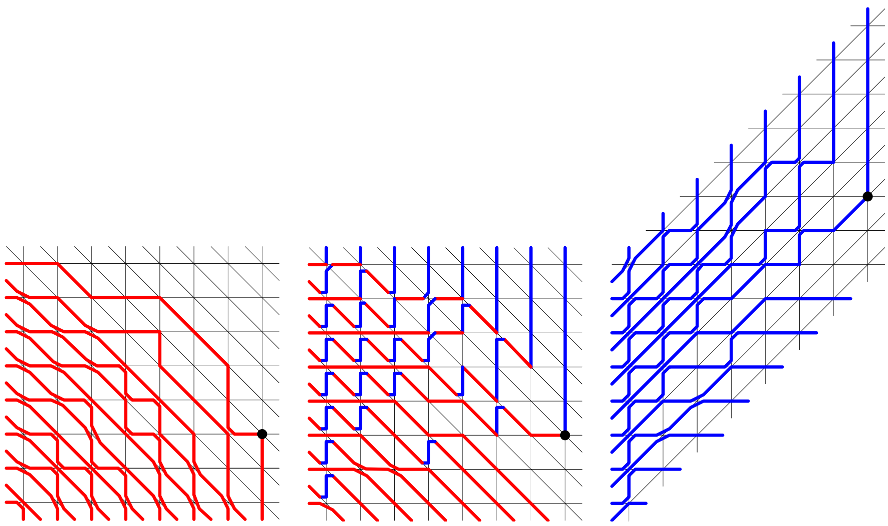


Figure 4.13: Shear trick. *Left*: Configuration of the 20V model with DWBC2. *Middle*: Occupied and non-occupied *vertical* edges are interchanged (the new ones are draw in blue). *Right*: A global shear of the lattice transforms the original square domain into a rhombus. The re-connection of the occupied steps into osculating paths shown in the middle is such that on the rhombus, all the paths are oriented from lower left to upper right at each node.

The shear turns the domain into a rhombus with particular boundary conditions, see Figure 4.13. Ordering the paths from bottom to top according to their starting (diagonal or horizontal) edge on the left boundary, the $(n-1)$ first paths (i.e. the lower "half" of the paths) now have their final step at the first $n-1$ successive horizontal edges along the lower (diagonal) boundary of the rhombus while the $n$ last paths (i.e. the upper "half" of the paths) have their final step at the successive vertical edges along on the upper (diagonal) boundary of the rhombus.

**Tangent method.** The desired new portion of arctic curve is obtained in the scaling limit as the frontier between the now empty region and that occupied by the paths *reaching the upper boundary of the rhombus*. The tangent method set-up used to access this portion is presented on Figure 4.14-right. The $n$-th path endpoint is moved from the rightmost vertical edge of the upper boundary at $(n, 2n)$ to a horizontally shifted position at $(n+M, 2n)$. It enters the original domain at some position $(n, n+L-1)$, defined as the position where the path hits the right boundary of the original rhombic domain for the *first time*. In the rescaled domain we work with $m \equiv M/n$ and $\ell \equiv L/n$.

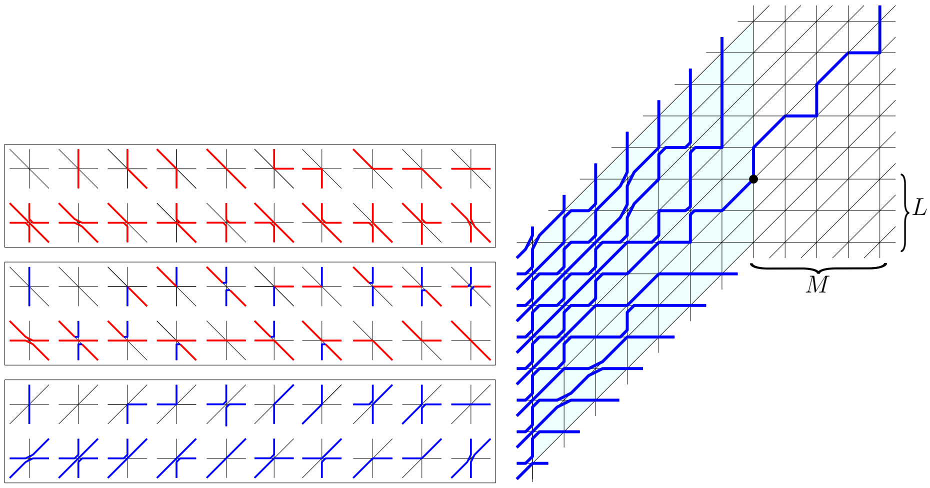



Figure 4.14: *Left*: Modification of the twenty vertices as a result of the shear trick of Figure 4.13. *Right*: Tangent method set-up. The uppermost endpoint of the domain is moved to the right (while remaining vertical) to $(n+M, 2n)$ and exits the original domain (in light blue) at position $(n, n+L-1)$, indicated by a black dot.

The most likely escape point position $\ell$ for a given $m$ is obtained by a saddle point analysis of the partition function of the extended problem,

$$\sum_{L=1}^{n} \frac{\mathcal{H}_n(\tau)|_{\tau^L}}{\varpi_0^{n-L}} \, \bar{\mathcal{Y}}_{(n,n+L-1)\to(n+M,2n)}. \tag{4.69}$$

In the first term, $\mathcal{H}_n(\tau)|_{\tau^L}$ corresponds to the contribution of the modified osculating path inside the rhombus. We use the one-to-one correspondence between the two families of osculating paths, that links the hitting points with the right boundary, see the black dots in Figure 4.13. If the hitting point is at $(n, L)$ in the original description, it is at $(n, n+L-1)$ after the shear trick. To be fully consistent, since empty vertices have a weight 1 in the inverted 20V description, we normalize $\mathcal{H}_n(\tau)|_{\tau^L}$ by $\varpi_0^{n-L}$ to factor out the weights of the vertical piece of path from $(n, n+L-1)$ to $(n, 2n)$ which are not present in the extended geometry. The second term $\bar{\mathcal{Y}}_{(n,n+L-1)\to(n+M,2n)}$ denotes the partition function for a single path from $(n, n+L-1)$ to $(n+M, 2n)$ (starting possibly with a number of vertical steps) *in the inverted 20V model setting* with, at each node along the path, the same weight as that of the corresponding inverted 20V configuration.

To compute the large $n$ asymptotics of $\bar{\mathcal{Y}}_{(n,n+L-1)\to(n+M,2n)}$, namely

$$\bar{\mathcal{Y}}_{(n,n+L-1)\to(n+M,2n)} \approx \mathrm{e}^{n\,\bar{S}(\ell,m)}, \tag{4.70}$$

we resort to the transfer matrix approach of Appendix D in a version adapted to the present shear geometry: this corresponds to setting $\alpha_1 = \varpi_0$, $\alpha_2 = 1$, $\alpha_3 = 2-\varpi_0$, $\alpha_4 = 1-\varpi_0$ and $\alpha_5 = \alpha_6 = 0$ and performing the change $\ell \to 1-\ell$ as the height difference for the path beyond the escape point is now measured from the top, hence equal to $1-\ell$ instead of $\ell$. We therefore have

$$\begin{aligned}
&\bar{S}(\ell,m) = \bar{S}(\ell,m,p_3,p_4)\\
&= \mathcal{L}((1-\ell)+m-p_3-2p_4) - \mathcal{L}(p_3) - \mathcal{L}(p_4)\\
&- \mathcal{L}((1-\ell)-p_3-2p_4) - \mathcal{L}(m-p_3-p_4)\\
&+ ((1-\ell)-p_3-2p_4)\log\varpi_0 + p_3\log(2-\varpi_0) + p_4\log(1-\varpi_0)
\end{aligned} \tag{4.71}$$

taken at the values of $p_3$ and $p_4$ which maximize $\bar{S}(\ell,m,p_3,p_4)$ for fixed $\ell$ and $m$. We used $\mathcal{L}(x) \equiv x\log(x)$.

The extremization conditions over $\tau$ and $\ell$ now read

$$\begin{aligned}
&\frac{d}{d\tau}\left(-\ell\log\tau + f(\sigma(\tau))\right) = 0,\\
&\frac{d}{d\ell}\left(-\ell\log\tau + \bar{S}(\ell,m) - (1-\ell)\log\varpi_0\right) = 0,
\end{aligned} \tag{4.72}$$

or equivalently

$$\ell = r(\tau), \quad \tau = \varpi_0\,\mathrm{e}^{\frac{d}{d\ell}\bar{S}(\ell,m)}. \tag{4.73}$$

In the new geometry of Figure 4.14-right, the corresponding tangent line passing trough $(1, 1+\ell)$ and $(1+m, 2)$ has equation $y+\frac{\ell-1}{m}(x-1)-\ell-1 = 0$. Therefore, *after performing the inverse shear transformation* $y \to y-x$ to bring back the path in the original geometry of a square domain, the equation of the tangent line reads: $y+\left(1-\frac{1-\ell}{m}\right)(x-1)-\ell = 0$. For the most likely escape point, this yields the tangent line equation:

$$y + \bar{s}(\tau)(x-1) - r(\tau) = 0, \tag{4.74}$$

where the slope $\bar{s}(\tau)$ is now given by

$$\bar{s}(\tau) = 1 - \frac{1 - r(\tau)}{m(\tau)}, \tag{4.75}$$

with $m(\tau)$ solution of $\tau = \varpi_0 \exp(\frac{\mathrm{d}}{\mathrm{d}\ell}\bar{S}(\ell, m))$ at $\ell = r(\tau)$. Writing the extremization conditions $\partial_{p_3}\bar{S} = \partial_{p_4}\bar{S} = 0$ and solving (4.75) yields now

$$\bar{s}(\tau) = \frac{\tau(1+\varpi_0)}{(\tau-1)(\tau+\varpi_0)} + \frac{\varpi_0}{\tau(1-\varpi_0)+\varpi_0}, \tag{4.76}$$

which implies

$$\bar{s}(\tau(\xi)) = \frac{\sin(\xi+6\eta)\sin(\xi+2\eta)}{\sin(\xi)\sin(\xi+4\eta)} - \frac{\sin(\xi+6\eta)\sin(\xi+2\eta)}{\sin(\xi-2\eta)\sin(\xi+2\eta)}. \tag{4.77}$$

We end up with a family of tangent lines parametrized by $\xi$ via

$$0 = \bar{F}(x, y, \xi) \equiv y + \bar{s}(\tau(\xi))(x-1) - r(\tau(\xi)). \tag{4.78}$$

The parameter $\xi$ spans the range $[-2\eta, 0]$. The range $[-2\eta, -\eta]$ corresponds to $\bar{s}(\tau(\xi))$ decreasing from 1 to 0, while the range $[-\eta, 0]$ corresponds to $\bar{s}(\tau(\xi))$ decreasing from 0 to $-\infty$, as dictated by the tangent method in the present geometry.

The corresponding V portion of arctic curve, is obtained by solving the linear system $\bar{F}(x, y, \xi) = \partial_\xi \bar{F}(x, y, \xi) = 0$ which yields a parametric expression $(x(\xi), y(\xi))$

$$x(\xi) = 1 + \frac{\partial_\xi r(\tau(\xi))}{\partial_\xi \bar{s}(\tau(\xi))}, \quad y(\xi) = r(\tau(\xi)) - \bar{s}(\tau(\xi))\frac{\partial_\xi r(\tau(\xi))}{\partial_\xi \bar{s}(\tau(\xi))}, \quad \xi \in [-2\eta, 0]. \tag{4.79}$$

This completes the derivation of the V portion, with the identifications $R_{\mathrm{V}}(\xi) = r(\tau(\xi))$ and $S_{\mathrm{V}}(\xi) = \bar{s}(\tau(\xi))$.

### 4.2.4 Characteristics of the arctic curve

We now study the behaviour of the arctic curve, for the general weights (4.44), whose (cumbersome) equation is given in [17] and was obtained using the tangent method in a very similar way as above. In the general case the reflection $x \leftrightarrow y$ is not a symmetry. However, the model is symmetric under the *simultaneous* exchange $x \leftrightarrow y$ and $\mu \leftrightarrow -\mu$ and the V portion again enables to compute three other portions. We describe the shape of the arctic curve as a function of the parameters $\mu$, $\lambda$ and $\eta$ which—we remind—are constrained by

$$0 < \eta < \lambda < \pi - \eta, \quad \eta - \lambda < \mu < \lambda - \eta. \tag{4.80}$$

We limit our analysis to the $\mu > 0$ case, without loss of generality since it is easily generalisable using the $x \leftrightarrow y$ and $\mu \leftrightarrow -\mu$ symmetry.

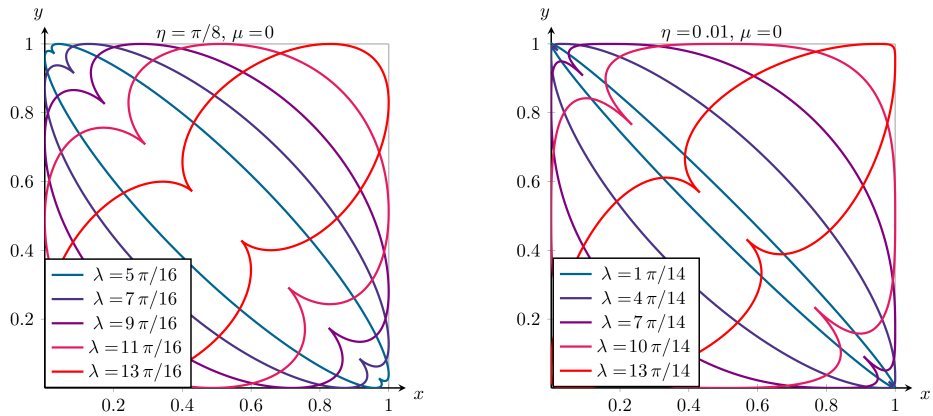



Figure 4.15: Arctic curves for $\mu = 0$ and various values of $\lambda \in ]\eta, \pi - \eta[$ . *Left*: $\eta = \pi/8$. *Right*: $\eta \to 0$. In practice we set $\eta = 0.01$.

**Reflection symmetry.** The arctic curves are all invariant under rotation by 180°. The parameter $\mu$ controls its asymmetry under the $x \leftrightarrow y$ transformation. Figure 4.15 displays arctic curves for $\mu = 0$, hence symmetric under $x \leftrightarrow y$. The left figure is for $\eta = \pi/8$ with a varying $\lambda$ in the admissible range $]\pi/8, 7\pi/8[$, while the right figure is for a value of $\eta \to 0$ with a varying $\lambda$ in the range $]0, \pi[$. In both cases, the limiting curve $\lambda \to \eta$ degenerates into the second diagonal $x + y = 1$. It may be understood in the light of the fact that for $\omega_2 = \omega_3 = \omega_5 = \omega_6 = 0$, only one configuration survives, with all edges occupied below the second diagonal and none above. Similarly, in the limit $\lambda \to \pi - \eta$, the arctic curve degenerates into the first diagonal $y = x$. In this case, it may be understood by noticing that when $\omega_0 = 0$, the only allowed configuration is one with occupied edges that are all horizontal above $y = x$ and all vertical below, with additional diagonal steps below $y + x = 1$.

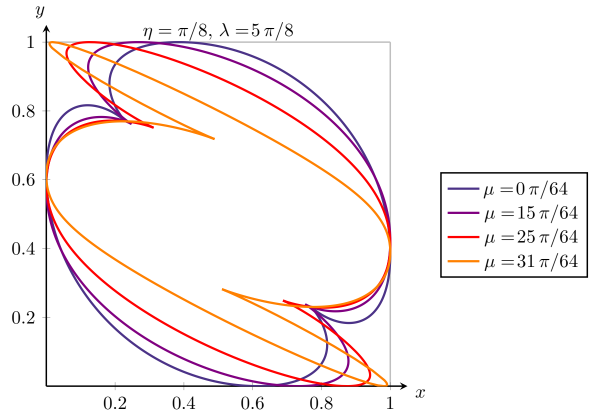



Figure 4.16: Arctic curves for $\eta = \frac{\pi}{8}$, $\lambda = \frac{5\pi}{8}$ and various values of $\mu \in [0, \lambda - \eta[$.

Figure 4.16 shows the increasing asymmetry of the arctic curve under $x \leftrightarrow y$ for increasing $\mu$ in the range $]0, \pi/2[$ at fixed values $\eta = \pi/8$ and $\lambda = 5\pi/8$. When the parameter $\mu$ tends to its maximal value $\lambda - \eta$, the arctic curve displays two outgrowths which become narrower and narrower and eventually degenerate into two segments so that the arctic curve is formed of a single convex curve (see Figure 4.17-left).

**Limiting degenerate cases.** For $\mu = \lambda - \eta$, the two aforementioned limiting segments are tangent to the arctic curve and have a slope $-1/2$ independently of $\lambda$ and $\eta$. This phenomenon has a simple explanation: for $\mu = \lambda - \eta$, we have $\omega_6 = 0$ so that the H and DV frozen phases cannot exist anymore. A direct transition then takes place between the frozen phases HD and E (resp. between HDV and V). The slope $-1/2$ of the transition line is directly dictated by the path geometry, as illustrated in Figure 4.17-left). The evolution of the $\mu = \lambda - \eta$ convex arctic curve with fixed $\eta = \pi/16$ and varying $\lambda$ is represented in Figure 4.17-right.

Figure 4.18-left displays the evolution of the arctic curve with $\mu$ for a very small $\eta$ and a value of $\lambda$ close to its maximal possible value $\pi - \eta$. When $\mu$ gets close to its maximal value $\lambda - \eta$, the arctic curve is formed of two symmetric convex pieces connected by a narrow isthmus. More precisely, a well-defined limit is reached by setting

$$\eta = \epsilon\, \Lambda_1, \quad \lambda = \pi - \epsilon(\Lambda_1 + \Lambda_2), \quad \mu = \pi - \epsilon(2\Lambda_1 + \Lambda_2 + \Lambda_3) \qquad (4.81)$$

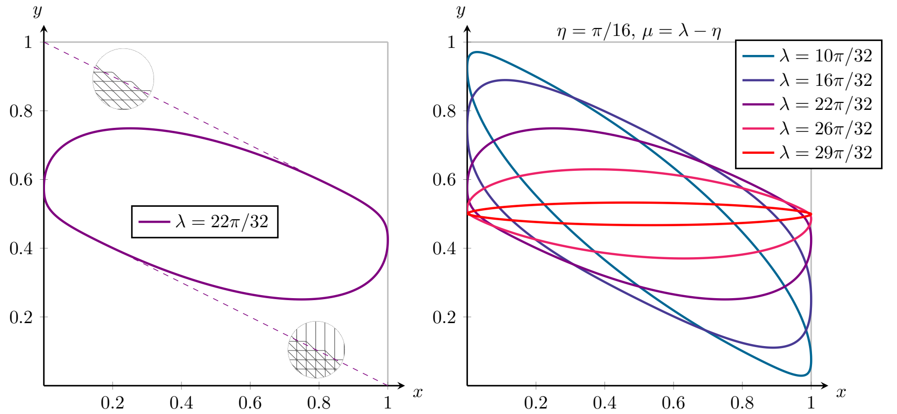



Figure 4.17: Arctic curves for $\eta = \pi/16$, $\mu = \lambda - \eta$ and various values of $\lambda \in ]\eta, \pi - \eta[$. At $\mu = \lambda - \eta$, a direct transition takes place between HD and E (resp. between HDV and V) along a segment of slope $-1/2$ dictated by the local path geometry, as shown on the left, for $\lambda = 22\pi/32$.

and sending $\epsilon \to 0$, with fixed positive real numbers $\Lambda_1$, $\Lambda_2$ and $\Lambda_3$ so that the parameters remain in the admissible range (4.45). Up to an overall $\epsilon^3$ factor, all the $\omega_i$'s remain finite and are positive homogeneous polynomials of degree 3 in the $\Lambda_i$'s. In this limit, the arctic curve tends to an algebraic curve made of two symmetric pieces: the first piece lies in the upper half of the rescaled square domain and is tangent to the lines $y = 1$, $x = 1$ and $y = 1/2$. The second piece is its image under 180° and lies in the lower half of the square. As for the isthmus observed in Figure 4.18, it degenerates into a horizontal segment joining the two tangency points along the $y = 1/2$ line. This segment corresponds to a new, direct transition between the HD and V frozen regions and its horizontality is dictated by the path geometry.

The same phenomenon of splitting of the arctic curve is in fact observed by keeping a finite value of $\eta$ whenever $\lambda$ and $\mu$ tend to their maximal admissible values, namely $\lambda = \pi - \eta - \epsilon\,\Lambda_2$ and $\mu = \pi - 2\eta - \epsilon(\Lambda_2 + \Lambda_3)$ with $\epsilon \to 0$, $\Lambda_2$ and $\Lambda_3$ finite and positive. This leads to a limiting arctic curve depending on $\Lambda_2$ and $\Lambda_3$ but *independent of* $\eta$ (assuming that $\eta$ itself is kept fixed and does not scale with $\epsilon$) corresponding to limiting weights $\omega_0, \omega_4, \omega_5, \omega_6 \to 0$ and (projectively) $\omega_1, \omega_2, \omega_3 \to 1$. This $\eta$-independent limit may be reached in the above setting (4.81) by sending $\Lambda_1 \to \infty$. Remarkably, the corresponding limiting arctic curve

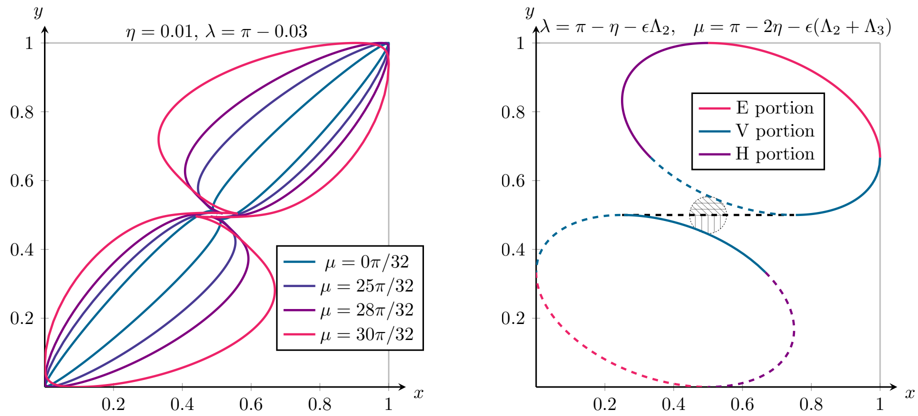



Figure 4.18: *Left*: Arctic curves for $\eta \to 0$ (here $\eta = 0.01$), $\lambda + \eta \to \pi$ (here $\lambda = \pi - 0.03$) and various values of $\mu \in [0, \lambda - \eta[$. *Right*: Arctic curve for $\lambda = \pi - \eta - \epsilon\,\Lambda_2$ and $\mu = \pi - \eta - \epsilon(\Lambda_2 + \Lambda_3)$ with $\epsilon \to 0$, for $\Lambda_3/\Lambda_2 = 2$ and some arbitrarily fixed $\eta$. The limiting shape is independent of $\eta$ and formed of two symmetric ellipses tangent to the $y = 1/2$ line. Note that the V portion gives rise to two portions in this limit, one in each ellipse. The black dashed segment is the locus of the transition between the HD and V frozen regions.

is made of two ellipses (exchanged by 180° rotation) with respective equations

$$\begin{aligned} &\big(\Lambda_3(x-y) + 2\Lambda_2(x+2y-2)\big)^2 + 8\Lambda_2\Lambda_3(1-y)(1-2y) = 0, \\ &\big(\Lambda_3(x-y) + 2\Lambda_2(x+2y-1)\big)^2 - 8\Lambda_2\Lambda_3 y(1-2y) = 0. \end{aligned} \tag{4.82}$$

In particular, these curves depend on the ratio $\Lambda_3/\Lambda_2$, hence on the precise way in which the weights $\omega_i$ reach their limiting values 0 or 1. The corresponding arctic curve for $\Lambda_3/\Lambda_2 = 2$ is displayed in Figure 4.18-right.

Finally, let us mention that the arctic curve of the 6V model with DWBC is a limiting case of the 20V arctic curve. The 6V model with DWBC may be realized as a particular instance of the 20V model with DWBC2 by setting $\omega_2 = \omega_5 = 0$. Indeed, the condition $\omega_2 = \omega_5 = 0$ forces the diagonal steps to be transmitted through each node so as to arrange into $n-1$ complete diagonal lines below the second diagonal. These lines may be removed and the remaining paths, made of horizontal and vertical steps only, form configurations of a 6V model with DWBC. Using this observation, the result of [14] can be derived (in the D regime) using the one of the 20V model with DWBC2, see [17].

**Uniform case.** To conclude this subsection, let us discuss the uniform case where all $\omega_i$'s are equal to 1, namely when $\mu = 0$, $\lambda = 5\eta$ and $\eta = \pi/8$. The solution for the E portion of the arctic curve of Theorem 4.2.4 is given explicitly by

$$\begin{aligned} x_{\mathrm{E}}(\xi) &= \frac{3}{18}\Big(5\cos\big(\frac{2\xi}{3}\big)+\cos\big(\frac{10\xi}{3}\big)\Big)-\frac{\sqrt{3}}{18}\Big(5\sin\big(\frac{2\xi}{3}\big)-\sin\big(\frac{10\xi}{3}\big)\Big), \\ y_{\mathrm{E}}(\xi) &= \frac{\sqrt{3}}{18}\Big(5\cos\big(\frac{2\xi}{3}\big)-\cos\big(\frac{10\xi}{3}\big)\Big)+\frac{3}{18}\Big(5\sin\big(\frac{2\xi}{3}\big)+\sin\big(\frac{10\xi}{3}\big)\Big), \end{aligned} \tag{4.83}$$

with $\xi \in [0, \pi/4]$. This describes a portion of some algebraic curve with equation

$$\begin{aligned} &3^{11}(x^2+y^2)^5 + 3^9\,10(x^2+y^2)^4 - 3^6\,5(x^2+y^2)^3 \\ &\quad + 6^2\,20(73(x^2+y^2)^2 - 5^4x^2y^2) - 2^8\,15(x^2+y^2) - 2^{12} = 0, \end{aligned} \tag{4.84}$$

which is exactly the (whole) arctic curve of the uniform ($\gamma = 1$) QTHADT. This is a direct consequence of the correspondence (first proved in [59], see Theorem 5.2) between the refined partition functions of these two models, which generalizes

$$\mathcal{Z}_n = A_n(\gamma=1) = Z_n[1, \sqrt{2}, 1] = 1, 3, 23, 433, 19705, \cdots . \tag{4.85}$$

As for the V portion of the 20V model with DWBC2, it is given by

$$(x_{\mathrm{E}}(\xi), y_{\mathrm{E}}(\xi) - x_{\mathrm{E}}(\xi) + 1), \tag{4.86}$$

with $\xi \in [-\pi/4, 0]$. Otherwise stated, in the uniform case, the V portion is obtained from the analytic continuation of the E portion by a simple shear transformation sending the $y = 0$ line onto the line $x + y = 1$.

Remarkably, as shown in Figure 4.19-left, we find exactly the same pattern of correspondences if we compare the arctic curve for rotationally symmetric rhombus tilings of a holey hexagon, in bijection with DPP [154] to that of the uniform 6V model (with weights $a = b = c$) with DWBC, in bijection with ASM (see Section 5.5). In the geometry of

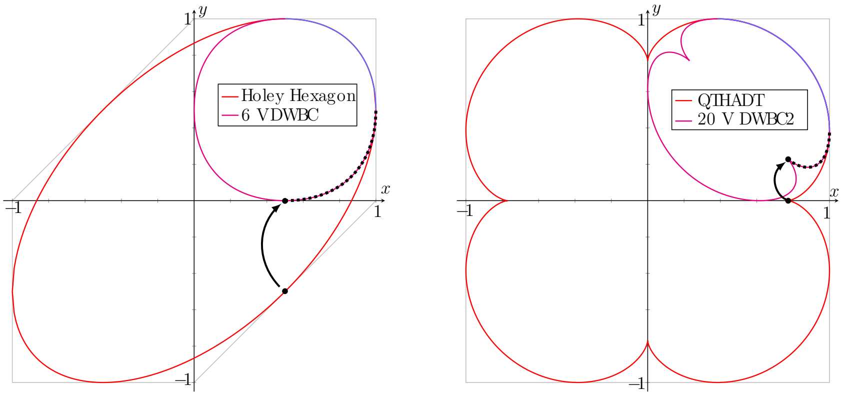



Figure 4.19: A shear transformation $y \to 1 - x + y$ relates the dashed portions of arctic curves of some *uniform* vertex models and the arctic curve of some *uniform* rotationally symmetric tilings of some holey domains. *Left*: 6V model with DWBC and rotationally symmetric rhombus tilings of a holey hexagon, whose arctic curve is an ellipse. *Right*: 20V model with DWBC2 and QTHADT. In both cases, the purple portion is common in both cases and the whole arctic curve of the tiling model is its analytic continuation, drawn in red.

Figures 2.12-right and 4.19-left, the arctic curve of the rotationally symmetric holey hexagon rhombus tiling is obtained via (2.56) with $\eta = \pi/6$, upon taking $\xi$ in the *extended* range $\xi \in [-5\pi/6, 7\pi/6]$, leading to the complete ellipse $x^2 + y^2 - xy = 3/4$. As to the uniform 6V model with DWBC, the north-east portion of its arctic curve is exactly a piece of this ellipse [13]. This is a direct consequence of the refined ASM-DPP correspondence, proved in its highest degree of refinement in [99, 155]. The three remaining portions are obtained by successive $90°$ rotations around the center of the domain. Again, as displayed in Figure 4.19-left, the south-east portion is the image by the same shear transformation $y \to 1 - x + y$ of a portion of ellipse. This recurrent pattern may hide a more general phenomenon for correspondences between osculating and non-intersecting paths problems, yet to be investigated. Finding such a correspondence would probably be hard. Indeed, since the discovery almost 40 years ago of the ASM-DPP correspondence, no bijection between the set of ASM and that of DPP has been found, despite recent promising bijective results [156–158].

### 4.2.5 Simulations

**Method.** Numerical studies of the arctic curve phenomenon in the 6V model with DWBC (or "domain-wall-like" boundary conditions) are numerous [104–106,109,111,141]. In this subsection, we adapt a Markov chain Monte Carlo method due to Allison and Reshetikhin discussed in [104–106]. This algorithm exploits the bijection with osculating paths to design a local-move Markov chain whose stationary distribution is that associated with the weights (4.44). For the sake of definiteness, we consider the 20V model with DWBC1, but the discussion is easily extended to other fixed boundary conditions, including DWBC2.

Let us first consider the model with uniform distribution, i.e. with all the $\omega_i$'s equal ($\eta = \pi/8$, $\lambda = 5\eta$ and $\mu = 0$). The Markov chain starts from an allowed configuration (for example, the "diagonal" configuration displayed in Figure 4.20-right) and preserves the non-crossing property at each step. At any given iteration, either the configuration remains unchanged or some elementary move is performed on a plaquette. The algorithm goes as follows: start by selecting a plaquette at random, and, if the plaquette has at least a section of path connecting its north-west to its south-east corner, randomly choose which section of path to update ($\diagdown$, $\llcorner$ or $\urcorner$). The four possible moves are $\diagdown \to \llcorner$, $\llcorner \to \diagdown$, $\diagdown \to \urcorner$ and $\urcorner \to \diagdown$, if allowed by the local environment. If no move is possible for this selected section of path, remain in the same configuration for this step of the chain (the attempted move is rejected). Otherwise, apply the rules $\mathrm{R}_i$, $i = 1, 2, \cdots, 7$ presented in Figure 4.20-left. Repeat the process until the number of steps is sufficient (see the discussion below).

Since every configuration of the model can be obtained from any other configuration by finitely many elementary moves and since for any configuration there always exists a move whose rejection probability is non-vanishing, the Markov chain is ergodic. One can also verify that the transition probability $p(\mathcal{C} \to \mathcal{C}')$ from a configuration $\mathcal{C}$ to a configuration $\mathcal{C}'$ is symmetric: $p(\mathcal{C}' \to \mathcal{C}) = p(\mathcal{C} \to \mathcal{C}')$. In particular, the two possibilities for the output in rules $\mathrm{R}_5$ and $\mathrm{R}_7$ are precisely designed so as to ensure this property. Hence, the detailed balance condition is satisfied for the uniform distribution, which implies that the stationary distribution of this Markov chain is uniform, as required.

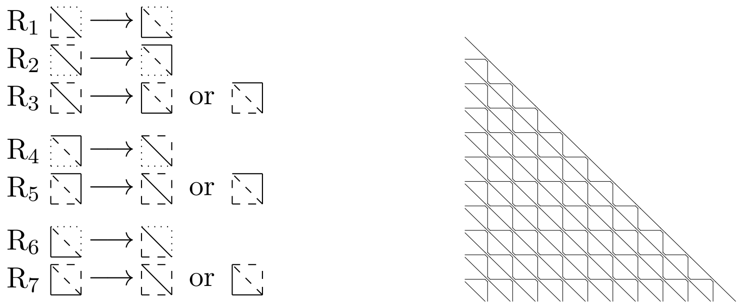



Figure 4.20: *Left*: Rules $\mathrm{R}_1, \cdots, \mathrm{R}_7$ used to perform the elementary moves of the Markov chain. Plain lines indicate occupied edges, dashed lines empty edges while at least one of the dotted line is occupied. When two outputs are drawn, one is chosen with probability 1/2. *Right*: "Diagonal" configuration used as initial state. It is made of a triangular HDV region and a triangular E region.

The uniform version of the algorithm is therefore such that, at every iteration, we either stay in the same configuration $\mathcal{C}$ (the move is rejected) or we perform a move and obtain a new configuration $\mathcal{C}'$. Generalizing the algorithm to the non-uniform case of arbitrary weights (4.44) can then be achieved by further rejecting some of the moves in a way that depends on the weight of the putative new configuration $\mathcal{C}'$. Assume that the move is attempted on a plaquette whose center is located at $(i, j)$ and call $W_{i,j}(\mathcal{C}')$ the product of the weights of the four nodes around this plaquette in the configuration $\mathcal{C}'$. The move is accepted (in a similar manner as in [104–106]) with a probability equal to

$$P = \frac{W_{i,j}(\mathcal{C}')}{W_0}, \tag{4.87}$$

where the normalisation $W_0 = (\max_k \omega_k)^4$ ensures that $P \leq 1$. This Markov chain is ergodic and satisfies the DBC for the probability distribution induced by the relative weights $\omega_i$. The ability of this algorithm to generate configurations with the correct frequency was tested for small sizes and for several choices of $\eta$, $\lambda$ and $\mu$ (see Figure 4.21 for an example).

The Markov chain converges to the desired distribution in the limit of an infinite number of iterations. In practice, we need a criterion to decide when to stop the simulation: the main criterion we used, also

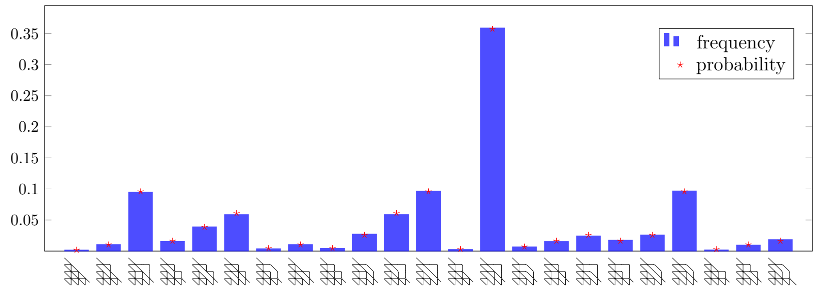



Figure 4.21: Observed frequencies of the 23 possible configurations of the 20V model with DWBC1 at $n = 3$, compared with the corresponding theoretical probabilities, here for $\eta = \pi/12$, $\lambda = 10\pi/12$ and $\mu = \lambda - 5\eta$. The frequencies were measured from 23000 generated configurations.

invoked in [104, 105], is the stabilization of the arctic curve, namely that no qualitative change in this curve is observed. In the cases where the convergence is slow, our estimation is checked against that obtained by running another Markov chain starting from a completely different configuration, as is done in [106], and we make sure that the results are comparable.

Notice that instead of (4.87) one can alternatively use the Metropolis probability

$$P = \min\left(1, \frac{W_{i,j}(\mathcal{C}')}{W_{i,j}(\mathcal{C})}\right). \tag{4.88}$$

For generic $\omega_i'$s, this choice lowers the rejection of moves and hence increases the thermalization speed. Unless stated otherwise we used (4.87).

**Results.** We take the average over multiple configurations and use as order parameter the local density of diagonal steps. Indeed this average density is 1 in the frozen regions HD, DV and HDV and 0 in the H, E and V regions, while it varies continuously in the disordered region. It is hence a good indicator for the position of the arctic curve. The evolution of the arctic curve with varying parameters is shown in Figures 4.22, 4.23 and 4.24. In all cases we also indicate the theoretical prediction (dashed curve): the agreement is quite good, despite what looks like an "attraction" of the finite-size arctic curve towards the disordered region.

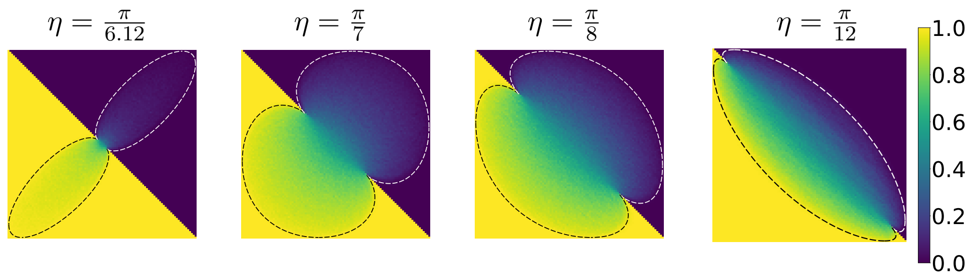



Figure 4.22: Local density of diagonal steps of the 20V model with DWBC1 for $\lambda = 5\eta$, $\mu = 0$ and $n = 100$, for several values of $\eta \in ]\pi/6, \pi/12]$. The order parameter is averaged over 1000 configurations. The dashed curve is the analytic prediction for the arctic curve.

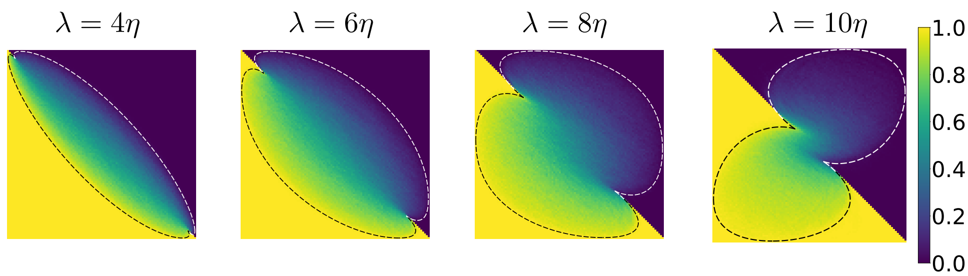



Figure 4.23: Local density of diagonal steps of the 20V model with DWBC1 for $\eta = \frac{\pi}{12}$, $\mu = \lambda - 5\eta$ and $n = 100$, for several values of $\lambda \in [4\eta, 10\eta]$. The order parameter is averaged over 1000 configurations. The dashed curve is the analytic prediction for the arctic curve.

The average outermost path is shown in Figure 4.25, for different sizes, in the uniform case. The figure shows finite size corrections to the arctic curve. As discussed in the previous chapter, in the case of Aztec diamonds, the correction is of order $n^\alpha$ with $\alpha = -2/3$, in the rescaled domain. We suppose that it is also the case for the model at hand. We estimate the curve reached asymptotically when $n \to \infty$, in the coordinates $u = (y + x)/2$ and $v = (y - x)/2$. For a fixed $v$ let us call $u_n(v)$ the corresponding $u$ on the average path in size $n$. We assume the scaling $u_n(v) = u(v) - n^{-2/3}\mathrm{corr}(v)$ and extract an estimation of the arctic curve $u(v)$, see the black dots in Figure 4.25. As a consistency check, one can alternatively estimate the scaling exponent $\alpha$ defined by $u(v) - u_n(v) = n^\alpha\mathrm{corr}(v)$ by using for $u(v)$ the theoretical prediction. It is found that on average $\alpha = -0.627$ with a standard deviation of $\sigma = 0.103$, which is consistent with the previous assumption, bearing in mind that next order corrections might play a role for the small sizes considered.

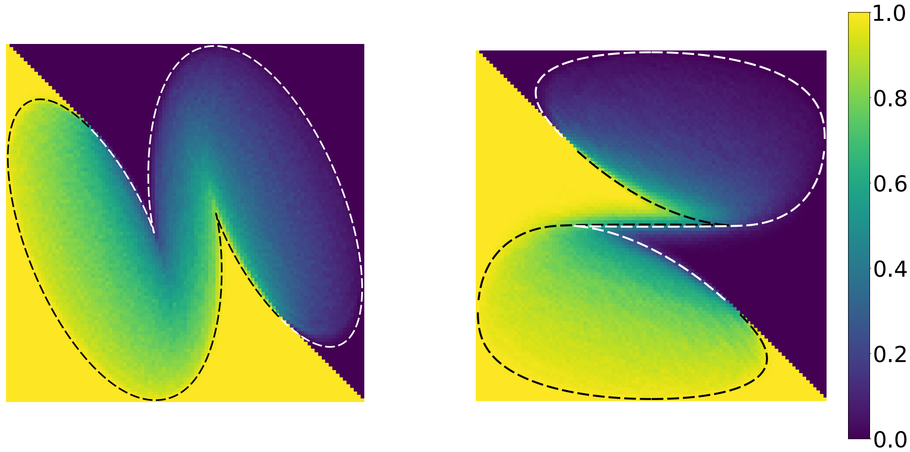



Figure 4.24: Local density of diagonal steps of the 20V model with DWBC1 for $n = 100$ in the case $\mu \neq \lambda - 5\eta$. The Metropolis version of the algorithm was used. *Left*: $\eta = \pi/6$, $\lambda = 9\pi/12$ and $\mu = -\pi/2$. *Right*: $\eta = 1/200$, $\lambda = \pi - 3\eta$ and $\mu = \lambda - 5\eta$.

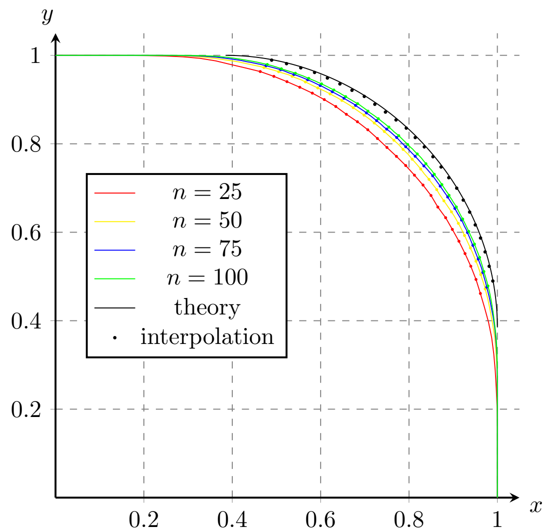



Figure 4.25: Averaged outermost path position (500 configurations) of the 20V model with DWBC1 and uniform weights for $n = 25, 50, 75, 100$. The interpolating curve is compared with the theoretical arctic curve.

# Chapter 5

# Variational approach

This chapter is based on [159, 160] written with Philippe Ruelle and Etienne Granet

---

The tangent method is believed to apply to a large variety of systems which can be formulated in terms of interacting directed lattice paths. In this chapter, we assume that the paths are non-crossing. Other types of interactions may also potentially lead to an arctic phenomenon that is tractable using the tangent method. Its domain of applicability is not precisely known. The method distinguishes the most external path from the other paths, which are responsible for the formation of the arctic curve itself. It essentially relies on two basic assumptions, as explained in Subsection 2.1.2. Assumption (i) states that in the scaling limit, the outermost path alone has no effect on the arctic curve, neither on its existence nor on its shape. The main argument for this is that the arctic curve is a volume effect that cannot be counterbalanced by a single path. Thus, the outermost path can be modified by changing its starting or ending point, or both, without interfering with the arctic curve. This brings up the second main assumption (ii), based on many observations: the modified external path has straight portions which connect tangentially to the arctic curve (in the scaling limit, the external path becomes a class $\mathcal{C}^1$ curve almost surely). It is this *tangency* *assumption* that makes the tangent method so efficient, allowing to determine the analytical shape of the arctic curve from a family of tangent straight lines.

The concrete implementation of the method is model-dependent but its principle is believed to be fairly universal. A general, model-independent, proof of assumptions (i) and (ii) is probably challengingly hard as it would be close to giving the shape of the arctic curve itself. Reference [21] presents a first rigorous (and rather technical) proof of these assumptions in a special model, namely the 6V model with uniform weights (at the ice point) and a generalization of DWBC. In this chapter we assume that (i) holds and expose general considerations on these assumptions using a *variational approach*. In the first two sections, we use (i) to give compelling arguments for (ii). In the remaining sections, we verify the validity of a natural extension of (i) and explore its consequences. Let us start by giving an overview of the main ideas developed in this chapter.

Our approach to justify (ii) may be summarized as follows. In a tangent method set-up, if one assumes that the arctic curve is deterministic and *fixed*, there remains only one random lattice path, travelling in a deterministic domain $\mathcal{D}$, bordered by pieces of the arctic curve and by the boundary in which the model is defined (the random path can be thought of as an extra path we introduce by hand). The problem is then reduced to study, in the scaling limit, the probability distribution of a random path that stays inside the domain $\mathcal{D}$ (it cannot cross the arctic curve due to the non-crossing property of the paths underlying the arctic curve), the distribution being induced by the weights associated to the elementary steps, in finite number. While the endpoint is fixed, the starting point can be varied and taken away from the arctic curve. The distribution over random paths can be complicated on the lattice but dramatically simplifies in the scaling limit. By using a variational approach, we write the measure on continuous trajectories in terms of a classical action, and find that, in the scaling limit, this measure condensates on the unique trajectory which maximizes the action. We show that the unicity of the maximizing trajectory as well as its complete and explicit description follow from a central observation: the Lagrangean function $L$ defining the classical action is a strictly concave function, a property which we prove to hold for any directed random walk (on $\mathbb{Z}^2$ for simplicity). This unique trajectory is found to satisfy the tangency assumption (ii).

In the second half of this chapter, we strengthen assumption (i) and also make it more concrete by observing a *factorization property*: if $Z_{n+1}^{\mathrm{ref}}$ denotes a 1-refined partition function of a system of $n+1$ non-crossing paths, with the starting point of the most external path (possibly) displaced, then in the large $n$ limit, it factorizes as $Z_{n+1}^{\mathrm{ref}} \simeq Z_n Z_1^{\mathrm{out}}$. In this expression, $Z_n$ is the partition function for $n$ unmodified paths and $Z_1^{\mathrm{out}}$ is the contribution of the most external path, constrained to stay inside a *fixed* domain $\mathcal{D}$ whose exact shape depends on the arctic curve induced by the $n$ other paths. Moreover, if the shape of the arctic curve is known, the variational approach predicts the asymptotic value of $Z_1^{\mathrm{out}}$, which is fully computable in terms of the Langrangean $L$. The factorization property can be extended to provide an efficient way to conjecture the asymptotics of *multirefined* partition functions. We present detailed verifications of this hypothesis in the Aztec diamond and for alternating sign matrices by using exact lattice results. In the last section, we reverse the argument and reformulate the tangent method in a way that no longer requires an extension of the domain, and which reveals the hidden role of the $L$ function.

## 5.1 Single free paths

The general framework which we consider is that of lattice statistical models which can be formulated entirely in terms of non-crossing random paths, and showing the phenomenon of arctic curve. In addition, and this assumption is crucial, the lattice paths can be viewed as *directed* random walks in the plane, which is always the case for models exhibiting an arctic phenomenon, to the best of our knowledge. The paths are defined in terms of a set $S = \{\vec{s}_1 = (u_1, v_1), \vec{s}_2 = (u_2, v_2), \ldots, \vec{s}_k = (u_k, v_k)\}$ of elementary discrete steps and a set $W = \{w_1, w_2, \ldots, w_k\}$ of associated weights $w_i > 0$. We take $k \geqslant 2$ as the case $k = 1$ is trivial; for the same reason we also suppose that not all the $\vec{s}_i$ are collinear; likewise we may assume without loss of generality that one of the $w_i$

is equal to 1. For convenience we assume that the steps belong to $\mathbb{Z}^2$, although this should not be a strong restriction.

By a directed walk, we mean a walk for which a return to a previously visited site is not permitted by its elementary step content. Thus, the available steps are such that the condition $\sum_{i=1}^{k} a_i \vec{s}_i = 0$ with $a_i$ non-negative integers has the only solution $a_i = 0$ for all $i$. Equivalently the convex hull of the points $\vec{s}_i$ in the plane does not contain the origin[1]. This implies the existence of a straight line passing through the origin, which separates the plane in two halves, one of which strictly contains $S$, viewed as a set of points. Then a unit vector $\hat{d} \in \mathbb{R}^2$ exists such that all step vectors have scalar products with $\hat{d}$ which are strictly positive, $\vec{s}_i \cdot \hat{d} > 0$ for all $i$. For simplicity, we may assume that $S$ is such that the possible paths are contained in a cone with an aperture strictly smaller than $\pi$ which is itself contained in the right half-plane $x \geqslant 0$ (thus $u_i \geqslant 0$ for all $i$). We formulate here a number of preliminary results.

This situation is fairly general, but the tangent method is believed to be applicable even more generally, such as for area-weighted paths [23, 25] or for vertex-weighted paths, as seen in the previous chapter. In [159], we generalize the argument given below to the area-weighted case. The argument can also be applied case-by-case for vertex-weighted paths using their Lagrangean function, which can be computed as in Appendix D.

### 5.1.1 Lagrangean function of a path

For the directed walk defined in terms of a set $S$ of possible steps $\vec{s}_i$ and weights $w_i$, we denote by $Z_{r,s}$ the weighted sum over all paths between the origin $(0, 0)$ and the site $(r, s)$, the weight of a path being the product of the weights of its elementary steps. The assumptions made above imply $r \geqslant 0$. The generating function of the numbers $Z_{r,s}$, for elementary steps $\vec{s}_i = (u_i, v_i)$ of weights $w_i$, is given by

$$G(x, y) = \sum_{r,s} Z_{r,s}\, x^r\, y^s = \frac{1}{1-P(x, y)}, \quad P(x, y) = \sum_{i=1}^{k} w_i\, x^{u_i}\, y^{v_i}. \tag{5.1}$$

[1]The previous argument shows that the origin is not in the rational convex hull. It is not difficult to show that it can neither be in the real convex hull.

Note that this general result yields as a particular case the generating function (1.45) already computed in Chapter 1. We will be mostly interested in the asymptotic behaviour of $Z_{r,s}$, when $r, s$ are both large with the ratio $s/r$ finite, and more specifically in its exponential growth rate. In what follows we use the notation $Z_{na,nb}$ with $a, b \in \mathbb{R}$ and $n \in \mathbb{N}$ defined to be $Z_{r,s}$ with $(r, s)$ one of the closest point to $(na, nb)$ *reachable* by the elementary steps $S$.

By a standard argument, we note that for $m, n$ positive integers, the inequality $Z_{(m+n)a,(m+n)b} \geqslant Z_{ma,mb} Z_{na,nb}$ implies that the sequence $c_n = \log Z_{na,nb}$ is superadditive (i.e. satisfies $c_{m+n} \geqslant c_m + c_n$). By Fekete's lemma, the following limit exists

$$\lim_{n\to\infty} \frac{1}{n} \log Z_{na,nb} = \sup_n \frac{1}{n} \log Z_{na,nb} \equiv L(a, b), \tag{5.2}$$

which we call the Lagrangean function for reasons that will be clear later on. It may also be called a large deviation function. The lemma does not guarantee that the limit is finite, but the following simple argument shows that it is finite, for all $a, b$, and grows at most linearly in $a$ and $b$.

Let $c = \min_{1\leqslant i\leqslant k}(\vec{s}_i \cdot \hat{d})$, a strictly positive number. To reach the point $(na, nb)$ requires at most $(na, nb) \cdot \hat{d}/c$ elementary steps, taken from the set $S$. We readily obtain the upper bound

$$Z_{na,nb} \leqslant \big(\sum_i w_i\big)^{(na,nb)\cdot\hat{d}/c} \quad \text{and} \quad L(a,b) \leqslant \frac{(a,b)\cdot\hat{d}}{c} \log\big(\sum_i w_i\big). \tag{5.3}$$

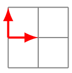
Unless the points $(na, nb)$ are reachable for no $n$, the function $L(a, b) \geqslant 0$ is also bounded below by any non-zero term in the sequence $\frac{1}{n}\log Z_{na,nb}$.

Repeating the above arguments in terms $n' = na$ and final point $(n', n'\frac{b}{a})$, we obtain the obvious but important scaling relation,

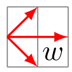
$$L(a, b) = a\, L(1, \tfrac{b}{a}) \equiv aL(\tfrac{b}{a}). \tag{5.4}$$

In the rest of this chapter, the $L$ function will be used with a single argument, meant to be a slope. The support of the function of one variable $L(t)$ depends on the walk considered; it is given by the interval $[t_{\min}, t_{\max}]$ where $t_{\min}$, $t_{\max}$ are respectively the minimal and maximal slopes among the elementary available steps. Our assumptions imply that at least one of $t_{\min}$, $t_{\max}$ is finite, so the support is a finite or semi-infinite interval.

In simple enough cases, the function $L(t)$ can be explicitly computed from the asymptotic form of the coefficients of the bivariate generating function $G(x,y)$, see for instance the review [161] (see below for more details). As an illustration, we find the following Lagrangean functions for two simple walks. All weights $w_i$ are equal to 1 in these, except the horizontal step $(2,0)$ in the second example, which has weight $w$:



$$\begin{aligned} &P_1(x,y) = x + y \\ &L_1(t) = (1+t)\log(1+t) - t\log t \qquad t \geqslant 0, \end{aligned} \tag{5.5}$$




$$\begin{aligned} &P_2(x,y) = xy + x/y + wx^2 \\ &L_2(t) = \log\tfrac{\sqrt{1+wt^2}+\sqrt{w+1}}{\sqrt{1-t^2}} + t\log\tfrac{\sqrt{1+wt^2}-\sqrt{w+1}\,t}{\sqrt{1-t^2}} \qquad -1 \leqslant t \leqslant 1. \end{aligned} \tag{5.6}$$

**Concavity.** For $(nx, ny)$ an intermediate passage point[2] between $(0,0)$ and $(na, nb)$, the simple inequality (which can be made strict except in degenerate cases)

$$Z_{na,nb} \geqslant Z_{nx,ny}\, Z_{n(a-x),n(b-y)}, \tag{5.7}$$

immediately yields $L(a,b) \geqslant L(x,y) + L(a-x, b-y)$ (which cannot be made strict in general, as we will see later). Using the scaling relation (5.4), it becomes

$$L\left(\frac{b}{a}\right) \geqslant \frac{x}{a} L\left(\frac{y}{x}\right) + \left(1 - \frac{x}{a}\right) L\left(\frac{b-y}{a-x}\right). \tag{5.8}$$

Setting $\alpha = \frac{x}{a}$, $t_1 = \frac{y}{x}$ and $t_2 = \frac{b-y}{a-x}$, the previous inequality is the statement that the function $L(t)$ is concave on its domain,

$$L\big(\alpha t_1 + (1-\alpha)t_2\big) \geqslant \alpha L(t_1) + (1-\alpha)L(t_2), \qquad 0 \leqslant \alpha \leqslant 1. \tag{5.9}$$

However, the function $L(t)$ satisfies a stronger property which turns out to be central in the subsequent arguments: *the Lagrangean function $L(t)$ is strictly concave.* The strict concavity means that

$$L\big(\alpha t_1 + (1-\alpha)t_2\big) > \alpha L(t_1) + (1-\alpha)L(t_2), \quad 0<\alpha<1,\ t_1 \neq t_2. \tag{5.10}$$

[2] Notice that we use $x$ and $y$, either as arguments of generating functions, or as rescaled coordinates. The meaning should be clear depending on the context.

The yet slightly stronger property, namely that $L''(t) < 0$ on its open support, which itself ensures the strict concavity, proves to be useful later on and can be obtained from the way the function $L(t)$ is to be computed. The reference [161] reviews a number of methods to compute the asymptotic value of the coefficients of multivariate generating functions. In our case, the relevant function is the two-variable generating function given in (5.1), for which the results quoted in [161] apply[3].

Let $x, y$ be two functions of $t$ satisfying the system (the indices of $P(x, y)$ denote partial derivatives)

$$P(x, y) = 1, \qquad\qquad tx\, P_x(x, y) = y\, P_y(x, y). \tag{5.11}$$

Theorem 1.3 and Corollary 3.21 of [161] state that (a) there is a unique solution for which $x(t), y(t)$ are real positive functions, and (b) the asymptotic value of $Z_{na,nb}$ is given by

$$Z_{na,nb} \sim \sqrt{\frac{N(t)}{2\pi na}}\, \exp\bigl\{naL(t)\bigr\}, \qquad t = \frac{b}{a}, \tag{5.12}$$

where we used the notation $f_n \sim g_n \Leftrightarrow \lim_{n\to\infty} f_n/g_n = 1$. Here $N(t)$ is an explicit function of $x(t)$ and $y(t)$, generally a rather complicated algebraic function of $t$, and

$$L(t) = -\log x(t) - t\, \log y(t). \tag{5.13}$$

In the two examples mentioned previously, these functions are found to be equal to

$$x_1(t) = \frac{1}{1+t}, \quad y_1(t) = \frac{t}{1+t}, \quad N_1(t) = \frac{1+t}{t}, \quad t > 0, \tag{5.14}$$

$$x_2(t) = \frac{\sqrt{1+w} - \sqrt{1+wt^2}}{w\sqrt{1-t^2}}, \quad y_2(t) = \frac{t\sqrt{1+w} + \sqrt{1+wt^2}}{\sqrt{1-t^2}}, \tag{5.15}$$

$$N_2(t) = \frac{(\sqrt{w+1} + \sqrt{1+wt^2})^2}{\sqrt{w+1}\,\sqrt{1+wt^2}\,(1-t^2)}, \quad -1 < t < 1.$$

[3] In the survey [161], the generating function is assumed to be non-singular at the origin and to expand in positive powers of $x$ and $y$ ($r, s \geqslant 0$). This assumption is not verified here if some of the elementary steps have a second component which is negative. When this is the case and since the step vectors are contained in a cone of aperture strictly less than $\pi$, they can be mapped by an affine transformation to steps contained in the first quadrant for which the results of [161] apply. However, one can show that this procedure and a naive application of the results of [161] yield exactly the same results. For this to hold, the directedness of the walks considered here is crucial.

The corresponding functions $x(t), y(t), L(t)$ are shown in Figure 5.1. It is a nice exercise to verify that the results derived by this general procedure indeed match the asymptotics of explicit lattice enumeration formulae, such as (1.44).

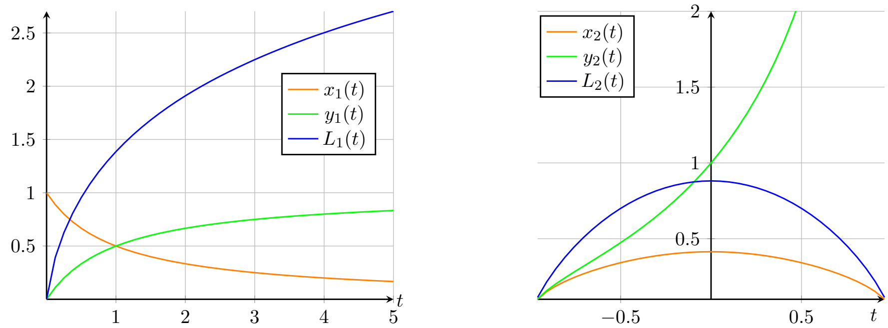



Figure 5.1: Plots of the functions $x(t)$, $y(t)$ and $L(t)$, for the two examples (5.5) and (5.6), first example on the left, second one, with $w = 1$, on the right. The plots show the typical behaviour of these functions for semi-infinite and finite domains respectively.

Before proving that $L''(t)$ is strictly negative, we show that the three functions $x(t), y(t), L(t)$ are smooth. The first equation of (5.11) yields a first implicit equation $f_1(x, y) = 0$, where $\partial_x f_1(x, y)$ nowhere vanishes in the domain $x, y > 0$. Moreover, since $f_1(x, y)$ is $C^\infty$ in both variables in the same domain, the implicit function theorem implies that $x = x(y)$ is a $C^\infty$ function of $y$. Similarly the second equation in (5.11), after substituting $x = x(y)$, yields an implicit equation $f_2(y, t) = 0$, which is $C^\infty$ in both $t$ and $y > 0$. The derivative $\partial_y f_2(y, t)$ is found to be equal to

$$\begin{aligned}\partial_y f_2(y,t) = &- \frac{t^2 x}{y}\big[P_x(x,y) + xP_{xx}(x,y)\big] + 2txP_{xy}(x,y) \\ &- P_y(x,y) - yP_{yy}(x,y),\end{aligned} \tag{5.16}$$

where we used the identity

$$x'(y) = -\frac{tx}{y}, \tag{5.17}$$

obtained by differentiating the first equation of (5.11) with respect to $y$, and using the second equation (here and in what follows, $f'$ denotes the derivative of $f$ with respect to its argument, made explicit). Recalling the form of the function

$$P(x,y)=\sum_{i=1}^{k} w_i\, x^{u_i}\, y^{v_i}, \qquad u_i \geqslant 0, \tag{5.18}$$

a simple computation shows that

$$\begin{aligned}&-\frac{t^2 x}{y}\big[P_x(x,y)+xP_{xx}(x,y)\big]+2txP_{xy}(x,y)-P_y(x,y)-yP_{yy}(x,y)\\&=-\sum_{i=1}^{k} w_i\,(v_i-tu_i)^2\, x^{u_i}\, y^{v_i-1}\end{aligned} \tag{5.19}$$

is strictly negative. Thus, $\partial_y f_2(y,t)<0$ nowhere vanishes in the domain $y>0$, $t\in\mathbb{R}$. A second use of the implicit function theorem implies that $y=y(t)$ is a smooth function of $t$ on its open domain $]t_{\min},t_{\max}[$. The statement that the functions $x(t)$ and $L(t)$ are smooth readily follows. We note that the double use of the implicit function theorem readily proves the uniqueness of the real functions $x(t),y(t)>0$.

In order to check that $L''(t)<0$, we differentiate the first equation of (5.11) with respect to $t$ (after the substitution $x=x(t)$ and $y=y(t)$) and use the second equation to get the condition $(\log x(t))'=-t\,(\log y(t))'$. It allows to write the first two derivatives of $L(t)$ as

$$L'(t)=-\log y(t), \qquad L''(t)=-\frac{y'(t)}{y(t)}. \tag{5.20}$$

Differentiating now the second equation of (5.11) with respect to $t$ and using (5.17) once more yields

$$y'(t)=\frac{xP_x(x,y)}{P_y(x,y)+yP_{yy}(x,y)+\frac{t^2x}{y}\big[P_x(x,y)+xP_{xx}(x,y)\big]-2txP_{xy}(x,y)}. \tag{5.21}$$

From the specific form of $P(x,y)$, we see that $xP_x(x,y)$ is strictly positive in the domain $x,y>0$, whereas the denominator, equal to $-\partial_y f_2(y,t)$ computed above, is also strictly positive. Thus, the numerator and denominator of (5.21) are both strictly positive for $x,y>0$, which implies $y'(t)>0$ and thus $L''(t)<0$ as claimed.

### 5.1.2 Passage probability distributions

In this subsection, we prove that in the scaling limit, by which all large lattice distances $na, nb, \dots$ are rescaled to $a, b, \dots$ (equivalently, the lattice spacing is shrunk from 1 to $\frac{1}{n}$), the probability measure on paths connecting two points concentrates on a unique trajectory, namely a straight line. We also investigate the size of the fluctuations around the straight line. The collapse onto a unique trajectory is eventually re-derived in a variational framework.

**Passage probabilities.** Given a distant site $(na, nb)$, the probability that a directed walk passes through the point $(nx, ny)$ is given by

$$\begin{aligned}\mathbb{P}_0\big[(nx,ny)\big|(na,nb)\big] &= \frac{Z_{nx,ny}\, Z_{n(a-x),n(b-y)}}{Z_{na,nb}} \\ &\approx \exp\left\{an\left[\frac{x}{a}L\Big(\frac{y}{x}\Big) + \Big(1-\frac{x}{a}\Big)\, L\Big(\frac{b-y}{a-x}\Big) - L\Big(\frac{b}{a}\Big)\right]\right\},\end{aligned} \tag{5.22}$$

up to the multiplicative non-exponential terms given in (5.12). The strict concavity of the Lagrangean function implies that the combination in the square brackets is strictly negative, except when the two points $\frac{y}{x}$ and $\frac{b-y}{a-x}$ coincide, namely when $\frac{y}{x} = \frac{b}{a}$, in which case it trivially vanishes. We therefore obtain that the above probability is exponentially depreciated with $n$ when the intermediate point is not on the straight line. An analogous though somewhat weaker result has been proved in [22].

**Fluctuations.** We have seen that in the rescaled domain, the path between $(0,0)$ and $(a,b)$ is almost surely a straight line. However, at finite but large $n$, the distribution shows fluctuations. It is natural and important for what follows to understand the size of these fluctuations.

In the asymptotic regime, the passage probability is given by

$$\mathbb{P}_0\big[(nx,ny)\big|(na,nb)\big] \sim \sqrt{\frac{aN(\frac{y}{x})N(\frac{b-y}{a-x})}{2\pi nx(a-x)N(\frac{b}{a})}} \times \exp\{anF(y)\}, \tag{5.23}$$

where $F(y)$ is the combination of three terms in the square brackets in (5.22). As noted above, $F(y)$ is strictly negative except at $y(x) = tx =$

$\frac{bx}{a}$ where it vanishes. Because of the factor $n$ in the exponential, only the values of $y$ very close to $y(x)$ will make a significant contribution to the probability. Therefore, we can expand, to leading order, $F(y)$ and the factor in front of the exponential around $y(x)$. From above we know that $F(tx) = 0$ but also that $F'(tx) = 0$ since $F(y)$ is strictly negative away from $y = tx$, which is therefore a maximum. Moreover, a simple computation yields $F''(tx) = \frac{1}{x(a-x)} L''(t) < 0$, a strictly negative number according to the results of the previous subsection. We therefore obtain the following Gaussian approximation,

$$\mathbb{P}_0\big[(nx, ny)\big|(na, nb)\big] \simeq \frac{1}{n}\sqrt{\frac{1}{2\pi\sigma^2}}\sqrt{\frac{N(t)}{-L''(t)}}\exp\left\{-\frac{(y-tx)^2}{2\sigma^2}\right\}, \tag{5.24}$$

for which we allow $y$ to take all real values (rather than those related to the reachability to and from $(nx, ny)$), and where

$$\sigma = \sqrt{\frac{x(a-x)}{-anL''(t)}}. \tag{5.25}$$

The above probability is not normalized upon summation over $y$, since one readily obtains

$$\begin{aligned}\sum_{ny=-xt_{\min}}^{ny=-xt_{\max}} \mathbb{P}_0\big[(nx, ny)\big|(na, nb)\big] &\sim \int_{-\infty}^{\infty} \mathrm{d}(ny)\mathbb{P}_0\big[(nx, ny)\big|(na, nb)\big] \\ &= \sqrt{\frac{N(t)}{-L''(t)}}.\end{aligned} \tag{5.26}$$

On general grounds, we do not expect it to be normalized. Indeed, if a vertical step of the form $\vec{s} = (0, v)$ is allowed (has non-zero weight), the walk may visit several values of $y$ for the same value of $x$, causing an overcounting (integral larger than 1). Likewise, if some of the possible steps have horizontal components $u \geqslant 2$, a fraction of walks will not visit any site on the vertical line at $x$, causing this time an undercounting (integral smaller than 1). Finally, there may be arithmetical constraints on the sites $(x, y)$ that are actually reachable, in which case the summation over $y$ must be further restricted. These various situations are reflected by the presence of the factor $\sqrt{N(t)/(-L''(t))}$ in (5.24).

Despite the (expected) absence of a proper normalization, the distribution (5.24) shows a dispersion of order $n^{-1/2}$ around the mean value

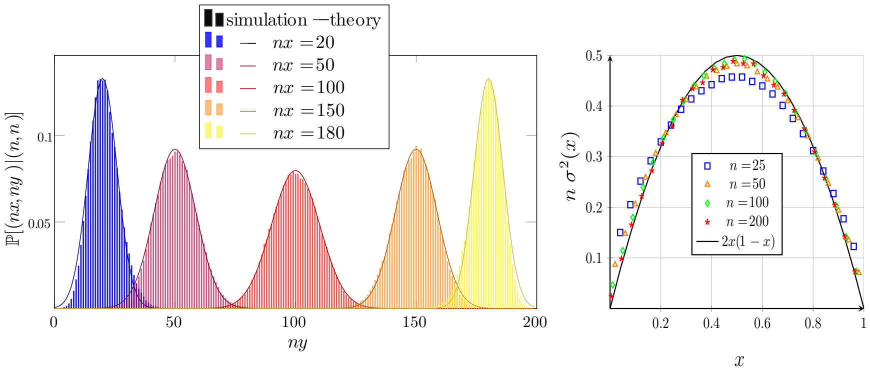



Figure 5.2: *Left:* Probability distributions of a passage at $y$ for five different values of $x$ ($a = b = 1$), for the directed paths defined by $S = \{(1,0),(0,1)\}$, each step having weight $w = 1$ (example 1 in Section 5.1). The bars are the results of simulations (36 000 configurations) for $n = 200$, whereas the solid lines represent the Gaussian approximation (5.24). *Right:* Comparison, for the same paths but for different values of $n$, of the quantity $n\sigma^2$ obtained from simulations with the one computed in (5.25) ($a = b = 1$, $t = 1$ and from (5.5), $L_1''(1) = -\frac{1}{2}$).

$y = tx$, in the rescaled domain. On the lattice, keeping the lattice variables $nx, ny, na, nb$, we expect a standard deviation of order $n^{1/2}$. This is well confirmed by numerical simulations, as shown in Figure 5.2 for the lattice paths of example 1 given in (5.5).

**Variational approach.** We now reformulate the problem in a continuous (variational) framework. Rather than counting the possible paths within a certain class dictated by the choice of possible elementary steps, and then doing statistical estimates about passage probabilities, we associate, in the continuum limit, a classical action to each path. The action allows to assign a statistical weight to a whole trajectory, represented as a continuous function $h(x)$, the weight being itself inherited from those of the lattice elementary steps. The resulting variational problem appears to be the appropriate tool to determine the extremal paths used in the tangent method to characterize arctic curves.

Let us consider all possible paths from (0,0) to $(na, nb)$, made of elementary steps taken from the set $S$. For $n$ large, their weighted sum $Z_{na,nb}$ has the asymptotic value given in (5.12). Let us now focus on all the paths which pass through $K+1$ intermediate points $\vec{x}_j$, with $\vec{x}_j = (\frac{ja}{K}n, h_j n)$ for $j = 0, \ldots, K$. Here the $h_j$ are fixed numbers independent of $n$ satisfying $h_0 = 0$ and $h_K = b$. Provided $\frac{n}{K}$ is large, the weighted sum $Z_n[\{h_j\}]$ over such paths is

$$Z_n[\{h_j\}] \approx \prod_{j=0}^{K-1} Z_{(\frac{an}{K},(h_{j+1}-h_j)n)} \approx \sqrt{\prod_{j=0}^{K-1} \frac{KN(t_j)}{2\pi na}} \exp\left\{\frac{an}{K}\sum_{j=0}^{K-1} L(t_j)\right\}, \tag{5.27}$$

where $t_j = \frac{(h_{j+1}-h_j)K}{a}$. Assuming that $h_j$ is the value at $\frac{ja}{K}$ of some real function[4] $h$, namely $h(\frac{ja}{K}) = h_j$, we obtain, for large enough $K$, that $t_j \simeq h'(\frac{ja}{K})$, and that the sum over $j$ in the exponential is well approximated by an integral, leading to

$$Z_n[h] \approx \sqrt{\prod_{j=0}^{K-1} \frac{KN(t_j)}{2\pi na}} \times \exp\left(n \int_0^a \mathrm{d}x \, L\big(h'(x)\big)\right). \tag{5.28}$$

The function $h$ needs not be everywhere differentiable. Looking at the way it was introduced, the mere existence of a right derivative is sufficient. More generally, we may consider the set of functions $h$ which have a left and a right derivative, hence piecewise of class $\mathcal{C}^1$.

The other dependence of $Z_n[h]$ in the trajectory is in the factor in front of the exponential, which can similarly be written as

$$\sqrt{\prod_{j=0}^{K-1} \frac{KN(t_j)}{2\pi na}} \simeq \exp\left(\frac{K}{2a}\int_0^a \mathrm{d}x \, \log \frac{KN\big(h'(x)\big)}{2\pi na}\right). \tag{5.29}$$

Since $\frac{K}{n}$ was assumed to be small (we will argue that choosing $K$ proportional to $\sqrt{n}$ is good enough), this term is subdominant, as expected.

We obtain that the fraction of paths going from the origin to the point $(na, nb)$, and which collapse in the scaling limit to a continuous trajectory described by the function $h(x)$, for $x \in [0, a]$, $h(0) = 0$ and $h(a) = b$, has a relative statistical weight given in terms of an action $S[h]$ associated with $h$,

$$Z_n[h] \approx \mathrm{e}^{nS[h]}, \qquad S[h] = \int_0^a \mathrm{d}x \, L\big(h'(x)\big). \tag{5.30}$$

[4]This is largely legitimate because the paths we consider are directed to the right. In case vertical steps are possible, a vertical macroscopic stretch is unlikely.

In other words, $Z_n[h]$ can be interpreted as the unnormalized probability that the path stays in the scaling neighbourhood of the graph of $h(x)$. Interestingly the weight of the paths depends on the underlying lattice walk only through the function $L$ (which was called Lagrangean for obvious reasons). A trivial but important remark is that not every function $h$ is permitted: it must be compatible with the way the discrete walk has been defined, implying that $h'(x)$ must be in the interval $[t_{\min}, t_{\max}]$ for all $x$.

Because the Lagrangean function depends on the derivative of $h$ only, the Euler-Lagrange equation, computed on the subdomains where $h$ is $\mathcal{C}^1$, takes the simple form

$$L''(h')\,h''(x) = 0, \qquad h(0) = 0,\, h(a) = b. \tag{5.31}$$

As $L''$ is strictly negative on its (open) domain, the trajectories $h(x)$ which make the action extremal satisfy $h''(x) = 0$, that is, $h(x)$ is piecewise rectilinear. The concavity of $L$ ensures that the straight line $h(x) = bx/a$ is the unique global maximum of $S[h]$, as proved below. In the limit $n \to \infty$, the straight line is therefore observed almost surely, consistently with the results obtained previously.

## 5.2 Tangency assumption

We now consider a typical tangent method set-up, see Figure 5.3. The argument that follows is applicable whether or not the original domain has been extended. We also focus *solely* on the frozen region that is *empty*, except for one path. Assuming the hypothesis (i), we view the compact cluster of paths collapsing to the arctic curve $\mathcal{C}$ in the scaling limit, as being given. It influences the shape and the distribution of the outermost path because it acts as an impenetrable frontier, due to the non-crossing property built in the model. The problem for the outermost path can then be posed in the following terms: *what is the distribution on the lattice paths from $b_0'$ to $b_1$ (see notations of Figure 5.3), made of elementary steps from the set $S$, and constrained not to cross a certain curve $\mathcal{C}$ ? In particular, in the scaling limit, does this distribution condensate on a deterministic trajectory with the properties given in (ii) ?*

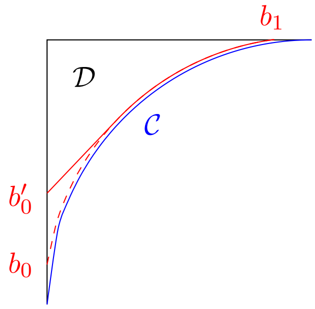



Figure 5.3: Sketch of a typical tangent method set-up in the scaling limit. The modified path, in red, has its starting point moved from $b_0$ to $b_0'$. It is constrained to stay in the domain $\mathcal{D}$ composed of portions of arctic curve $\mathcal{C}$ in blue and the boundaries of the (possibly extended) domain in black.

Compared to the results obtained in the previous section, the novel issue is of course the presence of the impenetrable curve $\mathcal{C}$ because in situations of interest, the endpoint $b_1$ cannot be reached from $b_0'$ by following a straight line, as it would cross $\mathcal{C}$. As the arctic curve is only defined in the continuum, we formulate the above question in the continuum. To do this we take advantage of the variational approach developed in the previous section, in which we assigned the action $S[h]$ to a trajectory $h$ between two points in the plane. For the present problem, we use the same action (5.30) as when there was no constraint at all, and restrict it to those $h$ that do not cross the given curve; a justification for keeping the same action is provided in Subsection 5.2.2.

### 5.2.1 Variational approach in a constrained domain

The following setting should be general enough. As we did before, we consider a trajectory in the plane as the graph of a continuous function $h(x)$. The general problem is, given a simply-connected domain $\mathcal{D}$, whose boundary is sufficiently smooth (say, it has a tangent everywhere), to find the trajectories $h$ from $(0,0)$ to $(a,b)$, wholly contained in $\mathcal{D}$, which maximize the action

$$S[h]=\int_0^a \mathrm{d}x L\big(h'(x)\big), \quad h(0)=0,\ h(a)=b,\ \big(x,h(x)\big)\in\mathcal{D}\ \forall x\in[0,a]. \tag{5.32}$$

We will find that this maximum is unique, let us call it $h^*$. The relative probability that the path follows any other trajectory $h$, is equal to $Z[h]/Z[h^*] \approx \mathrm{e}^{n(S[h]-S[h^*])}$. It goes to 0 exponentially fast with $n$; in the scaling limit, the single path follows the trajectory $h^*$ almost surely.

We recall that $L$ is the Lagrangean function, strictly concave, associated to the directed walk under consideration, and computable in terms of the sets $S$ and $W$. Also the function $h$ must have its (left- and right-) derivatives in the domain of $L$, namely the interval $[t_{\min}, t_{\max}]$, itself determined by the set $S$. This imposes certain constraints on $\mathcal{D}$ for the point $(a, b)$ to be reachable. The last two conditions are implicit in the rest of this section.

Our main result is the following

**Theorem 5.2.1.** *For the variational problem formulated above (connected and simply-connected domain $\mathcal{D}$, boundary $\partial\mathcal{D}$ of class $\mathcal{C}^1$), the action $S[h]$ given in (5.32) with $L$ strictly concave, has a global maximum in the set of continuous piecewise $\mathcal{C}^1$ functions, attained for a unique function $h^*$ of class $\mathcal{C}^1$. The maximizing $h^*$ is universal in the sense that it does not depend on the specific form of $L$. The function $h^*$ is best described as the shortest path in $\mathcal{D}$ between the points $(0, 0)$ and $(a, b)$: the graph of $h^*(x)$ has rectilinear pieces alternating with portions of the boundary of $\mathcal{D}$; moreover the linear parts of $h^*$ meet or exit the boundary $\partial\mathcal{D}$ tangentially.*

The proof relies in an essential way on the strict concavity of the Lagrangean function and on the simple form of the action $S[h]$, which depends on $h$ through its derivative only. It basically is a consequence the following lemma.

**Lemma 5.2.2.** *Fix $(x_i, y_i)$ and $(x_f, y_f)$ two points in the plane and consider the set of continuous and piecewise $\mathcal{C}^1$ functions connecting these two points, so that $h(x_i) = y_i$ and $h(x_f) = y_f$ (the functions $h$ are not restricted to be contained in a certain domain). Then the action $S[h] = \int_{x_i}^{x_f} \mathrm{d}x\, L\big(h'(x)\big)$ has a unique maximum for $h^*$ given by the straight line.*

*Proof.* Let $t^* = \frac{y_f - y_i}{x_f - x_i}$ be the slope of the straight line $h^*$. Assume first that $h$ is of class $\mathcal{C}^1$. In this case, the derivative $h'(x)$ for $x \in [x_i, x_f]$ takes all values in a finite interval $[h'_{\min}, h'_{\max}]$. Since the average value of $h'(x)$ over $[x_i, x_f]$ is equal to $t^*$, the interval $[h'_{\min}, h'_{\max}]$ must contain $t^*$. The function $L$ being strictly concave, its graph lies below that of any of its tangents, in particular it lies below its tangent at $t^*$, namely $L(h') < L'(t^*)(h' - t^*) + L(t^*)$ for all $h' \neq t^*$. If the interval $[h'_{\min}, h'_{\max}]$ is not reduced to the singleton $\{t^*\}$, that is, if $h$ is not the straight line $h^*$, then

$$S[h] < L'(t^*)(y_f - y_i) + (x_f - x_i)\big(L(t^*) - t^* L'(t^*)\big) = (x_f - x_i)L(t^*) = S[h^*]. \tag{5.33}$$

In case $h$ is only piecewise $\mathcal{C}^1$, the image of $[x_i, x_f]$ under $h'$ is the union of intervals (when $h$ is not linear) and singletons (when $h$ is linear). The inequality $L(h') < L'(t^*)(h' - t^*) + L(t^*)$ used above nonetheless remains valid for any value $h'$ in the domain of $L$, and leads to the same conclusion $S[h] < S[h^*]$. ■

*Proof.* To prove Theorem 5.2.1, we take $(0, 0)$ and $(a, b)$ to be the opposite vertices of a parallelogram (with slopes $t_{\min}, t_{\max}$), in which the allowed region is delimited by the boundary $\partial\mathcal{D}$, see Figure 5.4-left (as we can see, the $\mathcal{C}^1$ property is only required for certain portions of $\partial\mathcal{D}$). We repeatedly use Lemma 5.2.2: to go from point $A$ to point $B$, the optimal path is the straight line, whenever this is possible (if the line does not cross the boundary). In particular, if $(a, b)$ is reachable from $(0, 0)$ on a straight line, this will be the optimal $h^*$.

If not, the endpoint $(a, b)$, seen from the origin, is hidden by a portion of $\partial\mathcal{D}$. In this case, $h^*$ is a straight line starting from the origin that will necessarily make a first contact with $\partial\mathcal{D}$, say at $P_1$: it cannot go round $\partial\mathcal{D}$ without touching it because such a path could be optimized by following a shorter chord. Thus, $h^*$ must be composed of straight lines and portions of $\partial\mathcal{D}$. $P_1$ must be the farthest point of the boundary that is reachable by a straight line from the origin hence must be a tangential point. Indeed, the Lemma implies that the union of a straight line to another contact point, and the portion of $\partial\mathcal{D}$ from it to the tangential contact point, has a smaller action.

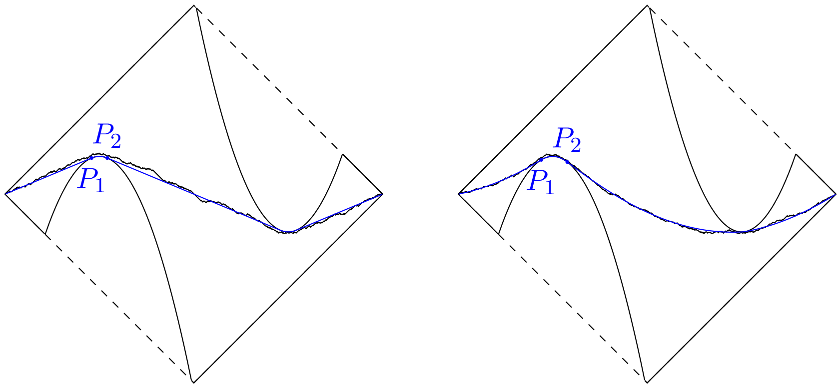



Figure 5.4: Numerical simulation of a single lattice path with elementary steps in $S = \{(1,1),(1,-1),(2,0)\}$ and a weight $w$ for the step $(2,0)$. The path starts from $(0,0)$ and ends at $(n,0)$ with $n = 960$, and is constrained to stay in the domain delimited by the solid curve. *Left:* The path is sampled with a weight $w = 0.5$. *Right:* We set $w = 1$ and in addition, the path is weighted by a factor $q^A$, with $A$ the area below it. The simulation is done using $q = (0.05)^{1/n}$. The blue curves are the corresponding optimal trajectories $h^*(x)$, made of portions of the boundary and free trajectories (straight lines on the left or known curves [159] on the right).

From $P_1$, the optimal path will follow $\partial\mathcal{D}$ for a while, exit the boundary at a certain point $P_2$ and follow a straight line towards some point $X$ (it could potentially follow the boundary till it reaches the endpoint, like for the Aztec diamond). $P_2$ is either a tangential point or else lies further down the boundary. Again, the Lemma shows that a tangential point is optimal. Iterating these arguments proves the claim. It should be noted that the successive exit and passage points $P_2, X, \ldots$ cannot be precisely determined without knowing the shape of the domain that the paths will encounter further ahead. These points can only be determined from the global shape of $\mathcal{D}$. The resulting global process leads to the optimal path as depicted in Figure 5.4-left.

The simply-connected assumption, usually satisfied in arctic curve situations, prevents the possibility to have two or more equally maximizing paths, ensuring the unicity of $h^*$. Indeed, let $h_1$ and $h_2$ be two distinct global maxima ($h_1 \neq h_2$) of the constrained problem, thus satisfying $S[h_1] = S[h_2]$. Then $h_3 = \alpha h_1 + (1-\alpha)h_2$ lies between $h_1$ and $h_2$ for $0 < \alpha < 1$. Since the domain is simply-connected and since $h_1$ and $h_2$ belong to it, $h_3$ is in the domain as well. It is therefore a valid trajectory and, since $L$ is strictly concave, satisfies $S[h_3] > \alpha S[h_1] + (1-\alpha)S[h_2] = S[h_1] = S[h_2]$, which contradicts the assumption. ■

The variational approach can also be applied in the area-weighted case, in which a path is attributed an additional weight $q^A$, with $A$ the area below this path. A non-trivial scaling limit is reached when $q = \mathfrak{q}^{\frac{1}{n}}$, with $\mathfrak{q} > 0$, in which case the Euler-Lagrange equation generally leads to curves rather than straight lines. Their precise shapes depend on the nature of $L(t)$ and hence on the elementary steps $S$ and their weights $W$. The tangency property also holds in the area weighted case (see Figure 5.4-right). The proof is then slightly more involved and is given in [159].

The typical shape of large uniformly weighted Young diagrams mentioned in the introduction is exactly one of the aforementioned curves, computed from $L_1(t)$, where $\lambda = \log(\mathfrak{q})$ is in this case a Langrange multiplier. This limit shape can also be used to compute the exponential growth of the number $p(n)$ of partitions of an integer $n$, namely $p(n) \simeq \mathrm{e}^{\pi\sqrt{2n/3}}$, a result first derived by Hardy and Ramanujan [1].

### 5.2.2 Discussion

We have taken a definite stand on how to treat the paths in presence of an impenetrable boundary, by restricting the variational problem obtained without boundary to trajectories contained in the allowed domain. If the endpoint is not reachable by a straight line, the optimal path will inevitably approach and coincide with portions of the constraining boundary. Since our derivation of the variational principle ignores the presence of a boundary, our approach calls for a justification. We provide it in this subsection, in the case of a *fixed* impenetrable boundary. This is further reinforced by the results of the following three sections, which suggest that the variational approach also holds when the impenetrable boundary is composed of other random paths.

An alternative route would be to take into account, in the discrete lattice formulation itself, the constraint that the paths be contained in the allowed region, by modifying the weights of the elementary steps so as to prevent the walk from entering the forbidden region. In that formulation, one would compute the number of acceptable paths and the asymptotic form of the passage probabilities in order to figure out on what sort of trajectories the distribution concentrates in the scaling limit. Though probably more natural and more satisfactory, this approach is likely to be impractical for general domains as those considered above in the continuum, as it would require to handle position-dependent weights $w_i(\vec{x})$. A (admittedly very) simple example is nonetheless worked out below which shows that the two approaches give identical results in the scaling limit.

The main reason which validates our approach is based on the following argument. If it is true that the optimal path coincides with portions of the boundary in the continuum limit, it is not so on the lattice. The lattice formulation of the problem considers random lattice paths from $(0,0)$ to $(na,nb)$, and asks for the distribution over these paths, for large $n$, equivalently the distribution of passage on any given vertical line. In regions where the constraint plays no role, we know from Subsection 5.1.2 that the distribution on paths has its support essentially contained in a window of size $n^{1/2}$ around the straight line (the average path). When the constraint is effective enough to force the lattice paths to deviate from their ballistic, straight line behaviour, one does not expect that the average path (or the likeliest, it does not make much difference) remains in a close neighbourhood of the constraint, because the entropy away from the constraint is higher: the paths which stay away from the constraint are more numerous. The entropic bias increases with the distance to the constraint and keeps the average path away from it. The average path will however not be pushed away from the constraint beyond a characteristic length scale of order $n^{1/2}$ since this is the scale of the fluctuations when there is no constraint. It is therefore natural to expect that the average path will be localized at a distance of order $n^{1/2}$ from the constraint. In this case, the derivation we made in Subsection 5.1.2 of the action associated with a macroscopic path remains valid if we partition the horizontal interval $[0,na]$ into small cells whose size $\frac{an}{K}$ is of order $n^{1/2}$. Over this horizontal distance, the vertical displacement $(h_{j+1}-h_j)n$ is of order $\frac{an}{K}t_j$, also proportional to $n^{1/2}$. It then follows that the path trajectory is partitioned into windows of linear size $n^{1/2}$,

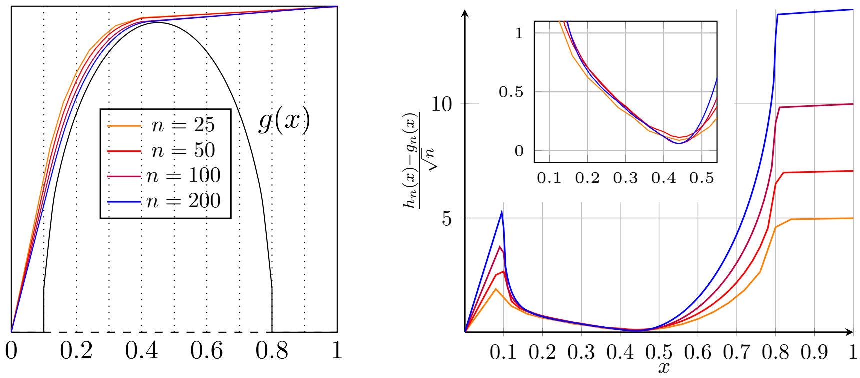



Figure 5.5: Results of numerical simulations for the walks with steps $(1, 0)$ and $(0, 1)$ as explained in the text. *Left:* Average path for various linear sizes $n$ in the rescaled domain. The average is performed using 100 000 random paths for $n = 25, 50$ and using 40 000 for $n = 100, 200$. *Right:* Difference between the average path $h_n$ and the boundary $g_n(x) = ng(x)$ rescaled by a factor $\sqrt{n}$. The inset focuses on the region where the paths stick to the boundary.

big enough to contain the main fluctuations, and small enough to stay clear of the constraint. Inside each window, the path combinatorics remains roughly the same as if there were no constraint. So on the lattice, the constraint has a repulsive effect that moves the windows away from itself, but the statistics inside the windows remains unchanged, suggesting that the Lagrangean function is the same whether or not there is a constraint (this will be confirmed in the explicit example below). In the continuum limit, namely after rescaling by a factor $1/n$, the repulsion scales away and the restriction that the paths do not enter the forbidden region is then all what remains of the constraint.

These expectations are borne out by the results of numerical simulations, shown in Figure 5.5. The simulations were carried out for the walks made of the two elementary steps $(1, 0)$ and $(0, 1)$, each with weight 1 (the example 1 in Section 5.1), conditioned to start at $(0, 0)$ and to end at the diagonal site $(n, n)$. The impenetrable boundary is represented by the parabola-like curve. The left panel shows the average paths, in the rescaled domain, for four values of $n$, whereas the right panel plots the distance between the average paths and the boundary $\partial\mathcal{D}$, rescaled by a factor $n^{1/2}$. In the region where the constraint is effective, one clearly

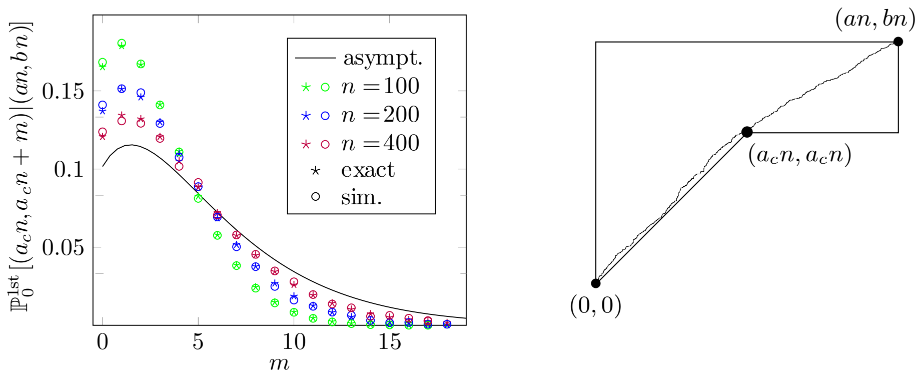



Figure 5.6: *Left:* Passage probability above $(a_c n, a_c n)$. The exact and asymptotic results $N_\lambda\left(\frac{m}{2}+1\right)\mathrm{e}^{-\lambda m}$ are compared to simulations and show good agreement. *Right:* Definition of the rectilinear domain under consideration, with a path sampled uniformly at random for $n = 1000$.

sees that the average paths stay at all times at a distance of order $n^{1/2}$ from it.

Exact calculations are often possible when the constraint is piecewise linear [162]. As a simple example, we consider the same walk as in the previous paragraph, with steps $(1,0)$ and $(0,1)$ each of weight 1. We condition the walk to start from the origin, to end at a distant site $(an, bn)$, with $\frac{b}{a} < 1$, and we constrain it not to visit sites under the diagonal line $y = x$ for $0 \leqslant x \leqslant a_\mathrm{c} n < an$, in such a way that the endpoint $(an, bn)$ is not reachable along a single straight line (see Figure 5.6-right). For $x > a_\mathrm{c} n$, the walk is no longer constrained although, due to the nature of the available steps, the walk will stay above (or on) the horizontal line $y = a_\mathrm{c} n$.

The lower boundary of the accessible domain $\mathcal{D}$ is made of two linear segments, a diagonal one and a horizontal one. We want to examine the alternative approach alluded to above, namely we solve the combinatorics of the constrained random paths, compute the continuous path $h^*$ on which the measure concentrates in the scaling limit and check that wherever the optimal path $h^*$ coincides with the diagonal constraint, the random lattice paths stay at a distance $\mathcal{O}(\sqrt{n})$ from it.

Let us denote by $Z_{r,s}$ the number of lattice paths from $(0,0)$ to $(r,s)$, and by $Z^{\mathrm{c}}_{r,s}$ the number of those which do not go under the diagonal line. They are explicitly given by

$$Z_{r,s} = \binom{r+s}{r}, \quad (r,s \geqslant 0), \qquad Z^{\mathrm{c}}_{r,s} = \frac{s-r+1}{s+1}\binom{r+s}{r}, \quad (s \geqslant r \geqslant 0). \tag{5.34}$$

Indeed they both satisfy the recurrence $a_{r,s} = a_{r-1,s} + a_{r,s-1}$ and the appropriate boundary conditions. These expressions make it clear that for $r, s \gg 1$ and fixed $t = \frac{s}{r} > 1$, $Z^{\mathrm{c}}_{r,s}$ is asymptotically equal to

$$Z^{\mathrm{c}}_{r,s} \sim \frac{t-1}{t}\sqrt{\frac{N_1(t)}{2\pi r}}\,\exp\{rL_1(t)\}, \tag{5.35}$$

where the functions $N_1(t)$ and $L_1(t)$, given in Section 5.1, pertain to the walk without the constraint. As anticipated, the Lagrangean function is the same in the constrained and unconstrained cases.

Let us first look at the passage probability distribution on the vertical line $x = a_{\mathrm{c}}n$. Because vertical steps are allowed, we consider the probability $\mathbb{P}^{\mathrm{1st}}_0[(a_{\mathrm{c}}n, a_{\mathrm{c}}n+m)|(an, bn)]$ that $(a_{\mathrm{c}}n, a_{\mathrm{c}}n+m)$ is the first visited site of the vertical line above the endpoint of the diagonal constraint, given that the path eventually reaches $(an, bn)$. This probability is proportional to the product of $Z^{\mathrm{c}}_{a_{\mathrm{c}}n-1, a_{\mathrm{c}}n+m}$ and $Z_{(a-a_{\mathrm{c}})n,(b-a_{\mathrm{c}})n-m}$. For large $n$, we find that the distribution is given by

$$\begin{aligned} &\mathbb{P}^{\mathrm{1st}}_0[(a_{\mathrm{c}}n, a_{\mathrm{c}}n+m)|(an, bn)] \sim N_\lambda \left(\frac{m}{2}+1\right)\mathrm{e}^{-\lambda m}, \\ &\lambda = \log\left[\frac{1}{2} + \frac{a-a_{\mathrm{c}}}{2(b-a_{\mathrm{c}})}\right], \end{aligned} \tag{5.36}$$

where $N_\lambda$ is the normalization factor. On Figure 5.6-left, this formula is compared with simulations. It follows that the average passage point remains at a distance $\mathcal{O}(1)$ above $(a_{\mathrm{c}}n, a_{\mathrm{c}}n)$, and implies that in the scaling limit, the continuous path $h^*$ visits the point $(a_{\mathrm{c}}, a_{\mathrm{c}})$ with probability 1. Since after that point there is no effective constraint, $h^*$ is a straight line from $(a_{\mathrm{c}}, a_{\mathrm{c}})$ to the final point $(a, b)$.

Let us now focus on the region above the diagonal constraint. As we have just seen that the passage distance to $(a_{\mathrm{c}}n, a_{\mathrm{c}}n)$ is of order $\mathcal{O}(1)$, we can consider the paths from $(0,0)$ to $(a_{\mathrm{c}}n, a_{\mathrm{c}}n)$ not going under the constraint. This exactly defines Dyck paths that are enumerated by the Catalan numbers $Z^{\mathrm{c}}_{r,r}$, which are ubiquitous in combinatorics. For $0 < \gamma < 1$, we look for the probability $\mathbb{P}^{\mathrm{1st}}_0[(\gamma a_{\mathrm{c}}n, \gamma a_{\mathrm{c}}n+m)|(a_{\mathrm{c}}n, a_{\mathrm{c}}n)]$ that

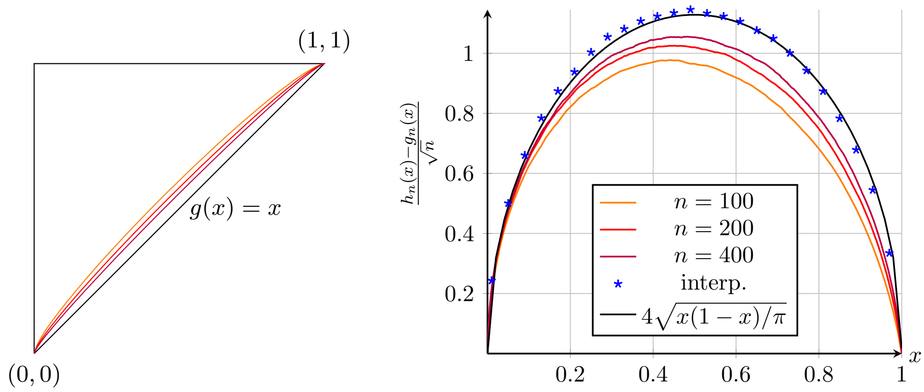



Figure 5.7: Dyck paths, having steps $\{(1,0),(0,1)\}$ constrained to stay above the diagonal. *Left:* Mean trajectories in the rescaled domain for various values of $n$, that tend toward the diagonal $g(x) = x$. *Right:* Lattice distance between the mean path $h_n(x)$ and the constraint $g_n(x) = ng(x)$, rescaled by $\sqrt{n}$. An interpolation shows good agreement with $\langle m \rangle \simeq \frac{4}{\sqrt{\pi}}\sqrt{\gamma(1-\gamma)a_c n}$.

the first passage on the vertical line $x = \gamma a_c n$ occurs at site $(\gamma a_c n, \gamma a_c n+ m)$. We obtain that this probability is equal to

$$\mathbb{P}_0^{1\text{st}}[(\gamma a_c n, \gamma a_c n+m)|(a_c n, a_c n)] = \frac{Z^{c}_{\gamma a_c n-1,\gamma a_c n+m} \times Z^{c}_{(1-\gamma)a_c n-m,(1-\gamma)a_c n}}{Z^{c}_{a_c n, a_c n}}. \tag{5.37}$$

For $n$ large and $m$ such that $\frac{m}{n}$ is small, we obtain the following asymptotic value,

$$\mathbb{P}_0^{1\text{st}} \sim \frac{(m+1)(m+2)}{2\sqrt{\pi}[(\gamma(1-\gamma)a_c n]^{3/2}} \exp\left\{\frac{-m^2}{4\gamma(1-\gamma)a_c n}\right\}, \qquad n \gg 1. \tag{5.38}$$

It implies that the average value $\langle m \rangle \simeq \frac{4}{\sqrt{\pi}}\sqrt{\gamma(1-\gamma)a_c n}$ is of order $\sqrt{n}$ above the diagonal constraint. Figure 5.7-right presents a numerical verification of this prediction for $\langle m \rangle$.

In this particular example, we can conclude that indeed the random lattice paths stay at a distance $\mathcal{O}(\sqrt{n})$ away from the constraint, with the exception of a small neighbourhood of the extremal point where the two linear segments of the boundary meet, which is also the only point where the boundary is not $\mathcal{C}^1$. We can however note that the entropy bias advocated above becomes negligible for such a small region: the gain of entropy is only substantial when the paths keep away from the constraint over a macroscopic length scale. We expect these characteristics to be generic.

The above combinatorial analysis shows that in the scaling limit, the constrained random lattice paths condense with probability 1 on the single continuous path $h^*$ going from the origin to the point $(a_c, a_c)$ by following the diagonal constraint, and then goes straight from $(a_c, a_c)$ to $(a, b)$. This is also what the variational problem in the continuum yields. Despite the fact that this boundary is not $\mathcal{C}^1$, it is easy to see that the concavity argument implies that the optimal path $h^*$ in the (rescaled) domain is piecewise linear, and made of the two linear pieces mentioned above.

## 5.3 Factorization conjecture

In this section, we strengthen assumption (i) into what we call the *factorization conjecture*. Similar ideas have been sketched, and written explicitly in special cases, by Sportiello [163, 164] under the names of entropic tangent method and 2-refined tangent method, described as alternative and tentatively more rigorous ways to compute the shape of arctic curves. The *factorization conjecture* that we propose in this section appears to be more general and more explicit.

In the previous section, we had a tangent method set-up in mind, where only the starting point was moved. The variational argument equally well describes the case where the endpoint is also moved. Let us consider a 2-refined partition functions $Z^{\mathrm{ref}}_{n+1}$, describing $n$ non-crossing lattice paths, to which is added an external path whose starting and ending points are modified (we will add more paths in a moment). We assume that the number of paths $n$ is directly related to the linear size of the system. The $n$ inner paths form the core of a configuration and produce an arctic curve. The factorization conjecture states that in the scaling limit, the full partition function $Z^{\mathrm{ref}}_{n+1}$ pertaining to the configurations of $n+1$ paths factorizes into the partition function $Z_n$ relative to the $n$ core paths and the contribution $Z^{\mathrm{out}}_1$ of the outermost path conditioned not to cross the arctic curve,

$$\log Z_{n+1}^{\mathrm{ref}} = \log Z_n + \log Z_1^{\mathrm{out}} + o(n). \tag{5.39}$$

We note that $\log Z_{n+1}^{\mathrm{ref}}$ and $\log Z_n$ ought to contain identical dominant terms (in particular a volume contribution proportional to $n^2$) up to order $n$, and differ by subdominant terms of order $n$, for which the contribution $Z_1^{\mathrm{out}}$ exactly compensates. So at the level of the entropies, we would expect the following identity,

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{Z_{n+1}^{\mathrm{ref}}}{Z_n} = \lim_{n\to\infty} \frac{1}{n} \log Z_1^{\mathrm{out}}. \tag{5.40}$$

We also note that in the full system of $n+1$ paths, the outermost one may or may not have its starting and ending points displaced. Such a modification, at the heart of the tangent method, will affect both $Z_{n+1}^{\mathrm{ref}}$ and $Z_1^{\mathrm{out}}$, but not $Z_n$. If the identity holds, the changes on $Z_{n+1}^{\mathrm{ref}}$ and $Z_1^{\mathrm{out}}$ are predicted to be exactly identical, at dominant order.

The above identity is reasonable and altogether not surprising. The underlying idea is that the behaviour of the bulk paths is essentially unaffected by the external path, viewed as a single random path confined to a domain which is itself random at finite volume, but becomes deterministic in the scaling limit.

There remains to evaluate the right-hand side of (5.40). We begin by considering a *single* lattice path that goes from $(na_i, nb_i)$ to $(na_f, nb_f)$, in a *fixed* domain. In the previous section, we used the action $S[f]$ to characterize the most likely trajectory $f^*$ reached in the scaling limit, which is found to be the shortest path from $(a_i, b_i)$ to $(a_f, b_f)$. In turn, once $f^*$ is known, it can be used to evaluate the dominant behaviour of the partition function,

$$\lim_{n\to\infty} \frac{1}{n} \log Z_{n(a_f-a_i),n(b_f-b_i)} = S[f^*] = \int_{a_i}^{a_f} \mathrm{d}x\, L\Big(\frac{\mathrm{d}f^*}{\mathrm{d}x}\Big). \tag{5.41}$$

It is quite tempting to apply the above result to compute $Z_1^{\mathrm{out}}$, with $f^*$ maximizing $S[f]$ in the domain induced by the arctic curve. Let us mention first a few caveats. The heuristic argument that justified the use of the variational approach in a fixed domain must be slightly adapted in this case. Indeed, we have seen in Chapter 3 that the lattice distance between the outermost paths is typically of order $n^{1/3}$. This behaviour is to be contrasted with the $n^{1/2}$ one that was observed for a fixed domain. In any case, the heuristics of Subsection 5.2.2 still applies; it suffices to change the scaling of $K$. Another complication might occur for osculating paths weighted in such a way as to favour kissing points so that paths are attracted to each other on the lattice. Even then, we expect the distance between the outermost paths to increase with $n$ due to entropic effects, provided that the weights do not scale with $n$. The heuristics hence still applies. We expect the following conjecture to hold for quite general models:

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{Z^{\mathrm{ref}}_{n+1}}{Z_n} = \int_{a_i}^{a_f} \mathrm{d}x\, L\Big(\frac{\mathrm{d}f^*}{\mathrm{d}x}\Big), \tag{5.42}$$

where $Z^{\mathrm{ref}}_{n+1}$ is the partition function for a system of $n+1$ lattice paths for which the outermost path goes from $(na_i, nb_i)$ to $(na_f, nb_f)$.

The discussion can be readily generalized to several external paths. If adding one path does not perturb the arctic curve, adding two, three, and in general $m$ external paths does not either, provided $m$ does not scale with $n$. It gives rise to a $2m$-refined partition function $Z^{\mathrm{ref}}_{n+m}$, depending on $m$ starting points and $m$ ending points. The natural generalization of (5.42) reads

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{Z^{\mathrm{ref}}_{n+m}}{Z_n} = \sum_{j=1}^{m} S[f^*_j] = \sum_{j=1}^{m} \int_{a_{i,j}}^{a_{f,j}} \mathrm{d}x\, L\Big(\frac{\mathrm{d}f^*_j}{\mathrm{d}x}\Big), \tag{5.43}$$

where $f^*_j$ is the likeliest trajectory of the $j$-th path. If the $m$-th path is the uppermost and the first one the lowermost, then the inequalities $f^*_1 \leqslant f^*_2 \leqslant \ldots \leqslant f^*_m$ hold. The maximizing function $f^*_j$ is in principle computed with respect to its own domain $\mathcal{D}_j$, bordered by the graph of $f^*_{j-1}$ (that is, the artic curve if $j = 1$). But in practice, the constraint from the rectilinear portions of $f^*_{j-1}$ are ineffective so that $f^*_j$ is only constrained by the arctic curve. In other words, each $f^*_j$ can be computed as in the 2-refined situation. In particular, each $S[f^*_j]$ is given by the same function $S[f^*](a_i, b_i; a_f, b_f)$ evaluated at $(a_{i,j}, b_{i,j}; a_{f,j}, b_{f,j})$. Examples are given in Section 5.4.

The formula (5.43) is the main point of this section and clearly remains conjectural. It is expected on the basis of the arguments reviewed above, but could possibly fail, like the tangent method itself, in situations where the interactions between the lattice paths are pathological. Many checks are presented in Section 5.4 and 5.5, namely for Aztec diamonds, in which the paths are strictly non-intersecting, and for alternating sign matrices, in which the paths can have kissing points. Specifically for the Aztec diamonds, the conjecture (5.43) will be verified in full generality, that is, for arbitrary $m$. In more general situations, the formula (5.43) allows to conjecture the asymptotic value of multirefined partition functions, provided the arctic curve is known.

As a special case of (5.42), the 1-to-0 discontinuity which is the core of the 2-refined tangent method discussed in [163] readily follows. The 2-refined partition function is $Z_{n+1}^{\text{ref}}$ computed with varying extremities $(na_i, nb_i)$ and $(na_f, nb_f)$ for the external path. One considers the asymptotic value of the following ratio,

$$\frac{Z_{n+1}^{\text{ref}}}{Z_n\, Z[f_{\text{st}}]} \approx \exp\left\{n\big(S[f^*] - S[f_{\text{st}}]\big)\right\}, \tag{5.44}$$

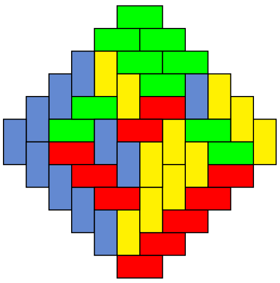
where $f_{\text{st}}$ is the straight trajectory between $(a_i, b_i)$ and $(a_f, b_f)$. If $(a_i, b_i)$ and $(a_f, b_f)$ are varied so that the straight line $f_{\text{st}}$ does not cross the arctic curve, then the maximizing $f^*$ *is* the straight line $f_{\text{st}}$ and the ratio is of order 1. The limiting values of $(a_i, b_i)$ and $(a_f, b_f)$ for this to be true are such that the line $f_{\text{st}}$ is tangent to the arctic curve. For $(a_i, b_i)$ and $(a_f, b_f)$ beyond these limiting values, the maximizing $f^*$ is no longer the straight line, with the consequence that $S[f^*] < S[f_{\text{st}}]$. It follows that the ratio becomes exponentially small in $n$ and goes to 0 in the scaling limit. The previous formula provides an explicit[5] expression for the exponential rate, namely $S[f^*] - S[f_{\text{st}}]$.

The observation underlying the entropic tangent method [164], which is based on a 1-refined partition function, also follows from the formula (5.42) and the nature of the maximizing trajectory $f^*$.

[5] That expression may not look so explicit. We will see in the following sections that it can be made completely explicit in concrete examples, provided the shape of the arctic curve is known.

## 5.4 Multirefinements for Aztec diamonds

It would be natural to start the verifications of the factorization hypothesis by considering tilings on the rectangoloid domain, considered in Section 2.1. Indeed, the Gelfand-Tsetlin formula (2.3) that enumerates these tilings directly yields refinement on the bottom boundary of the (extended) rectangoloid domain. Instead, we start the verification by considering a prototypical model in which an arctic curve occurs, namely the domino tilings of the Aztec diamond (see Figure 5.8). It was introduced in Section 1.2, where the LU decomposition of its LGV matrix was worked out. The reason for such a choice is that in this case, the lattice computation of multirefined partition functions—which is based on the LU decomposition—is not trivial and should be transposable to other tiling problems.



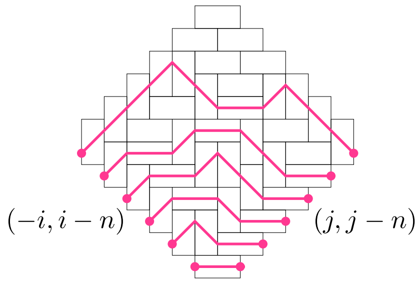



Figure 5.8: The bijection between a tiling of the Aztec diamond of order 6 and the corresponding 6-uple of NILP. The two kinds of horizontal and vertical dominoes are distinguished by the parity of their coordinates.

We briefly remind the salient features of this tiling model. We decide to weigh the dominoes according to their orientation, by assigning the horizontal and vertical dominoes a weight $\sqrt{w}$ and 1 respectively. The weight of a tiling is then equal to $w^{N_h/2}$ with $N_h$ the number of horizontal dominoes. The partition function for the Aztec diamond of order $n$ is given by

$$Z_n = \sum_{\text{tilings}} w^{N_h/2}. \tag{5.45}$$

The domino tilings of the Aztec diamond of order $n$ are in bijection with the $n$-uples of NILP [66,165], where the extremities of the paths are fixed on the lower left and lower right boundaries, as shown on Figure 5.8. We choose the coordinate system for which the center of the diamond is at $(0, \frac{1}{2})$ and such that the starting and ending points of the paths have coordinates $p_i \equiv (-i, i-n)$ and $q_j \equiv (j, j-n)$, for $i, j = 1, \ldots, n$. The paths themselves are made of three elementary steps: $(1, 1)$, $(1, -1)$ and $(2, 0)$ carried respectively by the blue, yellow and red dominoes (the lowest path from $p_1$ to $q_1$ cannot contain a down step). To get the correct weight, we give the diagonal steps $(1, 1)$ and $(1, -1)$ a weight 1, and the level step $(2, 0)$ a weight $w$ (there is an equal number of red dominoes, carrying a level step, and green dominoes, which carry no step at all). The LGV lemma 1.2.1 gives

$$Z_n = \det_{0 \leqslant i,j \leqslant n} (A_{i,j}), \quad A_{i,j} = \sum_{p=0}^{\min(i,j)} w^p \binom{i+j-p}{i-p, j-p, p}. \tag{5.46}$$

The LU decomposition of the semi-infinite matrix $(A_{i,j})_{i,j \geqslant 0}$ is

$$L_{i,j} = \binom{i}{j} \quad (i \geqslant j), \qquad U_{i,j} = (1+w)^i \binom{j}{i} \quad (i \leqslant j). \tag{5.47}$$

The semi-infinite LU decomposition directly yields that of any principal submatrix, by restriction; it also turns out to be crucial for the computation of general refined partition functions.

The calculation of $Z_n$ is now easy,

$$Z_n = \det_{0 \leqslant i,j \leqslant n} (U_{i,j}) = \prod_{i=0}^{n} U_{i,i} = (1+w)^{n(n+1)/2}. \tag{5.48}$$

Before proceeding to the verification of the factorization formula (5.43), we briefly discuss the scaling limit.

The scaling limit is obtained by dividing all distances by $n$ and taking the limit $n$ going to infinity. The Aztec diamond goes to a square centered at the origin and rotated by 45°, partly shown in Figure 5.9. For generic $w$, the arctic curve is an ellipse, with equation [7]

$$\frac{x^2}{w} + y^2 = \frac{1}{1+w}. \tag{5.49}$$

The interior of the ellipse, namely the temperate region, is represented in Figure 5.9 as the shaded area.

The $2m$-refinements we consider consist in imposing constraints on the $m$ uppermost paths. For convenience, we choose to keep $n$ unconstrained paths, so that the normalizing partition is always the same, and equal to $Z_n$. Different refinements can be considered, depending on the starting and ending points of the extra paths. The ensuing calculations in the scaling limit are not very sensitive to the actual choice, but the lattice calculations are. A natural choice is to constrain the $m$ paths to take their first step not equal to $(1, 1)$ at prescribed positions along the upper left boundary, and likewise to constrain them to have their last step not equal to $(1, -1)$ at some other positions along the upper right boundary, in a similar way as in Figure 2.4. Equivalently, we can actually move the starting and ending points from their original positions to new locations by inserting monomers at appropriate positions.

A simplifying choice turns out to let the $m$ paths start and end on the lines which form the continuations of the lower left and right boundaries of the diamond. More precisely, for two growing sequences $0 \leqslant r_1 \leqslant \ldots \leqslant r_m$ and $0 \leqslant s_1 \leqslant \ldots \leqslant s_m$, the $j$th path starts from the point $(-1 - r_j, r_j)$ and ends at $(1 + s_j, s_j)$, after the rescaling of the coordinates by $n$, as shown in Figure 5.9. Thus, the extra path with label 1 is the innermost, that with label $m$ the outermost. The refined partition function, corresponding to the $n + m$ paths, will be denoted by $Z_{n+m}(nr_1, nr_2, \ldots | ns_1, ns_2, \ldots)$.

This second way of defining the refinements is not really different from the first one. In the scaling limit, the paths starting at $(-1 - r_j, r_j)$ and ending at $(1 + s_j, s_j)$ will enter and exit the domain at certain points, making contact with the first refinements. Fixing one set of points or the other is merely a change of variables. The only difference is the contributions coming from the straight portions located outside the domain, which can easily be taken into account.

Let us first consider the case of a single extra path, with starting and ending points $P_i$ and $P_f$ specified by a pair $(r, s)$. For $r$ and $s$ large enough, the straight line from $P_i$ to $P_f$ does not cross the arctic ellipse. From the discussion of the previous section, it follows that in the scaling limit, the shape of the extra path is almost surely the straight line. This case is illustrated by the blue line in Figure 5.9. When $r$ and/or $s$ decrease, the straight line will at some point cross the arctic ellipse.

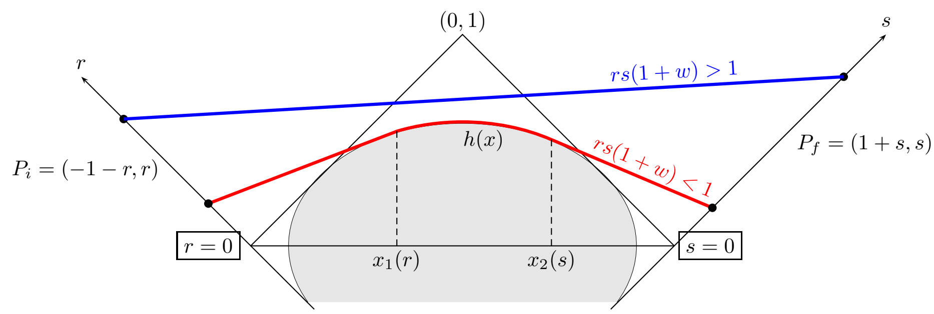



Figure 5.9: Graphical representation of a 2-refinement for the Aztec diamond. An extra path is added, starting from the left boundary and ending on the right boundary. In the scaling limit, the random path collapses almost surely on a trajectory that looks like the blue straight line or the red curve, depending on the initial and final positions.

Then the path follows almost surely a trajectory like that represented in red in Figure 5.9, which avoids the temperate region by minimizing the length from $P_i$ to $P_f$. The factorization conjecture predicts the following limit,

$$S(r,s) \equiv \lim_{n\to\infty} \frac{1}{n} \log \frac{Z_{n+1}(nr|ns)}{Z_n} = S[f^*] = \int_{-1-r}^{1+s} \mathrm{d}x \, L_2\Big(\frac{\mathrm{d}f^*}{\mathrm{d}x}\Big), \tag{5.50}$$

for the appropriate function $L_2$, given by (5.6), and where $f^*$ is computable in terms of $r, s$ and looks like the blue or the red trajectory. In the second case, the knowledge of the arctic curve is required to compute $S(r,s)$ explicitly.

In case we add $m$ paths, and depending on the values of the pairs $(r_j, s_j)$, paths will either be straight lines or follow portions of arctic curve, in the scaling limit. Any path above a straight one must also be straight. Each of them has an associated action given by $S[f_j^*] = S(r_j, s_j)$, and the following limit should hold,

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \log & \frac{Z_{n+m}(nr_1,\ldots,nr_m|ns_1,\ldots,ns_m)}{Z_n} \\ &= S(r_1,s_1) + S(r_2,s_2) + \ldots + S(r_m,s_m). \end{aligned} \tag{5.51}$$

### 5.4.1 The 2-refined entropy function

The function $S(r,s)$ appears as the entropy associated to a single path, or as the action $S[f^*]$ computed along the maximizing trajectory. To compute it, there only remains to have an explicit formula for $f^*(x)$.

We first consider the case when $r$ and $s$ are small enough so that the trajectory $f^*$ touches the arctic curve. It is then composed of two line segments with slopes $t_1(r)$ and $t_2(s)$ and of a portion of arctic curve between the two tangency points $\big(x_1(r), h(x_1(r))\big)$ and $\big(x_2(s), h(x_2(s))\big)$, with $h(x)$ the arctic curve, see Figure 5.9. It is not hard to show that

$$t_1(r) = \frac{1-r^2(1+w)}{1+r(2+r)(1+w)}, \quad x_1(r) = -wt_1(r)\sqrt{(1+w)(1+wt_1(r))}, \tag{5.52}$$

while $t_2(s) = -t_1(s)$ and $x_2(s) = -x_1(s)$ are obtained by symmetry. When $r$ and/or $s$ get larger, the portion of arctic curve spanned by $f^*$ decreases, until it is reduced to a single point when $t_1 = t_2$, implying $rs = \frac{1}{1+w}$. This condition is therefore the borderline between the two cases illustrated in Figure 5.9.

When $rs \leqslant \frac{1}{1+w}$, the action takes a surprisingly simple form,

$$\begin{aligned} S[f^*] &= (x_1+1+r)L_2(t_1) + \int_{x_1}^{x_2} \mathrm{d}x \; L_2\big(h'(x)\big) + (1+s-x_2)L_2(t_2) \\ &= \mathcal{L}(1+r) - \mathcal{L}(r) + \mathcal{L}(1+s) - \mathcal{L}(s) + \log(1+w), \end{aligned} \tag{5.53}$$

where we again used $\mathcal{L}(x) \equiv x\log(x)$. In case $rs \geqslant \frac{1}{1+w}$, the trajectory is a straight line of slope $\frac{s-r}{2+r+s}$, so that the integral giving $S[f^*]$ is straightforward.

We thus obtain

$$S(r,s) = \begin{cases} (r+s+2)\, L_2\Big(\dfrac{s-r}{r+s+2}\Big), & rs \geqslant \frac{1}{1+w}, \\[2ex] \mathcal{L}(r+1) + \mathcal{L}(s+1) - \mathcal{L}(r) - \mathcal{L}(s) + \log(1+w), & rs \leqslant \frac{1}{1+w}. \end{cases} \tag{5.54}$$

Let us already note that the factorization is satisfied in the case $r = s = 0$. Indeed, $Z_{n+1}(0|0) = Z_{n+1}$ is the large $n$ limit of the partition function of the Aztec diamond of order $n+1$, and the conjecture (5.51) reads

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{Z_{n+1}}{Z_n} = S(0,0) = \log(1+w). \tag{5.55}$$

The identity is satisfied in view of $Z_n = (1+w)^{n(n+1)/2}$. The same readily holds for the $2m$-refined partition function $Z_{n+m}(0,\ldots,0|0,\ldots,0)$.

We now proceed to verify the factorization conjecture for a general $2m$-refinement, starting with the case $m = 1$.

### 5.4.2 Lattice 2-refined partition functions

In terms of paths, the lattice 2-refined partition function is defined from the paths configurations of the Aztec diamond of order $n$ to which we add one extra path, with starting and ending points $p_{n+k} = (-n-k, k)$ and $q_{n+\ell} = (n+\ell, \ell)$ for $k, \ell \geqslant 1$. The corresponding partition function, which we denote by $Z_{n+1}(k|\ell)$, is the weighted sum of all configurations of $n+1$ non-intersecting paths starting from $\{p_1, p_2, \ldots, p_n, p_{n+k}\}$ and ending at $\{q_1, q_2, \ldots, q_n, q_{n+\ell}\}$. In the scaling limit, when $k, \ell$ are proportional to $n$, the lattice and continuous variables are related by $k = rn$ and $\ell = sn$.

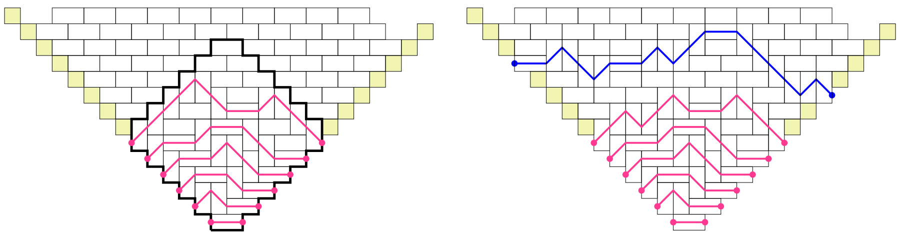


Figure 5.10: Graphical representation of how to add an extra path, with fixed starting and ending points, to the standard set of $n$ paths for an Aztec diamond of order $n$ ($n = 6$ here). The yellow cells represent the monomers.

In terms of tilings, the way an extra path can be added is depicted in Figure 5.10. Before adding the extra path, the Aztec diamond of order $n$ is embedded in a larger quadrant by placing one layer of monomers along the left boundary and one along the right boundary, and filling the region above the diamond with green dominoes. This makes sure that the extended domain contains no more paths than the original Aztec diamond. Instead of considering the infinite quadrant just discussed, an alternative is to cut it off horizontally at a large enough height. Dominoes in the extension are then deterministically frozen and are all green ones. The location of the horizontal boundary is chosen to be high above the path to be added, so that no path can touch it. Once the extension is realized, one removes one monomer on the left and one on the right, at the required positions; a new path, in blue in Figure 5.10, is then automatically created.

According to the LGV formula, the 2-refined partition function is given by the following determinant of rank $n+2$,

$$Z_{n+1}(k|\ell)=\det_{0\leqslant i,j\leqslant n+1}\left[A_{i,j}(k|\ell)\right],\;\; A_{i,j}(k|\ell)=\begin{cases}A_{i,j} & i,j,\leqslant n,\\ A_{n+k,j}, & i=n+1,\ j\leqslant n,\\ A_{i,n+\ell}, & i\leqslant n,\ j=n+1,\\ A_{n+k,n+\ell}, & i,j=n+1.\end{cases} \tag{5.56}$$

Defining the following combination of the $L$ and $U$ matrix entries given in (5.47),

$$g=\sum_{j=1}^{\min(k,\ell)} L_{n+k,n+j}\,U_{n+j,n+\ell}, \tag{5.57}$$

one can verify, by using the semi-infinite LU decomposition of $A$, that the LU decomposition of $A(k|\ell)=L(k|\ell)U(k|\ell)$ is given by

$$L_{i,j}(k|\ell)=\begin{cases}L_{i,j}, & i\leqslant n, j\leqslant n+1,\\ L_{n+k,j}, & i=n+1, j\leqslant n,\\ 1, & i=j=n+1,\end{cases}\;\; U_{i,j}(k|\ell)=\begin{cases}U_{i,j}, & i\leqslant n+1, j\leqslant n,\\ U_{i,n+\ell}, & i\leqslant n, j=n+1,\\ g, & i=j=n+1.\end{cases} \tag{5.58}$$

From this, we readily obtain

$$Z_{n+1}(k|\ell)=g\det_{0\leqslant i,j\leqslant n}\left[U_{i,j}(k|\ell)\right]=g\det_{0\leqslant i,j\leqslant n}\left[U_{i,j}\right], \tag{5.59}$$

and therefore

$$Z_1^{\text{out}}(k|\ell)=\frac{Z_{n+1}(k|\ell)}{Z_n}=g=\sum_{j=1}^{\min(k,\ell)}(1+w)^{n+j}\binom{n+k}{n+j}\binom{n+\ell}{n+j}. \tag{5.60}$$

In order to make contact with the function $S(r,s)$, we set $k=rn$, $\ell=sn$, and perform a saddle point analysis of the previous sum, for which we may assume $k\leqslant\ell$ by symmetry in $k,\ell$. Keeping the exponential terms only, we find

$$g \approx \int_0^r \mathrm{d}\xi \exp\big\{nF(\xi;r,s)\big\} \approx \exp\big\{nF(\xi^*;r,s)\big\},$$
$$F(\xi;r,s) \equiv \mathcal{L}(r{+}1) - \mathcal{L}(r{-}\xi) + (r \to s) + (\xi{+}1)\log(1{+}w) - 2\mathcal{L}(\xi{+}1), \tag{5.61}$$

where $\xi^*$ is the location of the maximum of $F(\xi;r,s)$ in the integration domain (if unique). The function $F$ is strictly concave over the entire domain $[0,r]$, and its derivatives at the endpoints are equal to $\partial_\xi F(0;r,s) = \log\big[rs(1+w)\big]$ and $\partial_\xi F(r;r,s) = -\infty$. The extremum condition $\partial_\xi F = 0$ yields the quadratic equation

$$(\xi^* + 1)^2 = (1+w)(r-\xi^*)(s-\xi^*). \tag{5.62}$$

Let us first assume $rs(1+w) \leqslant 1$. Since $\partial_\xi F(0;r,s) \leqslant 0$, the concavity of $F$ implies that it is strictly decreasing on $[0,r]$. Its maximum on the domain is therefore at $\xi^* = 0$, with the value

$$F(\xi^*;r,s) = \mathcal{L}(r{+}1) - \mathcal{L}(r) + \mathcal{L}(s{+}1) - \mathcal{L}(s) + \log(1{+}w), \tag{5.63}$$

exactly equal to $S(r,s)$.

When $rs(1+w) > 1$, the derivative of $F$ at $\xi = 0$ being strictly positive, there is a unique maximum in the interior of $[0,r]$. Because the product of the two roots of (5.62) is equal to $\big(rs(1+w)-1\big)/w$, hence positive, one is in $[0,r]$ and the other is strictly larger than $r$. Thus, the location of the maximum of $F$ in $[0,r]$ is at the smallest of the two roots,

$$\xi^* = \frac{1}{2w}\Big[2{+}(r{+}s)(1{+}w) - \sqrt{\big[2{+}(r{+}s)(1{+}w)\big]^2 + 4w\big[1{-}rs(1{+}w)\big]}\Big]. \tag{5.64}$$

Using the quadratic equation satisfied by $\xi^*$, the maximal value of $F$ can be written as

$$\begin{aligned} F(\xi^*;r,s) &= (r+1)\log\frac{r+1}{r-\xi^*} + (s{+}1)\log\frac{s+1}{s-\xi^*} \\ &= -(r{+}s{+}2)\log\Big[\frac{(r{-}\xi^*)(s{-}\xi^*)}{(r{+}1)(s{+}1)}\Big]^{\frac{1}{2}} - (s{-}r)\log\Big[\frac{(r{+}1)(s{-}\xi^*)}{(r{-}\xi^*)(s{+}1)}\Big]^{\frac{1}{2}}. \end{aligned} \tag{5.65}$$

This last expression can be identified with the function $S(r,s)$, namely

$$(r{+}s{+}2)L_2(t) = -(r{+}s{+}2)\log x_2(t) - (s{-}r)\log y_2(t), \quad t = \frac{s-r}{r+s+2}, \tag{5.66}$$

provided the arguments of the logarithms match. The two functions $x_2(t)$ and $y_2(t)$, are the only positive solutions of two algebraic equations, which read in this case,

$$xy + \frac{x}{y} + wx^2 = 1, \qquad xy - \frac{x}{y} = t(1 + wx^2). \tag{5.67}$$

The equality of (5.65) and (5.66) is thus verified if one can show that the following two functions,

$$X = \Big[\frac{(r-\xi^*)(s-\xi^*)}{(r+1)(s+1)}\Big]^{\frac{1}{2}}, \qquad Y = \Big[\frac{(r+1)(s-\xi^*)}{(r-\xi^*)(s+1)}\Big]^{\frac{1}{2}}, \tag{5.68}$$

satisfy the same algebraic equations, which is completely straightforward by using the quadratic equation satisfied by $\xi^*$.

We have shown that

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{Z_{n+1}(nr|ns)}{Z_n} = F(\xi^*; r, s) = S(r, s), \tag{5.69}$$

and confirmed the factorization conjecture for the 2-refined case.

### 5.4.3 Lattice multirefined partition functions

For the general multirefined case, we include $m$ additional paths to the $n$ paths forming the standard Aztec diamonds of order $n$, with $m = \mathcal{O}(1)$. Following the conventions used before, we set the starting points to $\{p_1, p_2, \ldots, p_n, p_{n+k_1}, \ldots, p_{n+k_m}\}$ and the ending points to $\{q_1, q_2, \ldots, q_n, q_{n+\ell_1}, \ldots, q_{n+\ell_m}\}$ for two strictly increasing sequences $1 \leqslant k_1 < \ldots < k_m$ and $1 \leqslant \ell_1 < \ldots < \ell_m$. The corresponding partition function is denoted by $Z_{n+m}(k_1, \ldots, k_m|\ell_1, \ldots, \ell_m)$ or the shorter $Z_{n+m}(\vec{k}|\vec{\ell})$.

It is given by the determinant of the $(n+1+m) \times (n+1+m)$ matrix

$$A_{i,j}(\vec{k}|\vec{\ell}) = \begin{cases} A_{i,j} & i, j \leqslant n, \\ A_{n+k_a,j} & i = n+a,\ j \leqslant n, \\ A_{i,n+\ell_b} & i \leqslant n,\ j = n+b, \\ A_{n+k_a,n+\ell_b} & i = n+a,\ j = n+b, \end{cases} \qquad 1 \leqslant a, b \leqslant m. \tag{5.70}$$

Let us write the LU decomposition of $A(\vec{k}|\vec{\ell})$ in the following form,

$$L(\vec{k}|\vec{\ell})=\begin{bmatrix} L & 0 \\ M & l \end{bmatrix},\ U(\vec{k}|\vec{\ell})=\begin{bmatrix} U & V \\ 0 & u \end{bmatrix},\ A(\vec{k}|\vec{\ell})=\begin{bmatrix} LU & LV \\ MU & MV+lu \end{bmatrix}, \quad (5.71)$$

where $L, U$ decompose the principal submatrix $(A_{i,j})_{0\leqslant i,j\leqslant n}$, $l, u$ are $m\times m$ lower and upper triangular blocks, and $M$ and $V$ are rectangular, $m\times(n+1)$ and $(n+1)\times m$ respectively.

It is not difficult to see that $M$ and $V$ are given by

$$M_{a,j}=L_{n+k_a,j}, \quad V_{i,b}=U_{i,n+\ell_b}, \qquad 1\leqslant a,b\leqslant m,\ 0\leqslant i,j\leqslant n. \quad (5.72)$$

One verifies for instance, for $j\leqslant n$,

$$(MU)_{a,j}=\sum_{i=0}^{n} L_{n+k_a,i}\,U_{i,j}=\sum_{i\geqslant 0} L_{n+k_a,i}\,U_{i,j}=A_{n+k_a,j}, \quad (5.73)$$

and likewise for $LV$. From this, we find that

$$\begin{aligned} g_{a,b}\equiv(lu)_{a,b}&=A_{n+k_a,n+\ell_b}(\vec{k}|\vec{\ell})-(MV)_{a,b} \\ &=\sum_{j=1}^{\min(k_a,\ell_b)} L_{n+k_a,n+j}\,U_{n+j,n+\ell_b}, \quad 1\leqslant a,b\leqslant m, \end{aligned} \quad (5.74)$$

is the matrix generalization of the quantity we called $g$ in the previous subsection. We thus obtain,

$$\frac{Z_{n+m}(\vec{k}|\vec{\ell})}{Z_n}=\det_{1\leqslant a,b\leqslant m}\left(Z_1^{\text{out}}(k_a|\ell_b)\right)=\det_{1\leqslant a,b\leqslant m}(g_{a,b}), \quad (5.75)$$

a formula that is surprisingly reminiscent of an LGV determinant (it is not really an LGV determinant since the numbers $g_{a,b}$ are not weighted sums over paths). Incidentally, and since all the entries of $L$ and $U$ are integers for $w=1$ (or any positive integer), the formula shows that in this case, all ratios $Z_{n+m}(\vec{k}|\vec{\ell})/Z_n$ are integers, something that was not a priori obvious.

In the scaling limit, we set $k_a=r_a n$ and $\ell_b=s_b n$, and using the results of the 2-refined case for the asymptotic value of $g_{a,b}$, namely,

$$g_{a,b}=f_{a,b}\,\exp\{nS(r_a,s_b)\}, \qquad \lim_{n\to\infty}\frac{1}{n}\log f_{a,b}=0, \quad (5.76)$$

we have,

$$\frac{Z_{n+m}(nr_1,\ldots,nr_m|ns_1,\ldots,ns_m)}{Z_n} = \det_{1\leqslant a,b\leqslant m}\big(f_{a,b}\,\mathrm{e}^{nS(r_a,s_b)}\big). \tag{5.77}$$

Expanding the determinant as a sum over the symmetric group $S_m$, we obtain

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\log\frac{Z_{n+m}(nr_1,\ldots,nr_m|ns_1,\ldots,ns_m)}{Z_n}&\\ \leqslant \max_{\pi\in S_m}\big\{S(r_1,s_{\pi(1)})+\ldots+S(r_m,s_{\pi(m)})\big\}.&\end{aligned} \tag{5.78}$$

In case the maximum is attained for more than one permutation, subtle cancellations could occur amongst the factors $f_{a,b}$, resulting in a strict inequality, although this is not to be expected. We will show that (1) the identity permutation is always amongst the permutations for which the maximum is reached, though it is in general not the only one, and (2) that no cancellations occur if, for those $a$ such that $r_a s_a > \frac{1}{1+w}$, there is no equality between two $r_a$ or between two $s_a$. The proofs of both claims rely on the property that the difference $S(r,s)-S(r,s')$, for any $s\geqslant s'$, is an increasing function of $r$; the proof is given in Appendix C.

For the statement (1), we first consider $m=2$. In this case, since $r_2\geqslant r_1$, we have indeed,

$$\begin{aligned}&[S(r_1,s_1)+S(r_2,s_2)]-[S(r_1,s_2)+S(r_2,s_1)]\\ &\qquad=[S(r_2,s_2)-S(r_2,s_1)]-[S(r_1,s_2)-S(r_1,s_1)]\geqslant 0.\end{aligned} \tag{5.79}$$

For general $m$, we proceed by recurrence, assuming that the identity permutation in $S_{m-1}$ yields a maximum. Let $\pi\in S_m$ be a permutation such that $\pi(m)\neq m$ (otherwise we are done). Then

$$\begin{aligned}&\sum_{a=1}^{m}[S(r_a,s_a)-S(r_a,s_{\pi(a)})]\\ &=[S(r_m,s_m)-S(r_m,s_{\pi(m)})]-[S(r_{\pi^{-1}(m)},s_m)-S(r_{\pi^{-1}(m)},s_{\pi^{-1}(m)})]+\ldots\\ &\geqslant[S(r_{\pi^{-1}(m)},s_m)-S(r_{\pi^{-1}(m)},s_{\pi(m)})]-[S(r_{\pi^{-1}(m)},s_m)-S(r_{\pi^{-1}(m)},s_{\pi^{-1}(m)})]+\ldots\\ &=[S(r_{\pi^{-1}(m)},s_{\pi^{-1}(m)})-S(r_{\pi^{-1}(m)},s_{\pi(m)})]+\sum_{a\neq m,\pi^{-1}(m)}[S(r_a,s_a)-S(r_a,s_{\pi(a)})]\\ &=\sum_{a=1}^{m-1}[S(r_a,s_a)-S(r_a,s_{\tilde\pi(a)})]\geqslant 0,\end{aligned} \tag{5.80}$$

where $\tilde{\pi} \in S_{m-1}$ is the permutation acting on $\{1, 2, \ldots, m-1\}$ as

$$\tilde{\pi}(a) = \begin{cases} \pi(a) & \text{for } a \neq \pi^{-1}(m), \\ \pi(m) & \text{for } a = \pi^{-1}(m). \end{cases} \tag{5.81}$$

As the claim is true for $m = 2$, it holds for every $m$.

Let $m_0$ be such that $r_1 s_1 \leqslant \ldots \leqslant r_{m_0} s_{m_0} \leqslant \frac{1}{1+w} < r_{m_0+1} s_{m_0+1} \leqslant \ldots \leqslant r_m s_m$. For the point (2), we first characterize, under the assumptions stated (no two $r$ and no two $s$ with labels strictly larger than $m_0$ are equal), the permutations for which the maximum is attained, namely those $\pi$ such that

$$\sum_{a=1}^{m} S(r_a, s_{\pi(a)}) = \sum_{a=1}^{m} S(r_a, s_a). \tag{5.82}$$

We note that the group $S_{m_0}$ of permutations of the labels $a \leqslant m_0$ satisfy this identity since $S(r, s)$ has the form $\sigma(r) + \sigma(s)$ for such pairs $(r, s)$. We prove that (5.82) is satisfied for no other permutation.

Let $\pi$ be such a non-trivial permutation, so that there exists a label $b$ with $\pi(b) \neq b$ such that either $b > m_0$ or $\pi(b) > m_0$, or both. Without loss of generality, we may assume that $b > m_0$ and also $s_b > s_{\pi(b)}$, $r_b > r_{\pi^{-1}(b)}$. In this case, we repeat the chain of (in)equalities in (5.80) in which we replace, in the second line, the label $m$ by $b$. However, on the third line, we obtain a strict inequality since with the conditions we have just stated, the difference $S(r_b, s_b) - S(r_b, s_{\pi(b)})$ is strictly larger than $S(r_{\pi^{-1}(b)}, s_b) - S(r_{\pi^{-1}(b)}, s_{\pi(b)})$, as shown in Appendix C.

Coming back to the determinant (5.77), we see that the terms having the same exponential rate in $n$ are given by

$$\begin{aligned} &\frac{Z_{n+m}(nr_1, \ldots, nr_m | ns_1, \ldots, ns_m)}{Z_n} \\ &\quad \approx \det_{1 \leqslant a,b \leqslant m_0} \left( f_{a,b} \, \mathrm{e}^{n\sigma(r_a) + n\sigma(s_b)} \right) \times \Big[ \prod_{a=m_0+1}^{m} f_{a,a} \, \mathrm{e}^{nS(r_a, s_a)} \Big] \\ &\quad = \det_{1 \leqslant a,b \leqslant m_0} \left( f_{a,b} \right) \times \Big[ \prod_{a=m_0+1}^{m} f_{a,a} \Big] \times \exp\Big\{ n \sum_{a=1}^{m} S(r_a, s_a) \Big\}. \end{aligned} \tag{5.83}$$

From the definition (5.76), none of the coefficient $f_{a,a}$ vanishes. The determinant does not vanish either because the partition function for the $m_0$ extra paths alone is proportional to it,

$$\begin{aligned}&\frac{Z_{n+m_0}(nr_1,\ldots,nr_{m_0}|ns_1,\ldots,ns_{m_0})}{Z_n}\\&\quad=\det_{1\leqslant a,b\leqslant m_0}\left(f_{a,b}\,\mathrm{e}^{nS(r_a,s_b)}\right)=\det_{1\leqslant a,b\leqslant m_0}\left(f_{a,b}\right)\exp\left\{n\sum_{a=1}^{m_0}S(r_a,s_a)\right\}.\end{aligned} \tag{5.84}$$

Under the assumptions we made, namely $r_a \neq r_b$ and $s_a \neq s_b$ for $a, b \geqslant m_0 + 1$, we have thus proved that the upper bound is (5.78) is reached. The completely general situation corresponds to the existence of subsets $T_i$ of labels larger than $m_0$ yielding identical $r$-values or $s$-values. Using the same arguments as above, we can extend the proof provided the subsets $T_i$ are disjoint. Whether the $T_i$ subsets are disjoint or not, the permutations satisfying (5.82) can be determined explicitly. However, when the $T_i$ are not disjoint[6], these permutations do not form a direct product of permutations subgroups (they do not even form a group), and therefore do not build full subdeterminants (as the group $S_{m_0}$ above did). The above arguments then fail. But as these special cases are limits of cases for which our proof works, the result presumably holds for them as well, by continuity.

To summarize, we have shown that the following equality generically holds,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\log&\frac{Z_{n+m}(nr_1,\ldots,nr_m|ns_1,\ldots,ns_m)}{Z_n}\\&=S(r_1,s_1)+S(r_2,s_2)+\ldots+S(r_m,s_m),\end{aligned} \tag{5.85}$$

thereby proving the general conjecture in the case of Aztec diamonds.

## 5.5 Multirefinements for alternating sign matrices

In this section, we show that the factorization conjecture appears to apply to osculating paths, i.e. non-crossing paths that may have kissing

[6]The simplest situation of that type involves three doublets $(r_1, s_1)$, $(r_2, s_2)$ and $(r_3, s_3)$ with $r_1 = r_2$ and $s_2 = s_3$, corresponding to $T_1 = \{1, 2\}$ and $T_2 = \{2, 3\}$. In this case, the permutations which satisfy (5.82) are $\pi_1(1, 2, 3) = (2, 1, 3)$, $\pi_2(1, 2, 3) = (1, 3, 2)$ and the identity. They do not form a subgroup of $S_3$.

points. We will verify it in the case of multirefined enumerations of ASM. Exact lattice results are harder to obtain than for the Aztec diamond case and the known formulae quickly become fairly complicated as the degree of refinement increases.

An ASM of size $n$ is an $n \times n$ matrix with entries $m_{i,j}$ taking values 1, 0, and $-1$ in such a way that the 1 and $-1$ entries alternate in every column and every row and such that all row and column sums are equal to 1 (first and last non-zero entries are 1's). We will denote by $\mathrm{ASM}_n$ the set of ASM of order $n$, and mainly consider the uniform distribution on $\mathrm{ASM}_n$.

A matrix in $\mathrm{ASM}_n$ can be represented bijectively as a set of $n$ non-crossing lattice paths drawn on an $n \times n$ grid. The presence of a 1 or a $-1$ in the matrix at position $(i, j)$ means that the site $(i, j)$ is visited by one and only one of the $n$ paths, which in addition makes at that site one of the two turns shown on Figure 5.11-middle. These fix completely the rest of the $n$ paths, required to connect the $n$ vertices of the last column and those on the last row without crossing each other, although they are allowed to touch (they can share sites but not edges). If we see the entries of the last column as entrance points and those on the last row as exit points, then, in the $(i, j)$ plane, the paths are made of the elementary steps $(1, 0)$ and $(0, -1)$. These elementary steps are respectively called "right" and "down", when viewed in the reference frame of Figure 5.11-right, were we look at the matrix with the $i$-axis horizontal and the $j$-axis vertical. This model is equivalent to the 6V model with DWBC and weights $a = b = c = 1$ ($\Delta = \frac{1}{2}$ and $t = 1$), see Figure 5.11-right where yet another equivalent osculating path description was used.

The path representation allows to apply the tangent method but prevents the use of the LGV lemma to compute the partition function, as the paths in general are not non-intersecting. The expression for the partition function for $\mathrm{ASM}_n$ is,

$$A_n = \prod_{j=0}^{n-1} \frac{(3j+1)!}{(n+j)!} = \prod_{j=0}^{n-1} \frac{\binom{3j+1}{j}}{\binom{2j}{j}}. \tag{5.86}$$

As in the case of the Aztec diamond, the simplicity of the partition function calls for an elementary proof and as for the Aztec diamond, it remains elusive. The original proof of (5.86) provided in [166] was

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

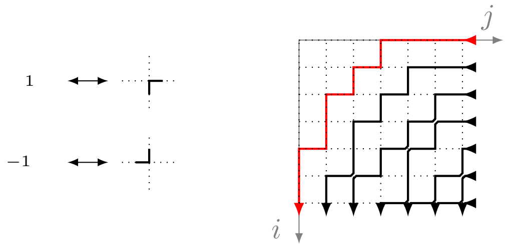



Figure 5.11: *Left:* Example of ASM of order $n = 7$. *Middle:* Rule used in the construction the corresponding osculating paths, by which the 1 and $-1$ are replaced by some portions of paths. *Right:* The osculating path are obtained by completing the remaining portions in a way to obtain $n$ paths that start from the right boundary and arrive at the bottom boundary without crossing. In the scaling limit, a portion of the outermost path, in red, condenses onto a portion of arctic curve.

very far from being elementary, see the book of Bressoud [167] for an historical account. The integrability machinery provides a substantially shorter proof [168].

In the scaling limit, ASM exhibit an arctic curve which separates the central region where the three entries $0, -1$ and $+1$ are present with non-zero densities, from four disconnected regions around the corners, which are filled with 0's with probability 1, see Figure 5.12. The exact shape of the arctic curve was first conjectured in [13] and rigorously computed in [21]. In the coordinate system used in Figure 5.12, in which $x$ and $y$ are the continuous counterparts of the row and column labels, the north-west portion of the arctic curve satisfies the conic equation $x(1-x) + y(1-y) + xy = \frac{1}{4}$, for $0 \leqslant x, y \leqslant \frac{1}{2}$, from which the upper branch is selected,

$$h(x) = \frac{1}{2}\big(1 + x - \sqrt{3x(2-x)}\big), \qquad 0 \leqslant x \leqslant \frac{1}{2}. \tag{5.87}$$

The other three portions can be deduced from mirror symmetry with respect to the horizontal and the vertical central axes. In what follows, we focus on the north-west portion. Other portions may also be studied by using other path descriptions.

An easy property of ASM is that they all have a single 1 (thus no $-1$) on the first and last rows and columns. For large and generic ASM, each of these 1's are almost surely located in the scaling neighbourhood of the middle point of their row or column. In the discrete setting, this means that the uppermost path takes its first step rightwards, i.e. $(1, 0)$, around the position $(i, j) = (1, \frac{n}{2})$ and its last step downwards around the position $(\frac{n}{2}, 2)$; in between these two positions it follows a zig-zag trajectory around the arctic curve. In the scaling limit and in the reference frame of Figure 5.12, the uppermost path starts from $(0, 1)$, joins the point $(0, \frac{1}{2})$ vertically, follows the arctic curve till $(\frac{1}{2}, 0)$, and finally moves horizontally to reach the point $(1, 0)$.

Two-refined partition functions can then be defined by considering ASM with a 1 at fixed positions on the first row and first column, $m_{1,k} = 1$ and $m_{\ell,1} = 1$, with $k, \ell \leqslant \frac{n}{2}$. For these, the uppermost path leaves the first row at position $(1, k)$, and enters the first column at position $(\ell, 1)$. In the scaling limit, depending on the values of $k = rn$ and $\ell = sn$, the uppermost path will follow a straight line from $(0, r)$ to $(s, 0)$, or will follow a portion of the arctic curve, as illustrated in Figure 5.12. The unrefined case is recovered for $r = s = \frac{1}{2}$.

If we denote by $A_{n+1}(k|\ell)$ the number of such 2-refined ASM of order $n + 1$, the factorization conjecture implies,

$$S(r,s) \equiv \lim_{n\to\infty} \frac{1}{n} \log \frac{A_{n+1}(rn|sn)}{A_n} = S[f^*] = \int_0^1 \mathrm{d}x\, L\Big(\frac{\mathrm{d}f^*}{\mathrm{d}x}\Big), \quad 0 \leqslant r, s \leqslant \frac{1}{2}, \tag{5.88}$$

where $f^*$ is the trajectory followed by the uppermost path in the scaling limit: it goes from $f^*(0) = 1$ to $f^*(1) = 0$ by passing through $f^*(0) = r$, $f^*(s) = 0$ and minimizes the distance without ever crossing the arctic curve. The relevant function $L$ is the one that pertains to the paths used for the ASM.

For multirefined partition functions, we would consider the number $A_{n+m}(k_1, \ldots, k_m|\ell_1, \ldots, \ell_m)$ of ASM such that in the scaling limit, their $j$-th uppermost path leaves the top boundary at $(0, r_j)$ and reaches the left boundary at $(s_j, 0)$, with $0 \leqslant r_1 \leqslant r_2 \leqslant \ldots \leqslant r_m \leqslant \frac{1}{2}$ and $0 \leqslant s_1 \leqslant s_2 \leqslant \ldots \leqslant s_m \leqslant \frac{1}{2}$. Then a formula similar to (5.51) is conjectured to hold, where the actions $S(r_i, s_i)$ for the individual paths are simply added.

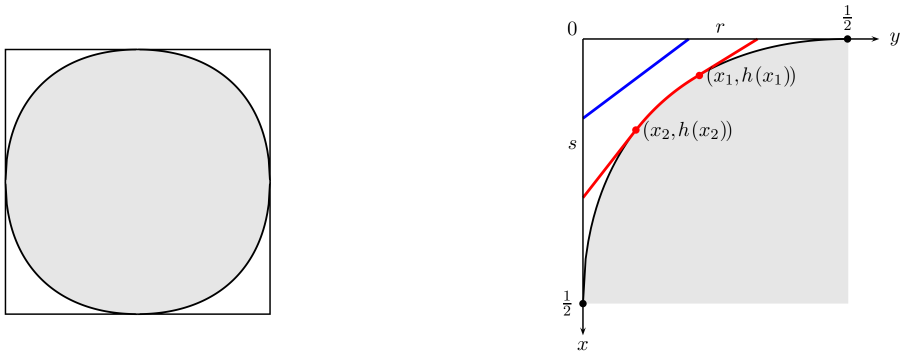



Figure 5.12: *Left:* Arctic curve of an ASM with the disordered region shaded in grey. *Right:* Magnification of the north-west quadrant. There are two different possibilities for the maximizing trajectory $f^*$: either it touches or it does not touch the arctic curve, depending on the sign of $2r+2s-rs-1$.

### 5.5.1 The 2-refined ASM entropy function

The paths that make up the correspondence with ASM are composed of two elementary steps $(1,0)$ and $(0,-1)$, both with weight 1. The appropriate Lagrangean function can be inferred from $L_1(t)$ given in (5.5),

$$L(t) = (1-t)\log(1-t) + t\log|t|, \qquad t \leqslant 0. \tag{5.89}$$

It satisfies $L(0) = 0$ and $\lim_{t\to-\infty} L(t) = 0$.

To compute the function $S(r,s)$, we first note that the horizontal and vertical portions of $f^*$ do not contribute to the integral since $L(0) = L(-\infty) = 0$ as we have just noted. If $r$ and $s$ are close enough to $\frac{1}{2}$, the trajectory $f^*$ is rectilinear from $(0,r)$ to a first tangency point $\big(x_1(r), h(x_1(r))\big)$, then follows the arctic curve till a second tangency point $\big(x_2(s), h(x_2(s))\big)$, where it tangentially leaves the arctic curve and moves rectilinearly to the endpoint $(s,0)$. Fixing $r$ and $s$, one finds, for $t_i = h'(x_i)$,

$$t_1 = -\frac{r(2-r)}{1-2r},\ x_1 = 1-\frac{1-2t_1}{2\sqrt{1-t_1+t_1^2}},\ t_2 = -\frac{1-2s}{s(2-s)},\ x_2 = 1-\frac{1-2t_2}{2\sqrt{1-t_2+t_2^2}}, \tag{5.90}$$

where $t_1 < t_2$ are both negative. As $r$ and $s$ decrease, the slopes $t_1$ and $t_2$ respectively increases and decreases, until they become equal, implying also that $x_1 = x_2$. If $r,s$ decrease even further, the trajectory $f^*$ is a single straight line going from $(0,r)$ to $(s,0)$. The borderline between the two cases corresponds to $t_1(r) = t_2(s)$, or $2r + 2s - rs = 1$, which, in the $(r,s)$-plane, is a concave curve slightly above the antidiagonal $2r + 2s = 1$.

When $r, s$ are such that $t_1(r) \leqslant t_2(s)$, we have

$$S(r,s) = S[f^*] = x_1 L(t_1) + \int_{x_1}^{x_2} \mathrm{d}x L\big(h'(x)\big) + (s - x_2) L(t_2), \qquad (5.91)$$

which, like the similar calculation in the Aztec diamonds, reduces to a simple symmetric function of $r, s$. For $t_1(r) \geqslant t_2(s)$, $f^*$ is a straight line so that the integral is straightforward.

The final results read, for $0 \leqslant r, s \leqslant \frac{1}{2}$,

$$S(r,s) = \begin{cases} \mathcal{L}(r+s) - \mathcal{L}(r) - \mathcal{L}(s), & 2r+2s-rs \leqslant 1, \\ \mathcal{L}(1+r) - \mathcal{L}(r) + \mathcal{L}(2-r) - \mathcal{L}(1-r) & \\ +(r \to s) - \mathcal{L}(3), & 2r+2s-rs \geqslant 1. \end{cases} \qquad (5.92)$$

As a first simple check, we consider the unrefined partition functions for ASM of order $n+1$ and order $n$. Since the unrefined case corresponds to $r = s = \frac{1}{2}$, the factorization implies

$$\lim_{n \to \infty} \frac{1}{n} \log \frac{A_{n+1}}{A_n} = \lim_{n \to \infty} \frac{1}{n} \log \frac{A_{n+1}(\frac{n}{2}|\frac{n}{2})}{A_n} = S(\tfrac{1}{2}, \tfrac{1}{2}) = \log \tfrac{27}{16}, \qquad (5.93)$$

where we have used the second form in (5.92). From (5.86), the exact ratio is equal to $A_{n+1}/A_n = \binom{3n+1}{n} / \binom{2n}{n}$, and is asymptotic to $\left(\frac{27}{16}\right)^n$, in agreement with the previous value.

Another limiting case is when the refined matrices in $\mathrm{ASM}_{n+1}$ have $m_{1,1} = 1$, in such a way that this 1 is shared between the first row and the first column. It corresponds to $k = \ell = 1$, and $r = s = 0$ in the scaling limit. As the number of such refined matrices in $\mathrm{ASM}_{n+1}$ is equal to the unrefined matrices in $\mathrm{ASM}_n$, the ratio $A_{n+1}(1|1)/A_n$ is exactly equal to 1, implying $S(0,0) = 0$. This is also the value obtained from the first form of (5.92).

### 5.5.2 Refined alternating sign matrices

In this section, we study a selection of multirefined enumerations of ASM. A quite extended literature exists on the subject, see for instance [169]. The results recalled here are collected in [170].

The simplest case is that of the 1-refined enumeration of ASM, by which one counts the number $A_n(k|-)$ of matrices in $\mathrm{ASM}_n$ having their unique 1 in the first row at position $k$, without any condition on the unique 1 in the first column. The result is surprisingly simple [171]

$$A_n(k|-) = \binom{n+k-2}{k-1} \frac{(2n-k-1)!}{(n-k)!} \prod_{j=0}^{n-2} \frac{(3j+1)!}{(n+j)!}. \tag{5.94}$$

It follows that the effective contribution of the uppermost path is given by

$$Z_1^{\mathrm{out}}(k|-) = \frac{A_{n+1}(k|-)}{A_n} = \frac{\binom{n+k-1}{n}\binom{2n-k+1}{n}}{\binom{2n}{n}}, \tag{5.95}$$

from which, upon setting $k = rn$ with $r \leqslant \frac{1}{2}$, one readily obtains

$$\lim_{n\to\infty} \frac{1}{n} \log Z_1^{\mathrm{out}}(rn|-) = \mathcal{L}(1+r) - \mathcal{L}(r) + \mathcal{L}(2-r) - \mathcal{L}(1-r) - \mathcal{L}(2). \tag{5.96}$$

As no condition on the rescaled position $s = \ell/n$ of the unique 1 in the first column implies $s = \frac{1}{2}$ almost surely, the previous limit is equal to $\lim_{n\to\infty} \frac{1}{n} \log Z_1^{\mathrm{out}}(rn|\frac{n}{2})$, expected to be equal to the function $S(r, \frac{1}{2})$ computed in (5.92), which it is.

There are two other refinements of ASM which are related to the 1-refined partition functions we have just discussed. The first one is the so-called Top-Bottom 2-refined enumeration $A_n^{\mathrm{TB}}(k, k')$, which is the number of matrices in $\mathrm{ASM}_n$ with their unique 1 in the top and bottom rows respectively in columns $k$ and $k'$. Strictly speaking, one cannot handle it within a single path description. However, because the two refinements concern regions of the domain which are far away, one would expect that they are weakly correlated, in such way that the following factorization would hold,

$$\frac{A_{n+2}^{\mathrm{TB}}(k, k')}{A_n} = \frac{A_{n+2}^{\mathrm{TB}}(k, k')}{A_{n+1}^{\mathrm{B}}(k')} \frac{A_{n+1}^{\mathrm{B}}(k')}{A_n} \approx \frac{A_{n+2}^{\mathrm{T}}(k)}{A_{n+1}} \frac{A_{n+1}^{\mathrm{B}}(k')}{A_n}. \tag{5.97}$$

In the scaling limit, and for $k = rn$ and $k' = r'n$ with $r, r' \leqslant \frac{1}{2}$, it leads to two independent 1-refined contributions,

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{A^{\mathrm{TB}}_{n+2}(rn, r'n)}{A_n} = S(r, \tfrac{1}{2}) + S(r', \tfrac{1}{2}). \tag{5.98}$$

The exact evaluation of $A^{\mathrm{TB}}_n(k, k')$ was carried out in [172, 173], and rewritten in [170] but the result does not easily lend itself to an asymptotic analysis. The previous equality can nonetheless be checked numerically for large $n$. The comparison is presented in the left panel of Figure 5.13, which shows a good agreement[7].

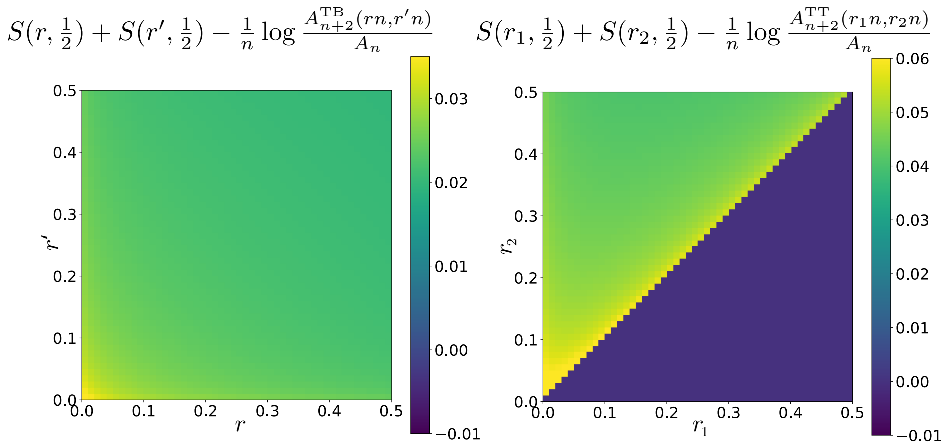



Figure 5.13: Verification of the factorization conjecture for two 2-refined enumeration of ASM, using exact lattice results. *Left:* Top-Bottom refinement, with $n = 400$. *Right:* Top-Top refinement, with $n = 200$.

The second refinement is called Top-Top, and is based on the fact that there is a unique pair of column labels, $k_1 < k_2$, such that $m_{1,k_1} + m_{2,k_1} = m_{1,k_2} + m_{2,k_2} = 1$ [174]. Let us denote by $A^{\mathrm{TT}}_{n+2}(k_1, k_2)$ the corresponding refined partition function. In terms of osculating paths, it means that the most upper path makes its first or second step rightwards at column $k_1$, and that the second most upper path makes its first step rightwards at column $k_2$. For large $n$, the partition function $A^{\mathrm{TT}}_{n+2}(k_1, k_2)$ can be assimilated to $A_{n+2}(k_1, k_2|-,-)$ defined earlier. For $k_1 = r_1 n$ and $k_2 =$

[7]We have used the exact expressions from [170], after correcting a typo in their equation (1.3), where the first term on the right-hand side should read $A_{n-1,j-i}$ instead of $A_{n,j-i}$.

$r_2 n$ with $r_1 \leqslant r_2 \leqslant \frac{1}{2}$, the factorization conjecture predicts the same result as for the Top-Bottom case,

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{A^{\mathrm{TT}}_{n+2}(r_1 n, r_2 n)}{A_n} = \lim_{n\to\infty} \frac{1}{n} \log \frac{A_{n+2}(r_1 n, r_2 n|-,-)}{A_n} \\ = S(r_1, \tfrac{1}{2}) + S(r_2, \tfrac{1}{2}). \tag{5.99}$$

The known determinations of $A^{\mathrm{TT}}_{n+2}(k_1, k_2)$ [175, 176] do not allow to easily extract its asymptotics, so that we again relied on numerical evaluations. The right panel of Figure 5.13 confirms the validity of this identity.

Another interesting 2-refinement corresponds to the enumeration of $\mathrm{ASM}_n$ with prescribed positions of the unique 1 in the first row and first column, a case also referred to as Top-Left [170]. If $m_{1,k} = 1$ and $m_{\ell,1} = 1$ for $1 \leqslant k, \ell \leqslant \frac{n}{2}$, it exactly corresponds to the refined partition function $A_n(k|\ell)$. The factorization conjecture would then imply that the limit $\lim_{n\to\infty} \frac{1}{n} \log A_{n+1}(rn|sn)/A_n$ is given by $S(r,s)$ in (5.92), with a different behaviour below or above the curve $2r + 2s - rs = 1$.

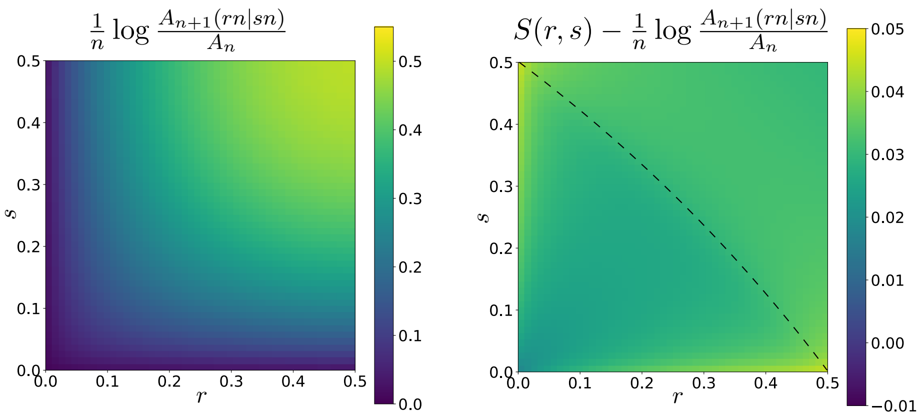



Figure 5.14: Numerical verification of the factorization conjecture for the Top-Left refinement of ASM for $n = 200$. *Left:* Exact estimation of $\frac{1}{n} \log \frac{A_{n+1}(rn|sn)}{A_n}$. *Right:* Difference with the predicted value $S(r,s)$ of the exponential growth rate, obtained from the factorization conjecture. Corrections below or above $2r + 2s - rs = 1$ (dashed curve) appear to be different.

The exact value of $A_n(k|\ell)$ was computed in [173], but again in a form that does not allow for an asymptotic estimate. A numerical verification is presented in Figure 5.14, which confirms the previous limit. Complicated exact formulae for 3- and 4-refinements exist in the form of generating functions [170], but their numerical evaluation is impractical for large $n$.

We have carried out similar checks on classes of ASM with symmetries, in particular the 1-refined partition functions $A_{\mathrm{V}}(2m+1,k)$ and $A_{\mathrm{HT}}(2m,k)$, counting respectively the vertically symmetric alternating sign matrices (VSASM) of order $2m+1$ with a 1 at position $k \leqslant m$ on the second row (see Figure 5.11-left for an instance of VSASM), and the half-turn symmetric alternating sign matrices (HTSASM) of order $2m$ with a unique 1 on the first row in column $k$. Their exact expressions are both given in [177], from which the asymptotic behaviour can be determined analytically. In the case of VSASM, the arctic curve is the same [22] as for ordinary ASM and it is also expected to be the case for HTSASM. In both cases, using the ASM arctic curve we check that the factorization conjecture is fulfilled.

In addition to the above checks of the factorization conjecture, many numerical checks, not reported here, based on the 1-refinements considered in [14, 15], have been successfully carried out for the six-vertex model, in the D and AF phases (and for values of $\Delta$ up to 0.998). These checks support the validity of the factorization conjecture, even in the attractive case mentioned above. The only numerical part in these verifications is the evaluation of fairly complicated integrals (that involve theta functions in the AF case). In the D regime, it should in principle be possible to evaluate these integrals explicitly since a closed form formula for the asymptotics of the one-point function was recently found in [30].

## 5.6 The tangent method revisited

This last section takes advantage of the factorization property to reformulate the tangent method in a simpler way. The core of the tangent method is to use a 1-refined partition function to compute a family of straight lines tangent to the arctic curve, which is in turn determined from it. In its usual form, information is extracted from the 1-refinement by moving the starting point of the outermost path from its original location to another point in an *extension* of the domain. The equation of the corresponding tangent line is then obtained by computing the likeliest entry point in the original domain, usually through the saddle point method. Here we show how the factorization conjecture allows to compute the family of tangent lines directly from the 1-refinements, thereby avoiding the calculation of the likeliest entry point and the saddle point analysis. It has the advantage of being directly applicable to any 1-refinement including those that *do not call for an extension* of the original domain.

We consider the generic situation shown in Figure 5.15, in which one intends to determine the portion of the arctic curve $y = h(x)$ between two contact points, $C_1$ and $C_2$. The white region below the arctic curve is supposed to be void of paths. The 1-refinement consists in moving the starting of the most external path from $C_1$ to the point $(0, r)$. The new external path will then look like the red curve, which comprises a straight segment hitting the arctic curve tangentially at $(x^*, h(x^*))$, with a slope $t^* = h'(x^*)$, before following the arctic curve towards $C_2$. Varying $r$ from 0 to $C_1$ yields a family of tangents to the arctic curve, which can be retrieved by solving the system of equations

$$y = t^* x + r, \qquad \frac{\mathrm{d}}{\mathrm{d}r}\big[t^* x + r\big] = 0, \tag{5.100}$$

where $t^* = t^*(r)$ is viewed as a function of $r$. The problem is to compute the function $t^*(r)$.

Denoting respectively by $Z_{n+1}(nr)$ and $Z_n$ the 1-refined and unrefined partition functions, we consider the effective contribution in the scaling limit of the outermost red path $f^*$ as given in terms of the action associated to $f^*$,

$$S(r) = \lim_{n\to\infty} \frac{1}{n} \log \frac{Z_{n+1}(nr)}{Z_n} = \int_0^{C_2} \mathrm{d}x\, L\Big(\frac{\mathrm{d}f^*}{\mathrm{d}x}\Big) = x^* L(t^*) + \int_{x^*}^{C_2} \mathrm{d}x\, L\big(h'(x)\big). \tag{5.101}$$

In this equation, the function $S(r)$ is seen as an input, obtained from independent lattice calculations, and $L$ is the Lagrangean associated to the particular lattice paths under consideration.

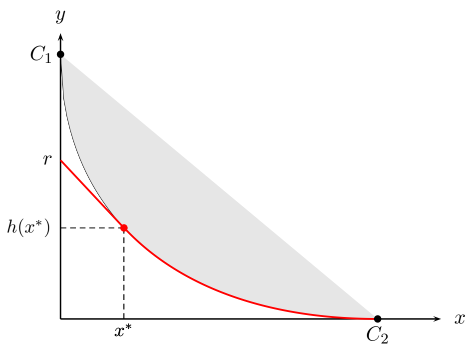



Figure 5.15: Typical configuration of a 1-refined calculation, shown in the scaling limit. A path starts from the point of coordinates $(0, r)$ and ends at the contact point $C_2$. The path is first a straight segment, which joins the arctic curve tangentially, at $(x^*, h(x^*))$, and follows it till $C_2$. The arctic curve is functionally given by $h(x)$.

To determine $t^*(r)$, we differentiate (5.101) with respect to $r$,

$$\frac{\mathrm{d}S(r)}{\mathrm{d}r} = x^* \frac{\mathrm{d}}{\mathrm{d}r} L(t^*) = x^* L'(t^*)\, h''(x^*)\, \frac{\mathrm{d}x^*}{\mathrm{d}r}. \tag{5.102}$$

Since the point $(x^*, h(x^*))$ belongs to the tangent to the arctic curve at $x^*$, given by $y = t^* x + r$, it follows that

$$r = h(x^*) - t^* x^* \quad \Longrightarrow \quad \frac{\mathrm{d}r}{\mathrm{d}x^*} = -h''(x^*) x^*. \tag{5.103}$$

From this, we readily obtain

$$\frac{\mathrm{d}S(r)}{\mathrm{d}r} = -L'(t^*), \qquad \text{or} \quad t^*(r) = (L')^{-1}\Big(-\frac{\mathrm{d}S(r)}{\mathrm{d}r}\Big), \tag{5.104}$$

provided the function $L'(t)$ is invertible. For the models formulated in terms of paths made of elementary steps that are weighted independently of each other, we proved that the function $L$ is strictly concave in Subsection 5.1.1. Moreover, one has $-L'(t) = \log y(t)$ for a positive and strictly increasing function $y(t)$, see (5.21). This class of models includes the Aztec diamond and ASM, but not the general six-vertex model which assigns each elementary step a weight that depends on the following step. Nevertheless, the explicit calculation of $L(t)$ for the general six-vertex model shows that is also strictly concave, see [160]. So assuming the expected strict concavity of $L$, it follows that $L'(t)$ is invertible since it is then strictly decreasing on its domain.

Let us illustrate the above procedure for ASM. The function $S(r)$, for $0 \leqslant r \leqslant \frac{1}{2}$, is given explicitly in (5.96), and $-L'(t) = \log y(t) = \log \frac{1-t}{-t}$ $(t \leqslant 0)$. The condition (5.104) is simple and easily solved,

$$\log \frac{1-r^2}{r(2-r)} = \log \frac{1-t^*}{-t^*} \quad \Rightarrow \quad t^*(r) = \frac{r(2-r)}{2r-1}. \tag{5.105}$$

The two conditions (5.100) yield the parametric form of the arctic curve in a straightforward way,

$$x = \frac{(2r-1)^2}{2(1-r+r^2)}, \quad y = \frac{3r^2}{2(1-r+r^2)}, \qquad 0 \leqslant r \leqslant \frac{1}{2}. \tag{5.106}$$

Eliminating $r$, we recover the ellipse of equation $x(1-x)+y(1-y)+xy = \frac{1}{4}$, found in [13].

# Conclusion

In this thesis, we have studied a geometrical technique allowing to compute the shape of large random objects, solely using some of their combinatorial properties. In Chapters 2 and 4, we have used the tangent method to compute arctic curves, respectively in particular instances of tiling models and vertex models. The fluctuations around a generic point of such curves have been numerically investigated in Chapter 3. The understanding of the tangent method and its natural consequences have been discussed in a variational framework in Chapter 5.

In Chapter 1, we have motivated the study of the arctic curve phenomenon and showed a hint of its appearance in tiling models, namely the dependence of the free energy per site on the boundary conditions. We have also introduced technical tools for enumerating tilings and provided an heuristics to understand the arctic phenomenon in terms of non-crossing lattice paths.

The tangent method has been applied to two different tiling models in Chapter 2. The first one is a tiling of a rectangoloid domain. Despite the freedom in its definition (we are free to choose an arbitrary defect distribution on one boundary), the partition function has a closed form: the Gelfand-Tsetlin formula. The second one is a domino tiling model on a cone, which may be seen as a generalization of descending plane partitions. While the first model is especially well suited for the application of the tangent method, it is a bit more involved for the second model, due to the lack of a closed form formula for the partition function, which is only given in terms of a determinant. Due to a seemingly coincidental connection with the twenty-vertex model, itself related to the six-vertex model, the asymptotics required by the tangent method can still be evaluated.

A numerical analysis has been instigated in Chapter 3. We have reviewed basic aspects of a general method to simulate statistical models, which has been extensively used to illustrate this thesis. In the particular case of domino tilings of the Aztec diamond, there exists an extremely efficient algorithm called the shuffling algorithm. It has been used to numerically study fluctuations around the uppermost point of the arctic curve of this model. Consequences and limitations of a theorem due to Johansson have been investigated. We have verified a conjecture that naturally extends Johansson's result and found universal distributions that originated from random matrix theory.

In Chapter 4, the tangent method has been applied to two different vertex models. In both cases, we have accessed one-refined partition functions via an almost homogeneous limit of inhomogeneous partitions function of those models. Firstly, we have studied the six-vertex model with partial domain wall boundary conditions at the free-fermion point. The second model was the twenty-vertex model with domain wall boundary conditions of type 2, in the disordered regime. The Yang-Baxter relation reduced the computation of the one-refined partition function of this model to that of the six-vertex with domain wall boundary conditions.

A variational approach has been used in Chapter 5 to shed some light on the assumptions of the tangent method. We have shown that the strict concavity of a certain Lagrangean function $L$ implies that a single path constrained to stay in a fixed domain is almost surely the shortest path between its starting and ending points, in the scaling limit. We have proposed a factorization conjecture for models which have a formulation in terms of non-crossing paths. The conjecture states that, in the scaling limit, the contribution of the few most external paths to the partition function factorizes from that of the bulk paths, responsible for the formation of the arctic curve. Moreover, the contribution of each external path is given by the action associated with $L$, and can be made fully explicit when the arctic curve is known. The factorization has been verified in full generality for tilings of Aztec diamonds, and for certain refinements of alternating sign matrices. In addition to providing a remarkably simple, efficient and explicit way to obtain the asymptotics of multirefined partition functions, the factorization conjecture has been used to provide a novel version of the tangent method that does not rely on an extension of the original domain.

## Outlook

There are two main directions in which the results presented in this thesis could naturally be extended. The first one is the study of arctic curves in other models and the second is the pursuit of the numerical analysis of fluctuations in models that show the limit shape phenomenon.

The arctic curve of the six-vertex model with partial domain wall boundary conditions, at $\Delta = 0$ and $t = 1$, has been determined in Subsection 4.1.4. It was shown to be a real algebraic curve of degree 6, with two symmetric lobes: one is contained in the $n \times s$ domain, the other is symmetrically reflected through the upper boundary, see Figure 5.16 after a 90° rotation. The full curve, with its two lobes, should be the arctic curve of a double copy of the same six-vertex model on the extended $n \times 2s$ domain, whose $n \times s$ upper part is completed by imposing vertical symmetry, whereby horizontal occupied edges are exchanged with empty ones when reflected, while vertical edges remain unchanged after reflection. Relaxing this symmetry is not expected to change the arctic curve but defines a new model. This new model can in turn be put in correspondence with a particular case of domino tilings of a double Aztec rectangle defined in [178] using the explicit construction used in Figure 4.3. In the notation of [178], this domain corresponds to $k = 1$, $m_1 = s - 1$, $n_1 = n - 1$, $m_2 = s$ and $n_2 = n$. The combinatorics of the model is tractable [178] and its refined enumeration is currently under investigation by Jean-François de Kemmeter[8]. Notice that the double Aztec diamonds studied in [179] look very similar, but differ by the position of the starting and ending points, leading to a completely different arctic phenomenon.

[8]Since the completion of this thesis, we have worked out this refined enumeration and have shown that the arctic curve is exactly as conjectured.

The relation between a model and its symmetrized version (the symmetry being imposed on configurations) raises several questions. It appears that when the model itself is symmetric (boundary conditions and weights), imposing this symmetry on configurations does not affect the arctic curve. Despite its intuitive nature, proving that this statement holds in general is not immediate. For tiling models with reflection symmetry, an idea would be to use general results such as Ciucu's factorization theorem [44] to show that the application of the tangent method to the symmetric and non-symmetric model yields the same result. Another interesting problem would be to study the arctic curve appearing when invariance under a given transformation is imposed on configurations of a model whose weights are not symmetric under this transformation.

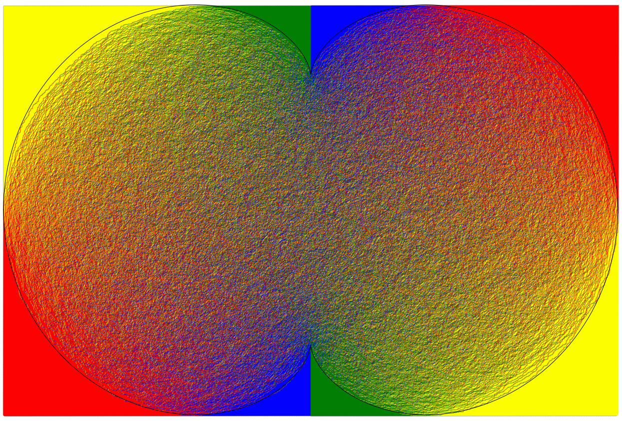


Figure 5.16: Domino tiling of a double Aztec rectangle for $n = 1000$ and $s = 750$. Simulations were performed using a generalization of the shuffling algorithm [86]. The black curve is the analytic continuation of (4.38), which is an algebraic curve of degree six. Courtesy of Jean-François de Kemmeter.

The computation of the arctic curve of the twenty-vertex model with domain wall boundary conditions of type 2 calls for natural generalizations. Clearly, the parametrization (4.44) for the twenty weights of the model covers only a small subset of the allowed values. In particular, the restriction to real values of $\eta$ confines the associated six-vertex model into its disordered phase. The result (4.56) derived in [14] describing the asymptotics of the one-point function $H_n(\sigma)$ was generalized in [15] for the anti-ferroelectric regime. It should therefore be possible to extend our results to this regime, namely, in the parametrization (4.44), to the case of imaginary values of $\eta$, $\lambda$ and $\mu$. The expressions of [15] involve elliptic functions and this extension might in practice lead to quite involved calculations. For convenience, we also restricted ourselves to symmetric weights under arrow reversal. This symmetry guarantees that the arctic curve is symmetric under 180° rotation. Contrary to the case of the six-vertex model with fixed boundary conditions, this restriction is an effective one and non-symmetric weights may lead to more general arctic curves without the 180° rotation symmetry, a situation yet to be explored. Another direction of exploration concerns other domain-wall-like boundary conditions such as that recently considered in [30]. The principal difficulty in such an endeavour is computing the asymptotics of the one-refined partition function of the six-vertex model obtained after performing the unravelling. Finally, a number of non-intersecting path problems were solved in a more general $q$-deformed framework, which consists in introducing an extra path weight $q^{\mathcal{A}}$ involving the area $\mathcal{A}$ below the path. In particular, the tangent method was applied successfully to some of these models [25, 27] to obtain the corresponding $q$-deformed arctic curve and we may wonder whether this generalization can be carried out for vertex models.

The numerical study of the fluctuations in domino tilings of the Aztec diamond around $(0, \frac{1}{\sqrt{2}})$ only scratched the surface of a rich subject. Certain aspects of the fluctuations around other generic points of the arctic curve were studied in [85]. If we zoom on non-generic points, other universal statistics appear, like around cusps [180], contact points with the domain boundary [114] or points of contact between two portions of arctic curve [179]. These various statistics should appear in the double aztec rectangle mentioned above, where the last case is achieved when $s \sim n$. A generalization of the shuffling algorithm [86] allows to efficiently generate such tilings. Generating a large enough collection should be possible, hence enabling an analysis in the same spirit as that of Chapter 3. The same algorithm could also be used to understand the fluctuations in a tangent method set-up, which might exhibit the same statistics as those appearing in [181]. In addition to the statistics around the arctic curve, other fluctuations prove to be interesting. The domino statistics in the temperate regions are believed to be described by some inhomogeneous field theory [56], see also [113] and [182] for other manifestations of the temperate region statistics. Finally, the numerical study of fluctuations of non-free-fermionic models (or even non-integrable ones) would be interesting. It would provide a valuable test for the universality of the various aforementioned statistics, for which exact results are rare and limited to very specific weights [183].

# Appendix A

# Shuffling algorithm

This appendix is based on a seminar given by Antoine Doeraene at KULeuven in June 2018. Its aim is to provide a simple understanding of the shuffling algorithm introduced by Propp and Kuperberg in [63]. The algorithm generates a random tiling of the Aztec diamond with uniform distribution, as defined in Section 1.2. It has many advantages over a Markov chain approach: it automatically stops, it is particularly fast and it is guaranteed to exactly produce the uniform distribution, under the assumption that the computer's pseudo-random number generator is perfect.

## A.1 Description of the algorithm

Instead of formulating the algorithm in terms of domino tilings of the Aztec diamond of order $n$, it will be more convenient to describe it in terms of perfect matching of its dual graph $G_n$, see Figure A.1-left. A perfect matching is a set of edges that cover every vertex of the graph once. Let $M_n$ be the set of perfect matchings $\mathcal{M}$ of $G_n$, which are clearly in bijection with domino tilings of the Aztec diamond, see Figure A.1-middle. Amongst the square faces of the dual graph, we distinguish the active faces, that will play a crucial role in the algorithm. By definition, the faces on the boundary are active and the active faces alternate in a chess board manner, see Figure A.1-right.

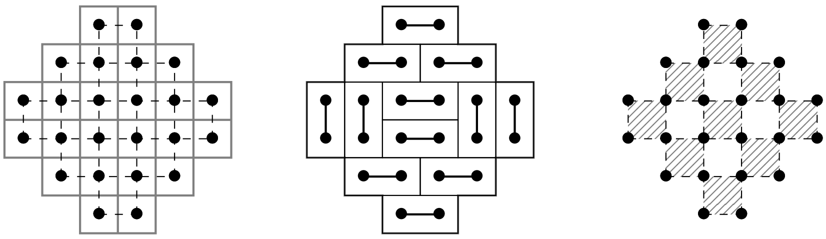


Figure A.1: *Left:* Dual graph (dashed lines) whose vertices (bullets) are at the center of the unit squares (in grey) that compose the Aztec diamond of order $n = 3$. *Middle:* Bijection between tilings of the Aztec diamond and perfect matchings of its dual graph. *Right:* The hatched squares are the active faces of the dual graph.

There are only two possible perfect matchings of $G_1$. The first step of the shuffling algorithm is to choose one of them with equal probability. In each subsequent step, a perfect matching of $G_{n+1}$ is generated from one of $G_n$. Each step is a random mapping $\sigma_n : M_n \to M_{n+1}$, that we now describe.

It starts by embedding $\mathcal{M} \in M_n$ into $G_{n+1}$, by placing it at the center of the dual graph of order $n+1$, causing a shift of the active faces in $\mathcal{M}$. Then, the rules $R_1$, $R_2$ and $R_3$ represented on Figure A.2 are applied once to each active face. The rules $R_1$ and $R_2$ are purely deterministic and respectively correspond, in the domino tilings picture, to the translation of dominoes (in a direction that depends on which edge of their active face they lie) and the removal of a pair of dominoes which would otherwise collide if they were to be translated as in $R_1$. The rule $R_3$ is the only probabilistic one and can be seen as the creation of blocks of two aligned dominoes, the two possible orientations being produced with probability $p = \frac{1}{2}$. Figure A.3 shows a possible outcome when the above procedure is applied to a perfect matching of $G_2$. Applying the rules $R_1$-$R_3$ on all active faces produces a perfect matching of $M_{n+1}$, as proved in the following section. Notice that the rules may be applied *independently* on all active faces. In practice, this allows for an efficient implementation.

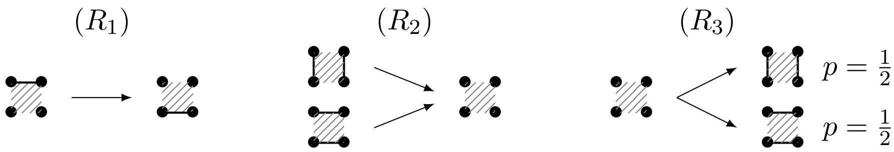



Figure A.2: Rules used (once) on each active faces, for each iteration of the shuffling algorithm. Notice that we only represented one of the four possibilities for the rule $R_1$, the other being obtained by 90° rotations.

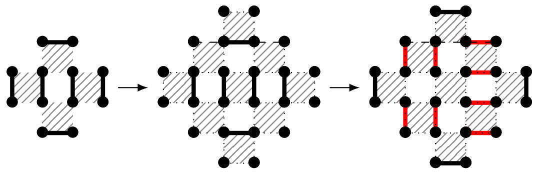


Figure A.3: Possible outcome of $\sigma_2 : M_2 \to M_3$ for a given $\mathcal{M} \in M_2$. Dimers that are newly created via the rule $R_3$ are shown in red.

## A.2 Proof

We now show that the shuffling algorithm works as announced, namely that it outputs a perfect matching $\mathcal{M} \in M_n$ with a probability $P(\mathcal{M} \in M_n) = \frac{1}{|M_n|}$. There are three points to verify:

1. The rules transform a perfect matching in $M_n$ into one in $M_{n+1}$.
2. Any $\mathcal{M}' \in M_{n+1}$ can be produced from a $\mathcal{M} \in M_n$.
3. The perfect matchings are generated uniformly.

We shall verify them separately.

**1. Production of perfect matchings.** For a fixed $\mathcal{M} \in M_n$, we call $\{\sigma_n(\mathcal{M})\}$ the set of all the possible outputs $\mathcal{M}' = \sigma_n(\mathcal{M})$, obtained by using the rule described above on $\mathcal{M}$. We must check that $\{\sigma_n(\mathcal{M})\} \subset M_{n+1}$. To do so we need to verify that each vertex $v$ of $G_{n+1}$ is tied to exactly one edge after the map. We treat two cases separately, depending on the position of $v$.

If $v$ is not on the boundary of $G_{n+1}$ (i.e one of the $4(n+1)$ exterior vertices of coordinate number 2), then it must be adjacent to two active faces $f_1$ and $f_2$. Directly after the embedding of $\mathcal{M}$, there must be exactly one vertex tied to $v$ in one of these two faces, say $f_1$. Due to the rules $R_1$ and $R_2$, any $\mathcal{M}' \in \{\sigma_n(\mathcal{M})\}$ will have no edge in $f_1$ covering $v$. However, due to $R_1$ and $R_3$, $v$ will be covered by an edge of $\mathcal{M}'$ in $f_2$. So after the mapping $v$ is covered by only one edge.

On the other hand, if $v$ is on the boundary of $G_{n+1}$, then no link is adjacent to it after the embedding. However, according to $R_1$ and $R_3$ one adjacent link will be occupied after the mapping.

We have shown that if $\mathcal{M}$ is a perfect matching of $G_n$, then $\mathcal{M}' \in \{\sigma_n(\mathcal{M})\}$ is a perfect matching of $G_{n+1}$. Since the algorithm starts from one of the two possible perfect matchings of $G_1$, it iteratively produces perfect matchings on dual graphs of increasing sizes $G_2, G_3, \cdots, G_n$.

**2. Pseudo-surjectivity of $\sigma_n$.** To understand why any $\mathcal{M}' \in M_{n+1}$ can be generated by applying $\sigma_n$ on some $\mathcal{M} \in M_n$, we consider the pseudo-inverse map $\sigma_n^{-1} : M_{n+1} \to M_n$. For $\mathcal{M}' \in M_{n+1}$, a possible outcome $\mathcal{M} = \sigma_n^{-1}(\mathcal{M}')$ is by definition obtained by *first* applying the same rules $R_1$-$R_3$ on $\mathcal{M}'$ and *then* restricting the result on the (smaller) dual graph $G_n$.

Such a restriction is always possible: after applying $R_1$-$R_3$ on $\mathcal{M}'$, one obtains a perfect matching of $G_n$, embedded in $G_{n+1}$. Indeed, for every vertex $v$ that is not on the boundary of $G_{n+1}$, the previous argument implies that $v$ is still covered by only one link after applying $R_1$-$R_3$. If $v$ is on the boundary of $G_{n+1}$, it is occupied before application of $R_1$-$R_3$, since $\mathcal{M}' \in M_{n+1}$. Then, due to $R_1$ and $R_2$, $v$ becomes unoccupied. So one can restrict the result to $G_n$ and obtain $\mathcal{M} \in M_n$.

By inspecting the rules $R_1$-$R_3$, it is not hard to see that there is a non-zero probability to recover $\mathcal{M}'$ by applying $\sigma_n$ on $\mathcal{M} = \sigma_n^{-1}(\mathcal{M}')$. This property explains why we called $\sigma_n^{-1}$ a pseudo-inverse. It also proves that $\sigma_n$ is pseudo-surjective, namely it has a non-zero probability of producing any $\mathcal{M}' \in M_{n+1}$, as required

**3. Uniform distribution.** We call $P_n(\mathcal{M})$ the probability of obtaining a perfect matching $\mathcal{M} \in M_n$ after the successive operations of $\sigma_j$, $j = 1, \cdots, n-1$. We want to prove that $P_n(\mathcal{M})$ does not depend on $\mathcal{M} \in M_n$. Gathering the results of the previous points this implies that the shuffling algorithm produces any $\mathcal{M} \in M_n$ with probability $P_n(\mathcal{M}) = P_n = 1/|M_n|$, hence proving the claim.

The proof that $P_n(\mathcal{M}) = P_n$ is done by induction on $n$. For $n = 1$ we generate the two possible configurations of $G_1$ with probability $p = \frac{1}{2}$, which is indeed the uniform distribution.

We now suppose that $P_{n-1}(\mathcal{M})$ is uniform. Then we have

$$P_n(\mathcal{M}) = \sum_{\mathcal{M}^* \in \{\sigma_{n-1}^{-1}(\mathcal{M})\}} P_{n-1}(\mathcal{M}^*) P(\mathcal{M}^* \to \mathcal{M}) = P_{n-1} \sum_{\mathcal{M}^* \in \{\sigma_{n-1}^{-1}(\mathcal{M})\}} P(\mathcal{M}^* \to \mathcal{M}), \tag{A.1}$$

with $P(\mathcal{M}^* \to \mathcal{M})$ the transition probability from $\mathcal{M}^*$ to $\mathcal{M}$ by applying $\sigma_{n-1}$. The sum over the possible pre-images $\mathcal{M}^* \in \{\sigma_{n-1}^{-1}(\mathcal{M})\}$ is non-empty due to the pseudo-surjection. It remains to prove that, despite the fact that each term of the last sum of (A.1) depends on $\mathcal{M}$, the overall sum does not. Let us respectively denote by $\#\mathrm{FAF}(\mathcal{M})$ and $\#\mathrm{EAF}(\mathcal{M})$ the number of full and empty active faces in $\mathcal{M}$. We show that the sum does not depend on $\mathcal{M}$ in three steps:

(a) $\forall \mathcal{M}^* \in \{\sigma_{n-1}^{-1}(\mathcal{M})\}$, $P(\mathcal{M}^* \to \mathcal{M}) = 1/2^{\#\mathrm{FAF}(\mathcal{M})}$,

(b) $|\{\sigma_{n-1}^{-1}(\mathcal{M})\}| = 2^{\#\mathrm{EAF}(\mathcal{M})}$,

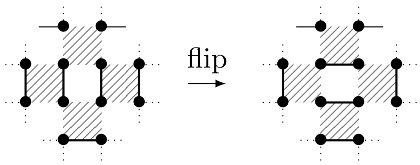
(c) $\#\mathrm{EAF}(\mathcal{M}) - \#\mathrm{FAF}(\mathcal{M})$ is independent of $\mathcal{M}$.

We prove every point separately:

(a) The only random part of the construction comes from $R_3$, which generates the full active faces in $\mathcal{M}$, for which a coin flip was done. Hence, $P(\mathcal{M}^* \to \mathcal{M}) = 1/2^{\#\mathrm{FAF}(\mathcal{M})}$.

(b) All the possible $\mathcal{M}^* \in \{\sigma_{n-1}^{-1}(\mathcal{M})\}$ can only differ by their full non-active faces, which become full active faces when embedded. Since there are two possibilities for every full non-active face of $\mathcal{M}^*$, and since each one of them transform into an empty active face of $\mathcal{M}$ via $R_2$, we get $|\{\sigma_{n-1}^{-1}(\mathcal{M})\}| = 2^{\#\mathrm{EAF}(\mathcal{M})}$.

(c) Since every $\mathcal{M}_1 \in M_n$ and $\mathcal{M}_2 \in M_n$ are related by a finite sequence of flips (as explained in Section 1.2), we only need to prove that a single flip doesn't change $\#\mathrm{EAF}(\mathcal{M}) - \#\mathrm{FAF}(\mathcal{M})$. If the flip occurs on an active face, the number of full active faces remains the same and so does the number of empty ones. If it occurs in a non-active face $f$, we check the different possible cases one by one. By definition, $f$ cannot be on the boundary and must be surrounded by active faces. An example of flip is:




$\#\mathrm{FAF} = 2$, $\#\mathrm{EAF} = 1$ $\qquad$ $\#\mathrm{FAF} = 1$, $\#\mathrm{EAF} = 0$

The other cases are similar. The proof is hence complete.

**The partition function as a side product.** By making the previous computation slightly more explicit, we readily find the partition function $Z_n = P_n^{-1} = |M_n| = 2^{n(n+1)/2}$.

*Proof.* The proof is by induction. It is true for $n = 1$. Suppose it is true for $n - 1$. We wish to evaluate a bit more explicitly (A.1). Since $\#\mathrm{EAF}(\mathcal{M}) - \#\mathrm{FAF}(\mathcal{M})$ does not depend on $\mathcal{M}$, we can chose a particular perfect matching to evaluate it. For $\mathcal{M}_h$ whose edges are all horizontal, there are no empty active faces, and the only full active faces are the one in the horizontal line in the middle of $G_n$. Hence, we have

$$P_n = P_{n-1}\, 2^{\#EAF(\mathcal{M}_h) - \#FAF(\mathcal{M}_h)} = \frac{2^{0-n}}{2^{n(n-1)/2}} = \frac{1}{2^{n(n+1)/2}}. \tag{A.2}$$

■

What we just presented as a side product was actually the original purpose of the shuffling algorithm, which was invented solely for enumerating domino tilings of the Aztec diamond [63, 65]. It is only later that it was implemented by Sameera Iyengar, an undergraduate student of Propp, whose simulations hinted toward the existence of the arctic phenomenon [65]. It was also by an analysis of the behaviour of the shuffling algorithm that the arctic circle theorem was established [31, 65].

A generalization of the shuffling algorithm allows to generate domino tilings of the Aztec diamond, with position-dependent weighting [86]. In particular, these weights can vanish. Hence, the algorithm allows to generate tilings on other domains, provided that they can be embedded in an Aztec diamond in such a way that its complement can be tiled as well. Antoine Doeraene implemented this algorithm for several such domains; it can be run on the website https://sites.uclouvain.be/aztecdiamond/.

# Appendix B

# Technicalities: 6V model with pDWBC

In this appendix, we gather some technicalities in the computation of the arctic curve of the 6V model with pDWBC at $\Delta = 0$ and $t = 1$. The core issue is the evaluation of sums involving factorials. There are several ways to deal with these sums. We will use two of them in this appendix. The first one exploits generating function and is based on lemmata 1.2.3 (generating function of binomials) and 1.2.2 (product of generating functions), in the spirit of what is done in [22]. The second approach is an automatic way of verifying such identity, based on Zeilberger's algorithm [184].

## B.1 Homogeneous limit

In [145], they show that in the free-fermion case ($\Delta = 0$), the partition function of the vertically inhomogeneous 6V model with pDWBC is

$$Z_{n,s}(t_1, \cdots, t_s) = \prod_{i=1}^{s} \frac{c_i}{1-t_i} \prod_{1\le i<j\le s} \frac{1+t_it_j}{(1-t_it_j)(t_j-t_i)} \det_{1\le i,j\le s} [f_i(t_j)] \,, \tag{B.1}$$

where the inhomogeneous weights are $a_j = 1$, $t_j = b_j$ and $c_j = \sqrt{1+t_j^2}$ and

$$f_i(t_j) = t_j^{i-1} - {t_j}^{n+s-i}. \tag{B.2}$$

Let us show how to take the homogeneous limit $t_i \to t = 1$. When we successively take the limits $t_1 \to 1$, $t_2 \to 1$, $t_3 \to 1$, *etc.* the prefactor develops poles of odd orders $\frac{1}{1-t_1}$, $\frac{1}{(1-t_2)^3}$, $\frac{1}{(1-t_3)^5}$, *etc.*. We will show that these poles are exactly matched by zeros of equal orders, coming from the dominant contribution in the determinant once the corresponding column is Taylor expanded

$$f_i(t_j) = f_i(1) + f_i^{(1)}(1)(t_j - 1) + \frac{f_i^{(2)}(1)}{2!}(t_j - 1)^2 + \cdots, \tag{B.3}$$

where

$$f_i^{(m)}(1) = \frac{(i-1)!}{(i-1-m)!} - \frac{(n+s-i)!}{(n+s-i-m)!}. \tag{B.4}$$

The (unusual) cancellation of zeros of even order stems from the identity

$$\frac{f_i^{(m)}(1)}{m!} = \frac{1}{2}\sum_{k=1}^{m-1}(-1)^{k+1}\binom{n+s-m-1-k}{k}\frac{f_i^{(m-k)}(1)}{(m-k)!} \tag{B.5}$$

that holds when $m$ is even. This relation follows from

$$\begin{aligned}
&\sum_{k=0}^{m}(-1)^{k+1}\binom{n+s-m-1-k}{k}\left[\binom{i-1}{m-k} - \binom{n+s-i}{m-k}\right] \\
&= \sum_{k=0}^{m} \frac{1}{(1+x)^{n+s-m}}\bigg|_{x^k} \left[(1+x)^{n+s-i} - (1+x)^{i-1}\right]\big|_{x^{m-k}} \\
&= \left[\frac{1}{(1+x)^{i-m}} - \frac{1}{(1+x)^{n+s-m-i+1}}\right]\bigg|_{x^m} \\
&= (-1)^m\left\{\binom{i-1}{m} - \binom{n+s-i}{m}\right\},
\end{aligned} \tag{B.6}$$

that is valid for any $m$. When $m$ is even, the relation can be used get (B.5).

Identity (B.5) states that any $f_i^{(m)}(1)$ with even $m$ is a linear combination (with coefficient that do not depend on $i$) of $f_i^{(\ell)}(1)$ with $\ell < m$ odd. When $m$ is odd however, since $f_i^{(m)}(1)$ is a polynomial in $i$ of order $m$, it cannot be written as a linear combination of $f_i^{(\ell)}(1)$ with $\ell < m$. Using these properties, it is now easy to evaluate the homogeneous limit and one gets

$$Z_{n,s}(1, \cdots, 1) = -2^{s^2/2}\prod_{k=1}^{s}\frac{1}{(2k-1)!}\det_{1\le i,j\le s}\left[f_i^{(2j-1)}(1)\right]. \tag{B.7}$$

## B.2 Proof of the LU decomposition

The matrix we want to decompose has matrix elements ($s \leqslant n$ are positive integers)

$$B_{ij} = \frac{(s-i)!}{(s+1-i-2j)!} - \frac{(n+i-1)!}{(n+i-2j)!}, \qquad 1 \leqslant i,j \leqslant s. \tag{B.8}$$

The lower triangular matrix $L$ has been conjectured as well as its inverse, which is given by

$$L^{-1}_{ij} = (-1)^{i+j} \binom{i-1}{j-1} \frac{(n-s+2i-1)!\,(n-s+j-1)!}{(n-s+i+j-1)!\,(n-s+i-1)!}. \tag{B.9}$$

The proof of the conjecture reduces to show that $U \equiv L^{-1}B$ is upper triangular. To do this, one introduce the following matrix

$$\begin{aligned} V_{ij}(s) = \sum_{k=1}^{s} (-1)^{i+k} & \frac{(n-s+k-1)!}{(k-1)!\,(i-k)!\,(n-s+i+k-1)!} \\ & \times \left[\frac{(s-k)!}{(s+1-k-2j)!} - \frac{(n+k-1)!}{(n+k-2j)!}\right], \end{aligned} \tag{B.10}$$

related to $U$ by $U_{ij} = -U_{ii}V_{ij}$ where $U_{ii} = -(i-1)! \prod_{l=i}^{2i-1}(n+l-s)$. So, proving that $U$ is upper triangular is the same as proving that $V$ is upper triangular; moreover one would like to show that the diagonal elements of $V$ are all equal to $-1$, as this would confirm the conjectured diagonal values of $U$.

That $V$ is upper triangular will be proved by recurrence: if $V(s_0)$ and $V(s_0+1)$ are upper triangular for some $s_0$, then $V(s)$ is upper triangular for all $s \geqslant s_0$. This is done by using recurrence relations generated by Zeilberger's algorithm. As these can be sometimes hard to verify explicitly, this is a kind of black box proof. A nice feature of Zeilberger's algorithm is that it can generate recurrence relations for $V_{ij}(s)$ with respect to any of the parameters, here $i, j, s$ and $n$.

Our first recurrence relation reads

$$\begin{aligned} &(s-2j)(n-s+1)V_{ij}(s) \\ &-\left(-2ij-in+3is-3i-2jn+2js-2j+2ns-n-2s^2+3s-1\right)V_{ij}(s-1) \\ &+(s-1-i)(2i+n-s+1)V_{ij}(s-2) = 0, \end{aligned} \tag{B.11}$$

which is valid for $i \leqslant s-1$ and $j \geqslant 1$. To understand the range of validity of this recurrence, we may observe that the above definition of $V_{ij}(s)$ makes sense for all values of $i, j$ in $\mathbb{Z}$ (though the infinite matrix $V(s)$ obtained this way is neither 0 for any $j < i$ nor equal to $-1$ for $i = j$). The recurrence (B.11) is more generally valid for all $i, j \in \mathbb{Z}$ (and $s$ a positive integer) with however a non-zero (and explicit) function of $i, j, s, n$ on the r.h.s. in case the two inequalities on $i$ and $j$ are not satisfied (an example of such a inhomogeneous recurrence relation is given below in (B.13)).

Assuming that $V(s-2)$ and $V(s-1)$ are upper triangular, the relation (B.11) for $1 \leqslant i \leqslant s-1$, readily shows that the strictly lower part of $V(s)$ vanishes identically, except perhaps on the its last row ($i = s$) and its $\frac{s}{2}$-th column (since the coefficient of $V_{ij}(s)$ vanishes when $2j = s$).

Let us first handle the last row by proving that $V_{s,j}(s) = 0$ for all $1 \leqslant j \leqslant s-1$ and $s \geqslant 2$ (we will then be done if $s$ is odd). To do this, we combine the following exact evaluation (from Mathematica),

$$V_{s,0}(s) = \frac{n!}{s!\,(n+s)!}, \tag{B.12}$$

with the following first-order recurrence on $j$,

$$\begin{aligned} &2(2s-2j-1)(s-j-1)\,V_{s,j+1}(s) \\ &-\ (s-2j-1)(s-2j)(n+s-2j)(n+s-2j-1)\,V_{s,j}(s) \\ &=\ \frac{(2j-1)\,n!\,\left(8j^2-2jn-8js+6j+ns-n+s^2-2s+1\right)}{(1-2j)!\,(s-1)!\,(n+s-1)!}, \quad j \in \mathbb{Z}. \end{aligned} \tag{B.13}$$

For $j = 0$, it simplifies to

$$2(2s-1)(s-1)V_{s,1}(s) - s(s-1)(n+s)(n+s-1)V_{s,0}(s) = -\frac{n!}{(s-2)!\,(n+s-2)!}. \tag{B.14}$$

Provided $s$ is at least 2 (there is nothing to prove if $s = 1$), it shows, using (B.12), that $V_{s,1}(s) = 0$. Feeding this value in (B.13) for $1 \leqslant j \leqslant s-2$, and noting that the r.h.s. vanishes, we find $V_{s,j}(s) = 0$ for $1 \leqslant j \leqslant s-1$.

To finish the proof, it remains to show that the lower triangular part of the $s$-th column of $V(2s)$ vanishes. By the previous argument on the last row, we know that $V_{2s,s}(2s) = 0$. That single value would be enough if the recurrence over $i$ of $V_{i,s}(2s)$ was first order. Unfortunately, that recurrence is second order (see below), so that we need the second value $a_s \equiv V_{2s-1,s}(2s)$, which itself can be computed by recurrence over $s$. Indeed, one finds that it satisfies the first-order following relation,

$$(n-2s)(n-2s-1)(2s-1)\, a_{s+1} = (2s+1)\, a_s, \qquad s \geqslant 1. \tag{B.15}$$

For $s = 2$, we check that $a_2 = V_{3,2}(4) = 0$ (that's easy !), from which it follows that $a_s = V_{2s-1,s}(2s) = 0$ for $s = 2, 3, \ldots$ up to $s = \frac{n}{2}$.

We are ready to examine the recurrence over $i$ of $V_{i,s}(2s)$,

$$\begin{aligned}
&i(1+i)(2s-i-1)(2+i+n-2s)(2+2i+n-2s)(3+2i+n-2s)V_{i+2,s}(2s)\\
&-\big(8i^3+8i^2n-32i^2s+20i^2+in^2-20ins+11in+40is^2-52is+18i\\
&\qquad -2n^2s+n^2+12ns^2-16ns+5n-16s^3+32s^2-24s+6\big)V_{i+1,s}(2s)\\
&-2(2s-2i-1)(s-i)\, V_{i,s}(2s) = 0, \qquad i \leqslant 2s-1.
\end{aligned} \tag{B.16}$$

Starting the recurrence from our known results $V_{2s,s}(2s) = V_{2s-1,s}(2s) = 0$ and setting successively $i = 2s-2$, $2s-3, \ldots$ down to $i = s+1$ yields the result, namely $V_{s+1,s}(2s) = V_{s+2,s}(2s) = \ldots = V_{2s,s}(2s) = 0$. Thus, the lower triangular part of the $s$-th column of $V(2s)$ is zero.

By similar arguments, one can also check the diagonal elements $V_{ii}(s) = -1$, for $1 \leqslant i \leqslant s$. One first uses second-order recurrence relations on $s$ to prove that $V_{11}(s) = V_{22}(s) = -1$ for every $s \geqslant 2$. Then one establishes another second-order recurrence on $i$ for $V_{ii}(s)$, which, despite being frighteningly complicated, shows unambiguously that $V_{ii}(s) = -1$ for all $1 \leqslant i \leqslant s$. Details available on request.

Other strictly upper elements of $V$, hence of $U$, can be computed by these recurrence methods, but their pattern appears to be more difficult to grasp. To illustrate this, we just give the results for the elements of $V$ that are on the first diagonal above the main one. These are found to be equal to

$$\begin{aligned}
V_{i,i+1}(s) = -\frac{i}{2}\big(&4i^3-4i^2n-4i^2s+6i^2+in^2+2ins\\
&-3in+is^2-5is+2i+n^2-n+s^2-s\big),
\end{aligned} \tag{B.17}$$

for $1 \leqslant i \leqslant s-1$. For all $s, n$ I have considered, the matrices $V$ were found to be integer-valued.

We also conjecture the entries of the $L$ matrix.

**Conjecture B.2.1.** *The matrix $L$ has entries*

$$L_{ij} = \binom{i-1}{j-1} \left(\frac{n-s-1+2i}{n-s-1+2j}\right) \prod_{l=1}^{j-1} \frac{n-s-1+l+i}{n-s-1+l+j} \quad \textit{for } i \geq j, \tag{B.18}$$

*and $L_{i,j} = 0$ otherwise.*

## B.3 Computation of alternating sums in $\tilde{U}_{ss}$

We now show how to get rid of the alternating sign in (4.27). We rewrite

$$\tilde{U}_{ss} = \sum_{p=1}^{s} (-1)^{s+p} \binom{s-1}{p-1} \prod_{l=p}^{s-1} \frac{n+l}{n-s+l} \left[(1+\xi)^{s-p} - (1+\xi)^{n+p-1}\right] \tag{B.19}$$

as $\tilde{U}_{ss} = (-1)^s \binom{n-1}{s-1}^{-1} (U_2 - U_1)$, with

$$U_1 \equiv \sum_{p=1}^{s} (-1)^{s-p} \binom{n-p}{s-p} \binom{n+s-1}{n+s-p} (1+\xi)^{p-1}, \tag{B.20}$$

and

$$U_2 \equiv \sum_{p=1}^{s} (-1)^{s-p} \binom{n-p}{s-p} \binom{n+s-1}{n+s-p} (1+\xi)^{n+s-p}, \tag{B.21}$$

and treat both terms separately. We start by $U_1$

$$\begin{aligned}
U_1 &= \sum_{p=1}^{s} \frac{1}{(1+z)^{n-s+1}}\bigg|_{z^{s-p}} (1+z)^{n+s-1}\Big|_{z^{p-1}} (1+\xi)^{p-1} \\
&= \sum_{p=1}^{s} \frac{1}{(1+z)^{n-s+1}}\bigg|_{z^{s-p}} (1+(1+\xi)z)^{n+s-1}\Big|_{z^{p-1}} \\
&= \sum_{p=0}^{s-1} \frac{1}{(1+z)^{n-s+1}}\bigg|_{z^{s-p-1}} (1+(1+\xi)z)^{n+s-1}\Big|_{z^{p}} \\
&= \frac{(1+z+z\xi)^{n+s-1}}{(1+z)^{n-s+1}}\bigg|_{z^{s-1}} \\
&= \sum_{l=0}^{n+s-1} \binom{n+s-1}{l} (z\xi)^l (1+z)^{2s-2-l}\Big|_{z^{s-1}} \\
&= \sum_{l=0}^{s-1} \binom{n+s-1}{l} (z\xi)^l (1+z)^{2s-2-l}\Big|_{z^{s-1}} \\
&= \sum_{l=0}^{s-1} \binom{n+s-1}{l} z^l \xi^l \sum_{m=0}^{2s-2-l} \binom{2s-2-l}{m} z^m\bigg|_{z^{s-1}} \\
&= \sum_{l=0}^{s-1} \binom{n+s-1}{l} \binom{2(s-1)-l}{s-1-l} \xi^l .
\end{aligned} \tag{B.22}$$

Similarly for $U_2$

$$\begin{aligned}
U_2 &= \sum_{p=1}^{s} (-1)^{s-p} \binom{n-p}{s-p} \binom{n+s-1}{n+s-p} (1+\xi)^{n+s-p} \\
&= (1+\xi)^{n+s-1} \sum_{p=1}^{s} \frac{1}{(1+z)^{n-s+1}}\bigg|_{z^{s-p}} (1+z)^{n+s-1}\Big|_{z^{p-1}} (1+\xi)^{-(p-1)} \\
&= (1+\xi)^{n+s-1} \sum_{p=1}^{s} \frac{1}{(1+z)^{n-s+1}}\bigg|_{z^{s-p}} \left(1+\frac{z}{1+\xi}\right)^{n+s-1}\Bigg|_{z^{p-1}} \\
&= \frac{(1+z+\xi)^{n+s-1}}{(1+z)^{n-s+1}}\bigg|_{z^{s-1}} \\
&= \sum_{k=0}^{n+s-1} \binom{n+s-1}{k} (1+z)^{2(s-1)-k} \xi^k \bigg|_{z^{s-1}} \\
&= \sum_{k=0}^{2(s-1)} \binom{n+s-1}{k} (1+z)^{2(s-1)-k} \xi^k + \sum_{k=2s-1}^{n+s-1} \binom{n+s-1}{k} \frac{1}{(1+z)^{k-2(s-1)}} \xi^k \bigg|_{z^{s-1}} \\
&= \sum_{k=0}^{s-1} \binom{n+s-1}{k} \binom{2(s-1)-k}{s-1} \xi^k + (-1)^{s-1} \sum_{k=2s-1}^{n+s-1} \binom{n+s-1}{k} \binom{k-s}{s-1} \xi^k.
\end{aligned} \tag{B.23}$$

Hence, at the end of the day we get

$$\tilde{U}_{ss} = -\binom{n-1}{s-1}^{-1} \sum_{k=2s-1}^{n+s-1} \binom{n+s-1}{k} \binom{k-s}{s-1} \xi^k. \tag{B.24}$$

# Appendix C

# Entropy function of the Aztec diamond

This short appendix is devoted to the proof of the property, used in Section 5.4, that for the function $S(r,s)$ in (5.54), the difference $S(r,s) - S(r,s')$ is increasing in $r$ for any fixed $s > s'$, and strictly increasing if $rs > \frac{1}{1+w}$. As the function $S(r,s)$ has a different functional form according to the value of $rs$, we distinguish three separate cases.

For $rs' < rs \leqslant \frac{1}{1+w}$, $S(r,s)$ is equal to $\sigma(r) + \sigma(s)$, with

$$\sigma(x) = (x+1)\log(x+1) - x\log x + \frac{1}{2}\log(1+w), \qquad \sigma'(x) = \log\frac{x+1}{x}. \tag{C.1}$$

It follows that $S(r,s) - S(r,s') = \sigma(s) - \sigma(s')$ is constant in $r$.

For $rs' \leqslant \frac{1}{1+w} < rs$, we have

$$S(r,s) - S(r,s') = (r+s+2)L(t) - \sigma(r) - \sigma(s'), \qquad t = \frac{s-r}{r+s+2}. \tag{C.2}$$

By using that $L(t) = -\log x(t) - t\log y(t)$ and the fact that $L'(t) = -\log y(t)$, we obtain that the derivative with respect to $r$ simplifies to

$$\frac{\mathrm{d}}{\mathrm{d}r}\big[S(r,s) - S(r,s')\big] = \log\frac{r\,y(t)}{(r+1)\,x(t)}. \tag{C.3}$$

The explicit forms of the functions $x(t)$ and $y(t)$ show that the argument of the logarithm is strictly larger than 1 for $r > \frac{1}{s(1+w)}$ (it is exactly 1 for $r = \frac{1}{s(1+w)}$ and tends to 1, from above, when $r \to +\infty$). Therefore, the difference $S(r,s) - S(r,s')$ is strictly increasing.

In the last case, namely for $\frac{1}{1+w} < rs' < rs$, we have

$$S(r,s) - S(r,s') = (r+s+2)L(t) - (r+s'+2)L(t'), \qquad t' = \frac{s'-r}{r+s'+2}. \tag{C.4}$$

The calculation of the previous paragraph readily yields

$$\frac{\mathrm{d}}{\mathrm{d}r}\big[S(r,s) - S(r,s')\big] = \log \frac{y(t)\,x(t')}{x(t)\,y(t')} > 0. \tag{C.5}$$

The inequality follows from the ratio $y(t)/x(t)$ being itself a strictly increasing function on its open domain $]-1,1[$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{y(t)}{x(t)} = \frac{y'(t)x(t) - y(t)x'(t)}{x^2(t)} = (1+t)\frac{y'(t)}{x(t)} > 0. \tag{C.6}$$

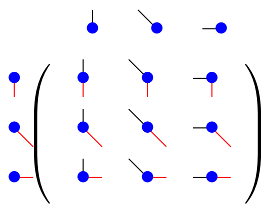
The fact that $s' < s$ implies the strict inequality $t' < t$ between the two slopes concludes the argument and shows that $S(r,s) - S(r,s')$ is strictly increasing in $r$.

# Appendix D

# Enumeration of vertex-weighted paths

We wish to compute $\mathcal{Y}_{(n,L)\to(n+M,0)} = \mathcal{Y}_{(n,L)\to(n+M,0)}(\beta_1, \beta_2, \beta_3)$ namely the partition function of a weighted path with elementary steps $(1, 0)$, $(1, -1)$ and $(0, -1)$ going from $(n, L)$ to $(n + M, 0)$. The weights are attributed to every visited vertex, and are inherited from the corresponding 20V model. The parameters $\beta_1$, $\beta_2$ and $\beta_3$ denote suitably chosen weights for the first node of the path. Note that the empty space around the path also receives a weight $\omega_0$ per empty vertex. Factoring those weights, the remaining weight is $\omega_i/\omega_0$ per vertex visited by the path, for which the local configuration corresponds to a vertex with weight $\omega_i$ in Figure 4.7. To compute $\mathcal{Y}_{(n,L)\to(n+M,0)}$, we use a *transfer matrix* technique. We introduce the $3 \times 3$ matrix $T$ with the following entries:

$T =$



$$= \frac{1}{\omega_0}\begin{pmatrix} \omega_1 u & \omega_2 u & \omega_4 u \\ \omega_2 uv & \omega_3 uv & \omega_5 uv \\ \omega_4 v & \omega_5 v & \omega_6 v \end{pmatrix}$$

Multiplication by $T$ on the left amounts to adding one extra step to paths, with the suitable 20V model normalized weights $\omega_i/\omega_0$ and some extra weights $u, uv, v$ per horizontal, diagonal and vertical step respectively so as to keep track of the global vertical and horizontal shifts. The generating function of the partition functions $\mathcal{Y}_{(n,L)\to(n+M,0)}$ then reads

$$\mathbb{Y}(u,v) = \sum_{L,M\geq 0} u^L v^M \mathcal{Y}_{(n,L)\to(n+M,0)} = 1 + (1,0,0)\ (\mathbb{I}-T)^{-1} \begin{pmatrix} \beta_1 u \\ \beta_2 uv \\ \beta_3 v \end{pmatrix}.$$

The final state $(1,0,0)$ corresponds to a vertical final step, as in the geometry of Figure 4.14-right. As for the initial state $(\beta_1 u, \beta_2 uv, \beta_3 v)^t$, it includes the special $\beta$ weights for the starting point depending on its local environment in the geometry at hand. In practice, the details of what happens at the extremities are irrelevant as they do not affect the large $n$ asymptotics, as shown below. The quantity $\mathcal{Y}(u,v)$ is a rational fraction $\mathbb{Y}(u,v) = g(u,v)/\Delta(u,v)$ where $g(u,v)$ is a polynomial which implicitly depends on the $\beta$ weights and

$$\Delta(u,v) = \det(\mathbb{I}-T) = 1 - \alpha_1 u - \alpha_2 v - \alpha_3 uv - \alpha_4 u^2 v - \alpha_5 uv^2 - \alpha_6 u^2 v^2$$

with

$$\begin{aligned}
\alpha_1 &= \frac{\omega_1}{\omega_0}, \quad \alpha_2 = \frac{\omega_6}{\omega_0}, \quad \alpha_3 = \frac{\omega_0\omega_3 + \omega_4^2 - \omega_1\omega_6}{\omega_0^2}, \\
\alpha_4 &= \frac{\omega_2^2 - \omega_1\omega_3}{\omega_0^2}, \quad \alpha_5 = \frac{\omega_5^2 - \omega_6\omega_3}{\omega_0^2}, \\
\alpha_6 &= \frac{2\omega_2\omega_4\omega_5 + \omega_1\omega_6\omega_3 - \omega_3\omega_4^2 - \omega_1\omega_5^2 - \omega_6\omega_2^2}{\omega_0^3}.
\end{aligned} \tag{D.1}$$

The large $n, L = \ell\, n, M = m\, n$ asymptotics of $\mathcal{Y}_{(n,L)\to(n+M,0)}$ are governed by $\Delta(u,v)$ only. Indeed, for large $L, M \propto n$:

$$\begin{aligned}
\mathcal{Y}_{(n,L)\to(n+M,0)} &= \oint \frac{\mathrm{d}u}{2\mathrm{i}\pi u^{L+1}} \frac{\mathrm{d}v}{2\mathrm{i}\pi v^{M+1}} \mathbb{Y}(u,v) \\
&\approx \oint \frac{\mathrm{d}u}{2\mathrm{i}\pi u^{L+1}} \frac{\mathrm{d}v}{2\mathrm{i}\pi v^{M+1}} g(u,v) \\
&\qquad \sum_{p\geq 0} (\alpha_1 u + \alpha_2 v + \alpha_3 uv + \alpha_4 u^2 v + \alpha_5 uv^2 + \alpha_6 u^2 v^2)^p \\
&= \oint \frac{\mathrm{d}u}{2\mathrm{i}\pi u^{L+1}} \frac{\mathrm{d}v}{2\mathrm{i}\pi v^{M+1}} g(u,v) \\
&\qquad \sum_{L',M'} u^{L'} v^{M'} \sum_{\substack{P_1,P_2,\ldots,P_6\geq 0 \\ P_1+P_3+2P_4+P_5+2P_6=L' \\ P_2+P_3+P_4+2P_5+2P_6=M'}} \binom{P_1+P_2+P_3+P_4+P_5+P_6}{P_1,P_2,P_3,P_4,P_5,P_6} \prod_{i=1}^{6} \alpha_i^{P_i}\ .
\end{aligned}$$

For large $n$, the integral selects values of $L'$ and $M'$ that differ from $L$ and $M$ by finite amounts bounded by the degree of the polynomial $g(u,v)$. Taking $L' = \ell\, n + O(1)$ and $M' = m\, n + O(1)$, we obtain the leading behaviour for large $P_i = np_i$:

$$\mathcal{Y}_{(n,n\ell)\to(n(1+m),0)} \approx \int_0^1 dp_3 dp_4 dp_5 dp_6 \; \mathrm{e}^{n\, S(\ell,m,p_3,p_4,p_5,p_6)},$$

where

$$\begin{aligned}
&S(\ell,m,p_3,p_4,p_5,p_6)\\
&= (\ell{+}m{-}p_3{-}2p_4{-}2p_5{-}3p_6)\log(\ell{+}m{-}p_3{-}2p_4{-}2p_5{-}3p_6)\\
&- (\ell{-}p_3{-}2p_4{-}p_5{-}2p_6)\log\left(\frac{\ell{-}p_3{-}2p_4{-}p_5{-}2p_6}{\alpha_1}\right)\\
&- (m{-}p_3{-}p_4{-}2p_5{-}2p_6)\log\left(\frac{m{-}p_3{-}p_4{-}2p_5{-}2p_6}{\alpha_2}\right) - \sum_{i=3}^{6} p_i \log\left(\frac{p_i}{\alpha_i}\right).
\end{aligned} \tag{D.2}$$

A saddle point estimate then allows to write

$$\mathcal{Y}_{(n,n\ell)\to(n(1+m),0)} \approx \mathrm{e}^{n\, S(\ell,m)},$$

where $S(\ell,m)$ is equal to $S(\ell,m,p_3,p_4,p_5,p_6)$ taken at the value of $p_3$, $p_4$, $p_5$ and $p_6$ which maximizes this latter quantity. The above expression can be adapted to a situation where the $\alpha_i$ for $i \in I \subset \{3,\cdots,6\}$ vanish. In this case, the recipe consists in simply dropping all the terms with indices $i \in I$.

# Appendix E

# Almost homogeneous unravelling

The aim of this appendix is to prove the general relations (4.52) and (4.53) relating the restricted refined partition functions

$$\mathcal{Z}_n^{-}(\tau) = \sum_{L=1}^{n} \mathcal{Z}_{n,L}^{-}\,\tau^{L-1}\ , \qquad \mathcal{Z}_n^{\diagdown}(\tau) = \sum_{L=1}^{n} \mathcal{Z}_{n,L}^{\diagdown}\,\tau^{L-1} \tag{E.1}$$

to the refined partition function

$$Z_n(\sigma) = \sum_{L=1}^{n} Z_{n,L}\,\sigma^{L-1}\ . \tag{E.2}$$

The relation (4.52) may be obtained by attaching a different spectral parameter $w\theta$ (instead of $w$) to the last column $j = n$. By unravelling the configurations of the 20V model with DWBC2 to a configuration of the 6V model on the sub-lattice 1, we have the relation, depicted in Figure E.1:

$$\mathcal{Z}_n[\theta] = \left(\frac{a_2 a_3}{t^{1/3}}\right)^{n^2} \left(\frac{a_3[\theta]}{a_3[1]}\right)^{n} Z_n[\theta]\ . \tag{E.3}$$

Here we shall use a different notation with *brackets* (as in $Z_n[\theta]$) to indicate that we deal with a model where the last column has a modified spectral parameter $w\theta$. It should not be confused with the notation with parentheses (as in $Z_n(\sigma)$) where all the column spectral parameters are

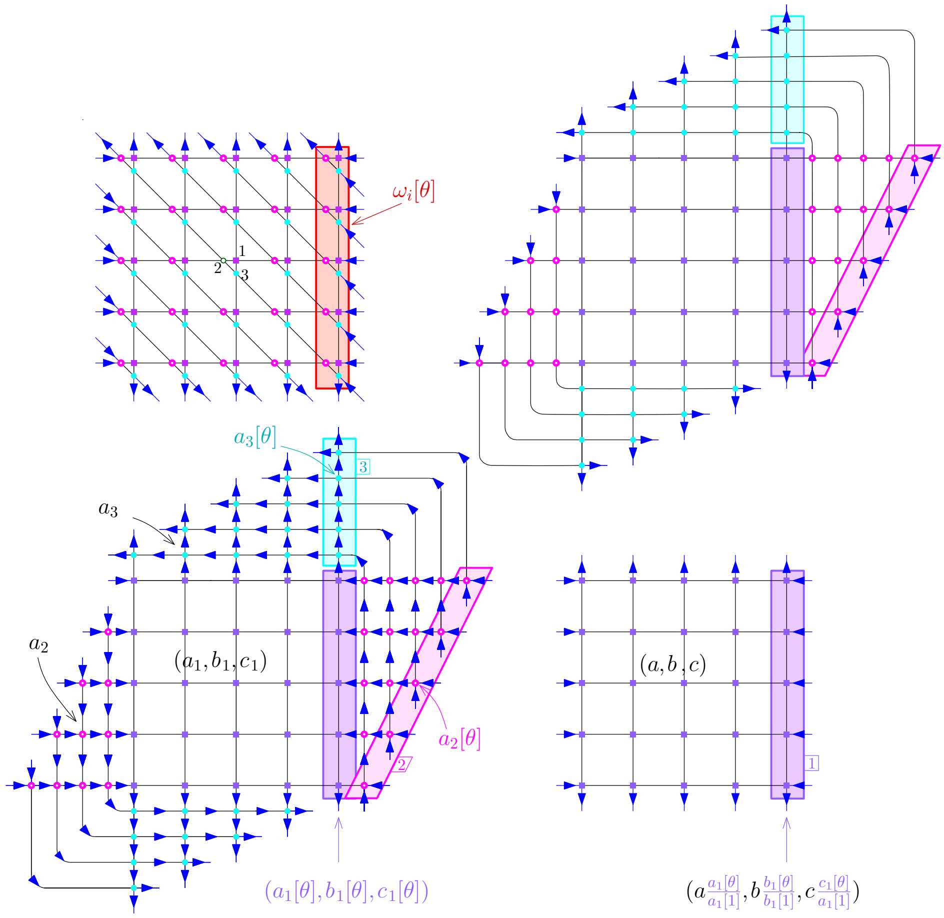



Figure E.1: The unravelling of Figure 4.12 in the presence of a spectral parameter $w\theta$ in the last column that modifies the weights for the $n$ vertices within the red box. In practice, only the weights of sub-lattices 1 and 3 are modified (since those of sub-lattice 2 do not involve the modified spectral parameter). Vertices have modified weight $a_3[\theta] = a_3 \times \frac{a_3[\theta]}{a_3[1]}$ in the cyan box (labelled 3), unmodified weight $a_2[\theta] = a_2$ in the pink box (labelled 2) and, after the same renormalization as in Figure 4.12 to go from $(a_1, b_1, c_1)$ to $(a, b, c)$, modified weights $(a \times \frac{a_1[\theta]}{a_1[1]}, b \times \frac{b_1[\theta]}{b_1[1]}, c \times \frac{c_1[\theta]}{c_1[1]})$ in the purple box labelled 1. This leads to the relation (E.3).

left equal to $w$ but where we deal with the refined partition function defined above keeping track of the position where the uppermost path hits the right boundary. The notation $a_i[\theta]$ (respectively $b_i[\theta]$ and $c_i[\theta]$), $i = 1, 2, 3$, also refers to the weights[1] obtained via the general formula (4.41) with $w \to w\theta$. Let us now discuss precisely the connection between the two objects $Z_n[\theta]$ and $Z_n(\sigma)$: by decomposing $Z_n[\theta]$ according to the position $L$ where the uppermost path hits the right boundary, we have

$$\begin{aligned} Z_n[\theta] &= \sum_{L=1}^{n} Z_{n,L} \left(\frac{b_1[\theta]}{b_1[1]}\right)^{L-1} \frac{c_1[\theta]}{c_1[1]} \left(\frac{a_1[\theta])]}{a_1[1]}\right)^{n-L} \\ &= \left(\frac{a_1[\theta]}{a_1[1]}\right)^{n} \frac{c_1[\theta]a_1[1]}{c_1[1]a_1[\theta]} \underbrace{\sum_{L=1}^{n} Z_{n,L}(\sigma[\theta])^{L-1}}_{=Z_n(\sigma[\theta])} \end{aligned} \tag{E.4}$$

with

$$\sigma[\theta] = \frac{b_1[\theta]a_1[1]}{b_1[1]a_1[\theta]} \ . \tag{E.5}$$

Similarly, decomposing $\mathcal{Z}_n[\theta]$ according to the position $L$ where the uppermost path hits the right boundary and to whether the step before the hitting point was horizontal or diagonal, we have

$$\begin{aligned} \mathcal{Z}_n[\theta] &= \sum_{L=1}^{n} \mathcal{Z}^{-}_{n,L} \left(\frac{\omega_1[\theta]}{\omega_1[1]}\right)^{L-1} \frac{\omega_4[\theta]}{\omega_4[1]} \left(\frac{\omega_0[\theta]}{\omega_0[1]}\right)^{n-L} \\ &\quad + \sum_{L=1}^{n} \mathcal{Z}^{\diagdown}_{n,L} \left(\frac{\omega_1[\theta]}{\omega_1[1]}\right)^{L-1} \frac{\omega_2[\theta]}{\omega_2[1]} \left(\frac{\omega_0[\theta]}{\omega_0[1]}\right)^{n-L} \\ &= \left(\frac{\omega_0[\theta]}{\omega_0[1]}\right)^{n} \left(\frac{\omega_4[\theta]\omega_0[1]}{\omega_4[1]\omega_0[\theta]} \mathcal{Z}^{-}_n(\tau[\theta]) + \frac{\omega_2[\theta]\omega_0[1]}{\omega_2[1]\omega_0[\theta]} \mathcal{Z}^{\diagdown}_n(\tau[\theta])\right), \end{aligned} \tag{E.6}$$

with

$$\tau[\theta] = \frac{\omega_1[\theta]\omega_0[1]}{\omega_1[1]\omega_0[\theta]} = \frac{b_1[\theta]a_2[\theta]b_3[\theta]a_1[1]a_2[1]a_3[1]}{b_1[1]a_2[1]b_3[1]a_1[\theta]a_2[\theta]a_3[\theta]} = \frac{b_1[\theta]b_3[\theta]a_1[1]a_3[1]}{b_1[1]b_3[1]a_1[\theta]a_3[\theta]} \ . \tag{E.7}$$

Using

$$\frac{\omega_0[\theta]}{\omega_0[1]} = \frac{a_1[\theta]a_2[\theta]a_3[\theta]}{a_1[1]a_2[1]a_3[1]} = \frac{a_1[\theta]a_3[\theta]}{a_1[1]a_3[1]} \tag{E.8}$$

[1] As before, when dealing with the 6V model, we use vertex weights normalized to $(a, b, c)$ as in (4.46). The change $w \to w\theta$ then affects the vertex weights in the last column, equal to $(a \times \frac{a_1[\theta]}{a_1[1]}, b \times \frac{b_1[\theta]}{b_1[1]}, c \times \frac{c_1[\theta]}{c_1[1]})$.

since $a_2$ is not changed by the replacement $w \to w\theta$, we obtain from (E.3)

$$\left(\frac{a_2 a_3}{t^{1/3}}\right)^{n^2} \; Z_n(\sigma[\theta]) = \mathcal{Z}_n^{-}(\tau[\theta]) + g(\sigma[\theta])\mathcal{Z}_n^{\diagdown}(\tau[\theta]) \tag{E.9}$$

with

$$\begin{aligned} g(\sigma[\theta]) &= \frac{c_1[1]a_1[\theta]}{c_1[\theta]a_1[1]} \frac{\omega_2[\theta]\omega_0[1]}{\omega_2[1]\omega_0[\theta]} = \frac{c_1[1]a_1[\theta]}{c_1[\theta]a_1[1]} \frac{(b_1[\theta]a_2[\theta]c_3[\theta])(a_1[1]a_2[1]a_3[1])}{(b_1[1]a_2[1]c_3[1])(a_1[\theta]a_2[\theta]a_3[\theta])} \\ &= \frac{c_1[1]b_1[\theta]c_3[\theta]a_3[1]}{c_1[\theta]b_1[1]c_3[1]a_3[\theta]} \ . \end{aligned} \tag{E.10}$$

Note the absence of prefactor in front of $\mathcal{Z}_n^{-}(\tau[\theta])$ in (E.9), due to the identity

$$\frac{c_1[1]a_1[\theta]}{c_1[\theta]a_1[1]} \frac{\omega_4[\theta]\omega_0[1]}{\omega_4[1]\omega_0[\theta]} = \frac{c_1[1]a_1[\theta]}{c_1[\theta]a_1[1]} \frac{c_1[\theta]a_2[\theta]a_3[\theta]a_1[1]a_2[1]a_3[1]}{c_1[1]a_2[1]a_3[1]a_1[\theta]a_2[\theta]a_3[\theta]} = 1 \ . \tag{E.11}$$

In particular, if we wish to impose a strict proportionality relation between $\mathcal{Z}_n(\tau[\theta])$ and $Z_n(\sigma[\theta])$, i.e. impose $g(\sigma[\theta]) = 1$, we must impose $\frac{\omega_2[\theta]}{\omega_2[1]} = \frac{\omega_4[\theta]}{\omega_4[1]}$ for all $\theta$, i.e. $\mu = \lambda - 5\eta$.

The relation between $\tau$ and $\sigma$ and that between $g$ and $\sigma$ are obtained by eliminating $\theta$. We obtain

$$\begin{aligned} \tau &= \sigma \frac{\sigma \sin(\lambda-\eta) \sin\left(\frac{\lambda+3\eta-\mu}{2}\right) - \sin(\lambda+\eta) \sin\left(\frac{\lambda-\eta-\mu}{2}\right)}{\sigma \sin(\lambda-\eta) \sin\left(\frac{\lambda-\eta-\mu}{2}\right) - \sin(\lambda+\eta) \sin\left(\frac{\lambda-5\eta-\mu}{2}\right)} \; \frac{\sin\left(\frac{\lambda+3\eta+\mu}{2}\right)}{\sin\left(\frac{\lambda-\eta+\mu}{2}\right)}, \\ g(\sigma) &= \frac{\sigma \sin(2\eta) \sin\left(\frac{\lambda+3\eta+\mu}{2}\right)}{\sigma \sin(\lambda-\eta) \sin\left(\frac{\lambda-\eta-\mu}{2}\right) - \sin(\lambda+\eta) \sin\left(\frac{\lambda-5\eta-\mu}{2}\right)}, \end{aligned} \tag{E.12}$$

which confirms (4.52) and (4.53), when combined with (E.9).

# Bibliography

[1] A. Okounkov, *Limit shapes, real and imagined*, Bull. Am. Math. Soc. (New Series) **53(2)** (2016).

[2] R. L. Dobrushin and S. B. Shlosman, *Droplet condensation in the Ising model: moderate deviations point of view*, in *Probability and phase transition*, pages 17–34, Springer, 1994.

[3] S. Miracle-Sole, *Facet shapes in a Wulff crystal*, in *Mathematical Results in Statistical Mechanics, Satellite Colloquium of STATPHYS 20*, pages 83–101, 1998. arXiv:1206.3736 [cond-mat.stat-mech].

[4] R. Cerf, *The Wulff Crystal in Ising and Percolation Models: École D'Été de Probabilités de Saint-Flour XXXIV-2004*, Springer, 2006.

[5] J. Olejarz, P. L. Krapivsky, S. Redner and K. Mallick, *Growth inside a corner: the limiting interface shape*, Phys. Rev. Lett. **108 1** (2012) 016102. arXiv:1109.1253 [cond-mat.stat-mech].

[6] R. Kenyon and A. Okounkov, *Limit shapes and the complex Burgers equation*, Acta Math. **199 2** (2007) 263–302. arXiv:math-ph/0507007.

[7] H. Cohn, N. Elkies and J. Propp, *Local statistics for random domino tilings of the Aztec diamond*, arXiv e-prints (2000), arXiv:math/0008243 [math.CO].

[8] H. Cohn, M. Larsen and J. Propp, *The shape of a typical boxed plane partition*, New York J. Math. **4** (1998) 137–165.

[9] M. Duits and A. Kuijlaars, *The two-periodic Aztec diamond and matrix valued orthogonal polynomials*, J. Eur. Math. Soc. **23 4** (2020) 1075–1131. arXiv:1712.05636 [math.PR].

[10] B. Pittel and D. Romik, *Limit shapes for random square Young tableaux*, Adv. Appl. Math. **38** (2007) 164–209.

[11] S. Caracciolo, G. Paoletti and A. Sportiello, *Explicit characterization of the identity configuration in an Abelian sandpile model*, J. Phys. A **41** (2008) 495003. arXiv:0809.3416 [cond-mat.stat-mech].

[12] I. Alevy and S. Mkrtchyan, *The limit shape of the leaky Abelian sandpile model*, arXiv preprint arXiv:2010.01946 (2020).

[13] F. Colomo and A. G. Pronko, *The limit shape of large alternating sign matrices*, SIAM J. Discrete Math. **24 4** (2010) 1558–1571. arXiv:0803.2697 [math-ph].

[14] F. Colomo and A. G. Pronko, *The arctic curve of the domain-wall six-vertex model*, J. Stat. Phys. **138 4-5** (2010) 662–700. arXiv:0907.1264 [math-ph].

[15] F. Colomo, A. G. Pronko and P. Zinn-Justin, *The arctic curve of the domain wall six-vertex model in its antiferroelectric regime*, J. Stat. Mech.: Theory Exp. **2010 03** (2010) L03002. arXiv:1001.2189 [math-ph].

[16] J. de Gier, R. Kenyon and S. S. Watson, *Limit shapes for the asymmetric five vertex model*, Commun. Math. Phys. (2021) 1–44. arXiv:1812.11934 [math.PR].

[17] B. Debin, P. Di Francesco and E. Guitter, *Arctic curves of the twenty-vertex model with domain wall boundaries*, J. Stat. Phys. **179** (2020) 33–89. arXiv:1910.06833 [math-ph].

[18] R. Kenyon, A. Okounkov and S. Sheffield, *Dimers and amoebae*, Ann. Math. **163** (2006) 1019–1056. arXiv:math-ph/0311005.

[19] N. Reshetikhin and A. Sridhar, *Limit shapes of the stochastic six vertex model*, Commun. Math. Phys. **363 3** (2018) 741–765. arXiv:1609.01756 [math-ph].

[20] F. Colomo and A. Sportiello, *Arctic curves of the six-vertex model on generic domains: the tangent method*, J. Stat. Phys. **164 6** (2016) 1488–1523. arXiv:1605.01388 [math-ph].

[21] A. Aggarwal, *Arctic boundaries of the ice model on three-bundle domains*, Invent. Math. (2019) 1–61. arXiv:1812.03847 [math.PR].

[22] P. Di Francesco and M. F. Lapa, *Arctic curves in path models from the tangent method*, J. Phys. A **51 15** (2018) 155202. arXiv:1711.03182 [math-ph].

[23] P. Di Francesco and E. Guitter, *Arctic curves for paths with arbitrary starting points: a tangent method approach*, J. Phys. A **51 35** (2018) 355201. arXiv:1803.11463 [math-ph].

[24] B. Debin and P. Ruelle, *Tangent method for the arctic curve arising from freezing boundaries*, J. Stat. Mech.: Theory and Exp. **2019 12** (2019) 123105. arXiv:1810.04909 [math-ph].

[25] P. Di Francesco and E. Guitter, *A tangent method derivation of the arctic curve for q-weighted paths with arbitrary starting points*, J. Phys. A **52 11** (2019) 115205. arXiv:1810.07936 [math-ph].

[26] F. Colomo, A. G. Pronko and A. Sportiello, *Arctic curve of the free-fermion six-vertex model in an L-shaped domain*, J. Stat. Phys. **174 1** (2019) 1–27. arXiv:1807.07549 [math-ph].

[27] P. Di Francesco and E. Guitter, *The arctic curve for Aztec rectangles with defects via the tangent method*, J. Stat. Phys. **176 3** (2019) 639–678. arXiv:1902.06478 [math-ph].

[28] S. Corteel, D. Keating and M. Nicoletti, *Arctic curves phenomena for bounded lecture hall tableaux*, Communications in Mathematical Physics **382 3** (2021) 1449–1493. arXiv:1905.02881.

[29] I. R. Passos and G. A. P. Ribeiro, *Arctic curve of the free-fermion six-vertex model with reflecting end boundary condition*, J. Stat. Mech.: Theory Exp. **2019 11** (2019) 113103. arXiv:1908.05773 [math-ph].

[30] P. Di Francesco, *Arctic curves of the reflecting boundary six vertex and of the twenty vertex models*, arXiv preprint arXiv:2106.02098 [math-ph] (2021).

[31] W. Jockusch, J. Propp and P. Shor, *Random domino tilings and the arctic circle theorem*, arXiv Mathematics e-prints (1998), math/9801068, arXiv:math/9801068 [math.CO].

[32] R. H. Fowler and G. S. Rushbrooke, *An attempt to extend the statistical theory of perfect solutions*, J. Chem. Soc. Faraday Trans. **33** (1937) 1272–1294.

[33] M. O. Blunt *et al.*, *Random tiling and topological defects in a two-dimensional molecular network*, Science **322 5904** (2008) 1077–1081.

[34] M. Gordon and W. H. T. Davison, *Theory of resonance topology of fully aromatic hydrocarbons. i*, J. Chem. Phys. **20 3** (1952) 428–435.

[35] H. Cohn, R. Kenyon and J. Propp, *A variational principle for domino tilings*, J. Am. Math. Soc. **14 2** (2001) 297–346. arXiv:math/0008220 [math.CO].

[36] P. Hall, *On representatives of subsets*, in *Classic Papers in Combinatorics*, pages 58–62, Springer, 2009.

[37] W. P. Thurston, *Conway's tiling groups*, The American Mathematical Monthly **97 8** (1990) 757–773.

[38] V. Gorin, *Lectures on random lozenge tilings*, Cambridge Studies in Advanced Mathematics (2021).

[39] E. H. Lieb, *Solution of the dimer problem by the transfer matrix method*, J. Math. Phys. **8 12** (1967) 2339–2341.

[40] P. W. Kasteleyn, *The statistics of dimers on a lattice: I. the number of dimer arrangements on a quadratic lattice*, Physica **27 12** (1961) 1209–1225.

[41] H. N. V. Temperley and M. E. Fisher, *Dimer problem in statistical mechanics-an exact result*, Philos. Mag. **6 68** (1961) 1061–1063.

[42] N. Elkies, G. Kuperberg, M. Larsen and J. Propp, *Alternating-sign matrices and domino tilings (part i)*, J. Algebr. Comb. **1 2** (1992) 111–132. arXiv:math/9201305 [math.CO].

[43] W. Jockusch, *Perfect matchings and perfect squares*, J. Comb. Theory, Series A **67 1** (1994) 100–115.

[44] M. Ciucu, *Enumeration of perfect matchings in graphs with reflective symmetry*, J. Comb. Theory, Series A **77 1** (1997) 67–97.

[45] M. Ciucu, *Perfect matchings of cellular graphs*, J. Algebr. Comb. **5 2** (1996) 87–103.

[46] J. Propp, *Enumeration of matchings: problems and progress*, arXiv preprint arXiv:math/9904150 [math.CO] (1999).

[47] M. Ciucu, *Symmetries of shamrocks, part i*, Journal of Combinatorial Theory, Series A **155** (2018) 376–397.

[48] M. Ciucu, *A new solution for the two dimensional dimer problem*, arXiv preprint arXiv:2102.07229 (2021).

[49] M. E. Fisher, *Statistical mechanics of dimers on a plane lattice*, Phys. Rev. **124 6** (1961) 1664.

[50] N. Allegra, *Critical Properties of Dimers, Spin Chains and Interface Models*, PhD thesis, Université de Lorraine, 2015.

[51] A. Van-Brunt and M. Visser, *Special-case closed form of the Baker–Campbell–Hausdorff formula*, J. Phys. A: Math. Theor. **48 22** (2015) 225207. arXiv:1501.02506 [math-ph].

[52] H. A. Kramers and G. H. Wannier, *Statistics of the two-dimensional ferromagnet. part i*, Phys. Rev. **60 3** (1941) 252.

[53] L. Onsager, *Crystal statistics. i. a two-dimensional model with an order-disorder transition*, Phys. Rev. **65 3-4** (1944) 117.

[54] R. J. Baxter, *Exactly solved models in statistical mechanics*, Academic Press, 1982.

[55] M. Gaudin, *La fonction d'onde de Bethe*, Masson, 1983.

[56] N. Allegra, J. Dubail, J.-M. Stéphan and J. Viti, *Inhomogeneous field theory inside the arctic circle*, J. Stat. Mech.: Theory Exp. **2016 5** (2016) 053108. arXiv:1512.02872 [cond-mat.stat-mech].

[57] R. J. Baxter, *Dimers on a rectangular lattice*, J. Math. Phys. **9 4** (1968) 650–654.

[58] M. Matone, *An algorithm for the Baker-Campbell-Hausdorff formula*, J. High Energy Phys. **2015 5** (2015) 113. arXiv:1502.06589 [math-ph].

[59] P. Di Francesco and E. Guitter, *Twenty-vertex model with domain wall boundaries and domino tilings*, Electron. J. Comb. (2020) 2–13. arXiv:1905.12387 [math.CO].

[60] J. Rasmussen and P. Ruelle, *Refined conformal spectra in the dimer model*, J. Stat. Mech.: Theory Exp. **2012 10** (2012) 10002. arXiv:1207.0385 [hep-th].

[61] M. E. Fisher and J. Stephenson, *Statistical mechanics of dimers on a plane lattice. ii. dimer correlations and monomers*, Phys. Rev. **132 4** (1963) 1411.

[62] D. Grensing, I. Carlsen and H.-Chr. Zapp, *Some exact results for the dimer problem on plane lattices with non-standard boundaries*, Philos. Mag. A **41 5** (1980) 777–781.

[63] N. Elkies, G. Kuperberg, M. Larsen and J. Propp, *Alternating-sign matrices and domino tilings (part ii)*, J. Algebr. Comb. **1 3** (1992) 219–234. arXiv:math/9201305 [math.CO].

[64] E. H. Kuo, *Applications of graphical condensation for enumerating matchings and tilings*, Theor. Comput. Sci. **319 1-3** (2004) 29–57. arXiv:math/0304090 [math.CO].

[65] J. Propp, *Generalized domino-shuffling*, Theor. Comput. Sci. **303 2-3** (2003) 267–301. arXiv:math/0111034 [math.CO].

[66] S.-P. Eu and T.-S. Fu, *A simple proof of the Aztec diamond theorem*, Electron. J. Comb. **12 R18** (2005) 1. arXiv:math/0412041 [math.CO].

[67] K. Johansson, *Non-intersecting paths, random tilings and random matrices*, Probab. Theory Relat. Fields **123 2** (2002) 225–280. arXiv:math/0011250 [math.PR].

[68] S. Karlin and J. McGregor, *Coincidence properties of birth and death processes*, Pac. J. Math. **9 4** (1959) 1109–1140.

[69] S. Karlin and J. McGregor, *Coincidence probabilities*, Pac. J. Math. **9 4** (1959) 1141–1164.

[70] B. Lindström, *On the vector representations of induced matroids*, Bull. London Math. Soc. **5 1** (1973) 85–90.

[71] I. Gessel and G. Viennot, *Binomial determinants, paths, and hook length formulae*, Adv. Math. **58** (1985) 300–321.

[72] M. E. Fisher, *Walks, walls, wetting, and melting*, J. Stat. Phys. **34 5-6** (1984) 667–729.

[73] P. John and H. Sachs, *Wegesysteme und linearfaktoren in hexagonalen und quadratischen systemen (path systems and linear factors in hexagonal and square systems)*, Graphen in Forschung und Unterricht (Festschrift K. Wagner), Franzbecker, Bad Salzdetfurth (1985) 85–101.

[74] M. Aigner and G. M. Ziegler, *Proofs from the Book*, Volume *274*, Springer, 2010.

[75] C. Krattenthaler, *Advanced determinant calculus*, in *The Andrews Festschrift*, pages 349–426, Springer, 2001. arXiv:math/9902004 [math.CO].

[76] C. Krattenthaler, *Advanced determinant calculus: a complement*, Linear Algebra Appl. **411** (2005) 68–166. arXiv:math/0503507 [math.CO].

[77] F. R. Gantmakher, *The theory of matrices*, Volume *131*, American Mathematical Soc., 1959.

[78] E. Duse and A. Metcalfe, *Asymptotic geometry of discrete interlaced patterns: part i*, Int. J. Math. **26 11** (2015) 1550093. arXiv:1412.6653 [math.PR].

[79] R. Kenyon and A. Okounkov, *Planar dimers and Harnack curves*, Duke Math. J **131 3** (2006) 499–524. arXiv:math/0311062 [math.AG].

[80] P. Di Francesco and R. Soto-Garrido, *Arctic curves of the octahedron equation*, J. Phys. A **47 28** (2014) 285204, 34. arXiv:1402.4493 [math-ph].

[81] R. Kenyon and R. Pemantle, *Double-dimers, the Ising model and the hexahedron recurrence*, J. Comb. Theory, Series A **137** (2016) 27–63. arXiv:1308.2998 [math-ph].

[82] L. Petrov, *Asymptotics of random lozenge tilings via Gelfand–Tsetlin schemes*, Probab. Theory Relat. Fields **160 3-4** (2014) 429–487. arXiv:1202.3901 [math.PR].

[83] E. Duse and A. Metcalfe, *Asymptotic geometry of discrete interlaced patterns: part ii*, in *Ann. Inst. Fourier*, Volume *70*, pages 375–436, 2020. arXiv:1507.00467 [math-ph].

[84] L. S. Levitov, *The rvb model as the problem of the surface of a quantum crystal*, Soviet Journal of Experimental and Theoretical Physics Letters **50** (1989) 469.

[85] S. Chhita, K.Johansson and B. Young, *Asymptotic domino statistics in the Aztec diamond*, Ann. Appl. Probab. **25 3** (2015) 1232–1278. arXiv:1212.5414 [math.PR].

[86] É. Janvresse, T. de La Rue and Y. Velenik, *A note on domino shuffling*, Electron. J. Comb. **13 1** (2006) R30.

[87] A. Doeraene, *Aztec diamond generator* 2018, https://sites.uclouvain.be/aztecdiamond/.

[88] N. M. Bogoliubov, A. G. Pronko and M. B. Zvonarev, *Boundary correlation functions of the six-vertex model*, J. Phys. A **35 27** (2002) 5525. arXiv:math-ph/0203025.

[89] L. Cantini, F. Colomo and A. G. Pronko, *Integral formulas and antisymmetrization relations for the six-vertex model*, in *Annales Henri Poincaré*, Volume *21*, pages 865–884, 2020. arXiv:1906.07636 [math-ph].

[90] M. Fulmek and C. Krattenthaler, *The number of rhombus tilings of a symmetric hexagon which contain a fixed rhombus on the symmetry axis, i*, Ann. Comb. **2 1** (1998) 19–41. arXiv:math/9712244 [math.CO].

[91] P. A. MacMahon, *Combinatory Analysis, Volumes I and II*, Cambridge University Press, 1915.

[92] E. Nordenstam and B. Young, *Domino shuffling on Novak half-hexagons and Aztec half-diamonds*, Electron. J. Comb. **18 1** (2011) 181. arXiv:1103.5054 [math.CO].

[93] L. P. Eisenhart, *A treatise on the differential geometry of curves and surfaces*, Courier Corporation, 2004.

[94] F. Colomo, V. Noferini and A. G. Pronko, *Algebraic arctic curves in the domain-wall six-vertex model*, J. Phys. A: Math. Theor. **44 19** (2011) 195201. arXiv:1012.2555 [math-ph].

[95] K. Astala, E. Duse, I. Prause and X. Zhong, *Dimer models and conformal structures*, arXiv preprint arXiv:2004.02599 (2020).

[96] C. Krattenthaler, *Descending plane partitions and rhombus tilings of a hexagon with a triangular hole*, Eur. J. Comb. **27 7** (2006) 1138–1146. arXiv:math/0310188.

[97] F. Bornemann, *On the numerical evaluation of Fredholm determinants*, Math. Comp. **79 270** (2010) 871–915. arXiv:0804.2543 [math.NA].

[98] P. Zinn-Justin, *Six-vertex model with domain wall boundary conditions and one-matrix model*, Phys. Rev. E **62** (Sep 2000) 3411–3418. arXiv:math-ph/0005008.

[99] R. E. Behrend, P. Di Francesco and P. Zinn-Justin, *On the weighted enumeration of alternating sign matrices and descending plane partitions*, J. Combin. Theory Ser. A **119 2** (2012) 331–363. arXiv:1103.1176 [math.CO].

[100] W. Krauth, *Statistical mechanics: algorithms and computations*, Volume *13*, OUP Oxford, 2006.

[101] W. Doeblin, *Exposé de la théorie des chaines simples constantes de Markov à un nombre fini d'états*, Mathématique de l'Union Interbalkanique **2 77-105** (1938) 78–80.

[102] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *Equation of state calculations by fast computing machines*, J. Chem. Phys. **21 6** (1953) 1087–1092.

[103] D. B. Wilson, *Mixing times of lozenge tiling and card shuffling Markov chains*, Ann. Appl. Probab. **14 1** (2004) 274–325.

[104] I. Lyberg, V. Korepin and J. Viti, *The density profile of the six vertex model with domain wall boundary conditions*, J. Stat. Mech.: Theory Exp. **2017 5** (2017) 053103. arXiv:1612.06758 [cond-mat.stat-mech].

[105] I. Lyberg, V. Korepin, G. A. P. Ribeiro and J. Viti, *Phase separation in the six-vertex model with a variety of boundary conditions*, J. Math. Phys. **59 5** (2018) 053301. arXiv:1711.07905 [cond-mat.stat-mech].

[106] D. Allison and N. Reshetikhin, *Numerical study of the 6-vertex model with domain wall boundary conditions*, in *Ann. Inst. Fourier*, Volume *55*, pages 1847–1869, 2005. arXiv:cond-mat/0502314 [cond-mat.stat-mech].

[107] J. G. Propp and D. B. Wilson, *Exact sampling with coupled Markov chains and applications to statistical mechanics*, Random Structures & Algorithms **9 1-2** (1996) 223–252.

[108] W. Krauth and R. Moessner, *Pocket Monte Carlo algorithm for classical doped dimer models*, Phys. Rev. B **67 6** (2003) 064503. arXiv:cond-mat/0206177 [cond-mat.str-el].

[109] O. F. Syljuåsen. and M. B. Zvonarev, *Directed-loop Monte Carlo simulations of vertex models*, Phys. Rev. E **70** (2004) 016118. arXiv:cond-mat/0401491 [cond-mat.stat-mech].

[110] R. Keesman, J. Lamers, R. A. Duine and G. T. Barkema, *Finite-size scaling at infinite-order phase transitions*, J. Stat. Mech.: Theory Exp. **2016 9** (2016) 093201. arXiv:1605.08876 [cond-mat.stat-mech].

[111] R. Keesman and J. Lamers, *Numerical study of the F model with domain-wall boundaries*, Phys. Rev. E **95 5** (2017) 052117. arXiv:1702.05474 [cond-mat.stat-mech].

[112] D. Keating and A. Sridhar, *Random tilings with the GPU*, J. Math. Phys. **59 9** (2018) 091420. arXiv:1804.07250 [cs.OH].

[113] K. Johansson, *The arctic circle boundary and the Airy process*, Ann. Probab. **33 1** (2005) 1–30. arXiv:math/0306216 [math.PR].

[114] B. J. Fleming and P. J. Forrester, *Interlaced particle systems and tilings of the Aztec diamond*, J. Stat. Phys. **142 3** (2011) 441–459. arXiv:1004.0474.

[115] J.-M. Stéphan, *Extreme boundary conditions and random tilings*, arXiv preprint arXiv:2003.06339 [cond-mat.stat-mech] (2020).

[116] S. Delvaux and A. B. Kuijlaars, *A phase transition for nonintersecting Brownian motions, and the painlevé ii equation*, Int. Math. Res. Not. **2009 19** (2009) 3639–3725. arXiv:0809.1000.

[117] I. Corwin and A. Hammond, *Brownian Gibbs property for Airy line ensembles*, Invent. Math. **195 2** (2014) 441–508. arXiv:1108.2291 [math.PR].

[118] C. A. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Commun. Math. Phys. **159 1** (1994) 151–174. arXiv:hep-th/9211141.

[119] D. S. Dean, P. Le Doussal, S. N. Majumdar and G. Schehr, *Noninteracting fermions in a trap and random matrix theory*, J. Phys. A **52 14** (2019) 144006. arXiv:1810.12583 [cond-mat.stat-mech].

[120] F. Bornemann, *On the numerical evaluation of distributions in random matrix theory: a review*, arXiv e-prints (2009), arXiv:0904.1581 [math.PR].

[121] M. Prähofer and H. Spohn, *Exact scaling functions for one-dimensional stationary KPZ growth*, J. Stat. Phys. **115 1** (2004) 255–279. arXiv:cond-mat/0212519.

[122] M. Prähofer and H. Spohn, *Scale invariance of the PNG droplet and the Airy process*, J. Stat. Phys. **108 5-6** (2002) 1071–1106. arXiv:math/0105240 [math.PR].

[123] K. Johansson, *Discrete polynuclear growth and determinantal processes*, Commun. Math. Phys. **242 1-2** (2003) 277–329. arXiv:math/0206208 [math.PR].

[124] I. Corwin, J. Quastel and D. Remenik, *Continuum statistics of the Airy 2 process*, Commun. Math. Phys. **317 2** (2013) 347–362. arXiv:1106.2717 [math.PR].

[125] P. Deift, *Universality for mathematical and physical systems*, in *Proceedings of the International Congress of Mathematicians Madrid, August 22–30, 2006*, pages 125–152, 2007. arXiv:math-ph/0603038.

[126] K. A. Takeuchi and M. Sano, *Universal fluctuations of growing interfaces: evidence in turbulent liquid crystals*, Phys. rev. lett. **104 23** (2010) 230601. arXiv:1001.5121 [cond-mat.stat-mech].

[127] G. Makey *et al.*, *Universality of dissipative self-assembly from quantum dots to human cells*, Nature Physics **16 7** (2020) 795–801.

[128] J. Baik, P. Deift and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, J. Am. Math. Soc. **12 4** (1999) 1119–1178. arXiv:math/9810105 [math.CO].

[129] L. Pauling, *The structure and entropy of ice and of other crystals with some randomness of atomic arrangement*, J. Am. Chem. Soc. **57 12** (1935) 2680–2684.

[130] E. H. Lieb, *Exact solution of the problem of the entropy of two-dimensional ice*, Phys. Rev. Lett. **18 17** (1967) 692.

[131] E. H. Lieb, *Residual entropy of square ice*, Phys. Rev. **162 1** (1967) 162.

[132] B. Sutherland, *Exact solution of a two-dimensional model for hydrogen-bonded crystals*, Phys. Rev. Lett. **19 3** (1967) 103.

[133] W. Zhou *et al.*, *The observation of square ice in graphene questioned*, Nature **528 7583** (2015) E1–E2.

[134] G. Algara-Siller *et al.*, *Square ice in graphene nanocapillaries*, Nature **519 7544** (2015) 443–445.

[135] V. Korepin and P. Zinn-Justin, *Thermodynamic limit of the six-vertex model with domain wall boundary conditions*, J. Phys. A **33 40** (2000) 7053. arXiv:cond-mat/0004250 [cond-mat.stat-mech].

[136] V. E. Korepin, *Calculation of norms of Bethe wave functions*, Commun. Math. Phys. **86 3** (1982) 391–418.

[137] N. A. Slavnov, *Algebraic Bethe ansatz*, arXiv preprint arXiv:1804.07350[math-ph] (2018).

[138] A. G. Izergin, *Partition function of a six-vertex model in a finite volume*, Dokl. Akad. Nauk SSSR **297 2** (1987) 331–333.

[139] A. G. Izergin, D. A. Coker and V.E. Korepin, *Determinant formula for the six-vertex model*, J. Phys. A **25 16** (1992) 4315–4334.

[140] F. Colomo and A. G. Pronko, *The arctic circle revisited*, Contemporary Mathematics **458** (2008) 361–376. arXiv:0704.0362 [math-ph].

[141] L. F. Cugliandolo, G. Gonnella and A. Pelizzola, *Six–vertex model with domain wall boundary conditions in the Bethe–Peierls approximation*, J. Stat. Mech.: Theory Exp. **2015 6** (2015) 06008. arXiv:1501.00883 [cond-mat.stat-mech].

[142] I. R. Passos, *Phd thesis: Separação espacial de fases no modelo de seis vértices*, (2020).

[143] P. Bleher and K. Liechty, *Six-vertex model with partial domain wall boundary conditions: ferroelectric phase*, J. Math. Phys. **56 2** (2015) 023302. arXiv:1407.8483 [math-ph].

[144] A. Aggarwal, *Limit shapes and local statistics for the stochastic six-vertex model*, Comm. Math. Phys. **376 1** (2020) 681–746. arXiv:1902.10867 [math.PR].

[145] A. G. Pronko and G. P. Pronko, *Off-shell Bethe states and the six-vertex model*, J. Math. Sci. **242** (2019) 742–752.

[146] S. Brasseur and C. Hagendorf, *Sum rules for the supersymmetric eight-vertex model*, J. Stat. Mech.: Theory Exp. **2021 2** (2021) 023102. arXiv:2009.14077 [math-ph].

[147] J.-M. Maillet, G. Niccoli and L. Vignoli, *Separation of variables bases for integrable $gl_{M|N}$ and Hubbard models*, (2019). arXiv:1907.08124 [math-ph].

[148] C. Babenko, F. Göhmann, K. K. Kozlowski, J. Sirker and J. Suzuki, *Exact real-time longitudinal correlation functions of the massive XXZ chain*, Phys. Rev. Lett. **126 21** (2021) 210602.

[149] S. Belliard, R. A. Pimenta and N. A. Slavnov, *Scalar product for the XXZ spin chain with general integrable boundaries*, arXiv preprint arXiv:2103.12501 [math-ph] (2021).

[150] P. Ruggiero, P. Calabrese, B. Doyon and J. Dubail, *Quantum generalized hydrodynamics*, Phys. rev. lett. **124 14** (2020) 140603. arXiv:1910.00570 [cond-mat.quant-gas].

[151] S. Bocini and J.-M. Stéphan, *Non-probabilistic fermionic limit shapes*, J. Stat. Mech.: Theory Exp. **2021 1** (2021) 013204. arXiv:2007.06621 [cond-mat.stat-mech].

[152] A. Borodin, I. Corwin and V. Gorin, *Stochastic six-vertex model*, Duke Math. J. **165 3** (2016) 563–624. arXiv:1407.6729 [math.PR].

[153] S. B. Kelland, *Twenty-vertex model on a triangular lattice*, Australian Journal of Physics **27** (dec 1974) 813–829.

[154] C. Krattenthaler, *Descending plane partitions and rhombus tilings of a hexagon with a triangular hole*, Eur. J. Comb. **27 7** (2006) 1138–1146. arXiv:math/0310188 [math.CO].

[155] R. E. Behrend, P. Di Francesco and P. Zinn-Justin, *A doubly-refined enumeration of alternating sign matrices and descending plane partitions*, J. Combin. Theory Ser. A **120 2** (2013) 409–432. arXiv:1202.1520 [math.CO].

[156] I. Fischer and M. Konvalinka, *The mysterious story of square ice, piles of cubes, and bijections*, Proc. Natl. Acad. Sci. U.S.A. **117 38** (2020) 23460–23466.

[157] I. Fischer and M. Konvalinka, *A bijective proof of the ASM theorem part i: the operator formula*, Electron. J. Comb. **27 3** (2020) 3–35. arXiv:1910.04198.

[158] I. Fischer and M. Konvalinka, *A bijective proof of the ASM theorem part ii: ASM enumeration and asm–dpp relation*, Int. Math. Res. Not. (2020). arXiv:1912.01354.

[159] B. Debin, E. Granet and P. Ruelle, *Concavity analysis of the tangent method*, J. Stat. Mech.: Theory Exp. **2019 11** (2019) 113107. arXiv:1905.11277 [math-ph].

[160] B. Debin and P. Ruelle, *Factorization in the multirefined tangent method*, arXiv preprint arXiv:2105.02257 [math-ph] (2021).

[161] R. Pemantle and M. C. Wilson, *Twenty combinatorial examples of asymptotics derived from multivariate generating functions*, Siam Rev. **50 2** (2008) 199–272. arXiv:math/0512548 [math.CO].

[162] C. Krattenthaler, *Lattice path enumeration*, in *Handbook of Enumerative Combinatorics*, pages 613–702, Chapman and Hall/CRC, 2015. arXiv:1503.05930 [math.CO].

[163] A. Sportiello, *Simple approaches to arctic curves for alternating sign matrices*, talk given at the conference "Statistical Mechanics, Integrability and Combinatorics", Florence 2015.

[164] A. Sportiello, *The tangent method: where do we stand ?*, talk given at the conference "Integrability, Combinatorics and Representations", Hyères 2019.

[165] M. Luby, D. Randall and A. Sinclair, *Markov chain algorithms for planar lattice structures*, SIAM J. Comput. **31** (2001) 167–192.

[166] D. Zeilberger, *Proof of the alternating sign matrix conjecture*, Electron. J. Combin. **3 2** (1996) R13. arXiv:math/9407211 [math.CO].

[167] D. M. Bressoud, *Proofs and confirmations: the story of the alternating-sign matrix conjecture*, Cambridge University Press, 1999.

[168] G. Kuperberg, *Another proof of the alternative-sign matrix conjecture*, IMRN **1996 3** (1996) 139–150. arXiv:math/9712207 [math.CO].

[169] R. E. Behrend, *Multiply-refined enumeration of alternating sign matrices*, Adv. Math. **245** (2013) 439–499. arXiv:1203.3187 [math.CO].

[170] A. Ayyer and D. Romik, *New enumeration formulas for alternating sign matrices and square ice partition functions*, Adv. Math. **235** (2013) 161–186. arXiv:1202.3651 [math.CO].

[171] D. Zeilberger, *Proof of the refined alternating sign matrix conjecture*, New York J. Math. **2** (1996) 59–68.

[172] F. Colomo and A. G. Pronko, *On two-point boundary correlations in the six-vertex model with DWBC*, J. Stat. Mech. (2005) 05010. arXiv:math-ph/0503049.

[173] Yu. Stroganov, *Izergin-korepin determinant at a third root of unity*, Theor. Math. Physics **146** (2006) 53–62.

[174] I. Fischer and D. Romik, *More refined enumerations of alternating sign matrices*, Adv. Math. **222** (2009) 2004–2035. arXiv:0903.5073 [math.CO].

[175] M. Karklinsky and D. Romik, *A formula for a doubly refined enumeration of alternating sign matrices*, Adv. Appl. Math. **45** (2010) 28–35. arXiv:0906.3405 [math.CO].

[176] I. Fischer, *Refined enumerations of alternating sign matrices: monotone $(d, m)$-trapezoids with prescribed top and bottom rows*, J. Alg. Comb **33** (2011) 239–257. arXiv:0907.0401 [math.CO].

[177] I. Fischer and M. P. Saikia, *Refined enumeration of symmetry classes of alternating sign matrices*, J. Comb. Theory A **178** (2021) 105350. arXiv:1906.07723 [math.CO].

[178] T. Lai, *Double Aztec rectangles*, Adv. Appl. Math. **75** (2016) 1–17. arXiv:1411.0146 [math.CO].

[179] M. Adler, K. Johansson and P. van Moerbeke, *Double Aztec diamonds and the tacnode process*, Adv. Math. **252** (2014) 518–571. arXiv:1112.5532 [math.PR].

[180] E. Duse, K. Johansson and A. Metcalfe, *The cusp-Airy process*, Electron. J. Probab. **21** (2016). arXiv:1510.02057 [math.PR].

[181] M. Adler, J. Delépine and P. Van Moerbeke, *Dyson's nonintersecting Brownian motions with a few outliers*, Comm. Pure Appl. Math **62 3** (2009) 334–395. arXiv:0707.0442 [math.PR].

[182] E. Granet, L. Budzynski, J. Dubail and J. L. Jacobsen, *Inhomogeneous Gaussian free field inside the interacting arctic curve*, J. Stat. Mech.: Theory Exp. **2019 1** (2019) 013102. arXiv:1807.07927.

[183] V. Gorin, *From alternating sign matrices to the Gaussian unitary ensemble*, Commun. Math. Phys. **332 1** (2014) 437–447. arXiv:1306.6347 [math.PR].

[184] D. Zeilberger, *A fast algorithm for proving terminating hypergeometric identities*, Discrete Math. **80 2** (1990) 207–211.