Université catholique de Louvain

Faculté des sciences, École de Physique
Institut de Recherche en Mathématique et Physique, GPP

Enumeration of alternating sign matrices via the six-vertex model

Mémoire présenté en vue de l'obtention du grade académique de Master [120] en sciences physiques, finalité approfondie

Supervisor :
Prof. Christian Hagendorf

Author :
Mathias Feijtel

Readers :
Prof. Philippe Ruelle
Prof. Jan Govaerts

Academic year 2021-2022

Remerciements

Le long travail que représente un mémoire ne peut être le fruit d'une seule personne. Je profite donc de ces quelques lignes pour remercier tous ceux qui, de près ou de loin, m'ont assisté dans la réalisation de ce projet.

Je commence tout naturellement par mon promoteur, Christian Hagendorf. Ses plusieurs cours en physique statistique que j'ai eu l'honneur de suivre ainsi que notre toute première réunion cette année m'ont convaincu sans efforts de travailler avec lui. Ses nombreuses indications ainsi que le temps (si précieux aux chercheurs) qu'il m'a consacré ont été une aubaine pour moi. C'est sans parler des relectures et des nombreux commentaires dont Christian m'a fait part et qui font de ce manuscrit ce qu'il est dans sa version finale que vous allez lire. L'écriture scientifique, au même titre que le travail scientifique lui-même, est une compétence en soi et je lui dois énormément autant pour l'un que pour l'autre.

Je remercie également mes lecteurs, Philippe Ruelle et Jan Govaerts, dont les questions pertinentes lors de ma prédéfense auront aiguisé davantage ma défense à venir.

Je n'aurais pas pu travailler comme je l'ai fait pendant cette longue année sans le soutien indéfectible de mes proches, à commencer par Marie, ma compagne. Les mots d'encouragement de ma Maman, de mes frères, cousins, amis et parents de Marie, qui se sont tous intéressés malgré eux à un sujet si lointain de leurs horizons, m'ont été droit au coeur. Je ne peux terminer ces remerciements sans mentionner Papa qui de là-haut m'aura soutenu tout du long, en brillant scientifique qu'il était.

À la mémoire de Tom, Corinne, Kees et Rie.

Contents

Chapter 1

Introduction

The story of *alternating sign matrices* starts in the 1980's, when David Robbins and Howard Rumsey investigated a specific method for computing determinants known as *Dodgson condensation* [10]. It turns out that Charles Dodgson, better known as Lewis Carroll—the famous author of the *Alice* books—was also a mathematician. His way of calculating determinants is not only elegant and somewhat mysterious but also very efficient computer-wise. Surprisingly, it has gained little interest and notoriety since its discovery in the nineteenth century. The reader may consult [1, 7] for brief historical accounts of Dodgson condensation and its utility.

The way to compute the determinant of an $n \times n$ matrix $A = (a_{i,j})_{i,j=1}^{n}$ using Dodgson's method is as follows:

1. Build an $(n-1) \times (n-1)$ matrix $B = (b_{i,j})_{i,j=1}^{n-1}$ by taking all connected 2×2 subdeterminants of A:
$$b_{i,j} = \begin{vmatrix} a_{i,j} & a_{i,j+1} \\ a_{i+1,j} & a_{i+1,j+1} \end{vmatrix}. \tag{1.1}$$

2. Reiterate step 1 to find an $(n-2) \times (n-2)$ matrix $C = (c_{i,j})_{i,j=1}^{n-2}$, and divide each term by its corresponding "interior" term in the original matrix A:
$$c_{i,j} = \frac{1}{a_{i+1,j+1}} \begin{vmatrix} b_{i,j} & b_{i,j+1} \\ b_{i+1,j} & b_{i+1,j+1} \end{vmatrix}. \tag{1.2}$$

3. Repeat step 2 (dividing by the corresponding interior term of the matrix from 2 steps prior) until a 1×1 matrix is reached. Its sole entry is the determinant of the original matrix A.[1]

Example 1.1. As a way of illustrating this process, here is the computation for an arbitrary 3×3 matrix A:

$$\begin{aligned} A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \longrightarrow B = \begin{pmatrix} ae - bd & bf - ce \\ dh - eg & ei - fh \end{pmatrix} \longrightarrow \\ \longrightarrow C = \frac{1}{e}\left[(ae^2i - aefh - bdei + bdfh) - (bdfh - befg - cdeh + ce^2g)\right] \\ = aei - afh - bdi + (0)bde^{-1}fh + bfg + cdh - ceg = \det A. \end{aligned} \tag{1.3}$$

[1] If, at any step, a 0 is found in the interior of a matrix, one should rearrange the rows and columns of A such as to avoid this problem.

In their study, Robbins and Rumsey were led to define a new mathematical object that generalized determinants, they called it the λ-determinant.[2] Their main discovery was that this new type of determinant could be written as a sum indexed over a specific type of matrix [23, 29], much like the normal determinant can be seen as a sum over a set of matrices, namely permutation matrices. Using Dodgson's method to compute regular determinants, they found that the non-vanishing terms in the expansion corresponded to permutation matrices, as expected, but the vanishing terms could be associated to matrices with -1's in them. This can be seen in Example 1.1: the term bdi for instance is associated to the permutation matrix with ones in the positions b, d and i while the (vanishing) term $bde^{-1}fh$ is associated to the matrix $\left(\begin{smallmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{smallmatrix}\right)$ with a -1 in the position of e. These observations are what motivated the definition of alternating sign matrices, a new set of matrices that encompasses permutations but also allows for negative entries in the following specific way:

Definition 1.2. *An alternating sign matrix of order n, written $A = (a_{ij})_{i,j=1}^{n}$, is a square matrix that verifies the following properties:*

- $a_{ij} \in \{-1, 0, +1\}$,
- $\sum_{i=1}^{n} a_{ij} = 1$; $\sum_{j=1}^{n} a_{ij} = 1$,
- the non-zero entries alternate in sign along each row and along each column.

It is not hard to see from this definition that all permutation matrices do in fact obey these rules. For example, as can be seen in Figure 1.1a, there are seven ASMs of order 3: the six permutation matrices and only one matrix that has a -1 in it.
Another elementary observation is that ASMs only have a single $+1$ in their first row. Having a -1 in the first row would indeed violate the sum rule along the corresponding column.[3] The position of this 1 can thus be used to classify ASMs of order n into n categories.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

(a) All ASMs of order 3 arranged according to the position of the 1 in their top row.

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

(b) Example of a 5×5 ASM.

Figure 1.1: Examples of ASMs

[2] The way they defined it made it so that the value $\lambda = -1$ would recover the "classic" determinant.
[3] This fact is also true for the last row and similarly for the first and last columns.

The two mathematicians, who were then joined by a third one named William Mills, thought to ask how many of these matrices existed for arbitrary n. This simply-stated enumeration problem would prove to require enormous work and give rise to a whole new set of conjectures, connecting fields of mathematics which had not previously been related.

The first thing they did was to run numerical simulations to count the number of ASMs of increasing size. Here is the sequence that came out

$$1, 2, 7, 42, 429, 7436, 218348, 10850216, 911835460, \ldots \tag{1.4}$$

The lack of large prime factors in these numbers hinted towards the existence of a closed formula involving products of factorials. Robbins then came up with an original idea: he arranged the data in a triangle whose k^{th} entry in the n^{th} row represents the number of $n \times n$ ASMs with a 1 in the k^{th} position of their first row. The problem of counting $n \times n$ ASMs that have their 1 at the top of column k is called a *refined* enumeration.
The sum of the entries of the n^{th} row in the triangle is then equal to the total number of ASMs of order n. Taking ratios of consecutive entries lead Robbins to what he was looking for: a Pascal-like triangle [24].

					1					
				1	2/2	1				
			2	2/3	3	3/2	2			
		7	2/4	14	5/5	14	4/2	7		
	42	2/5	105	7/9	135	9/7	105	5/2	42	
429	2/6	1287	9/14	2002	16/16	2002	14/9	1287	6/2	429

Figure 1.2: Number of ASMs of order n with a 1 at the top of the k^{th} column (in black) and ratios of consecutive entries (in orange).

The third row of Figure 1.2, for instance, can be compared to what we found in Figure 1.1a. It is remarkable that the ratios of consecutive entries shown in orange are formed by summing separately the numerators and denominators of the ratios located directly above it. This fact is what ultimately allowed Robbins and his colleagues to conjecture a formula for the so-called refined count of ASMs.

Conjecture 1.3. (The refined ASM Conjecture) *Let $A_{n,k}$ represent the number of ASMs of order n with their first-row-1 in the k^{th} column. Then, for $1 \le k < n$,*

$$\frac{A_{n,k}}{A_{n,1}} = \frac{(n+k-2)!\ (2n-k-1)!}{(2n-2)!\ (k-1)!\ (n-k)!}. \tag{1.5}$$

Looking at the black entries of Figure 1.2, we also observe that the first entry of any given row is the sum of the entries in the previous row. It is not hard to see why this is true: placing the 1 of the first row in the first column of an ASM freezes the rest of the first row and column and we are left with an ASM of order $(n-1)$ to be filled in.
In other words, if we write A_n the total number of ASMs of order n, we have

$$A_n = A_{n+1,1}. \tag{1.6}$$

This process is illustrated in Figure 1.3.

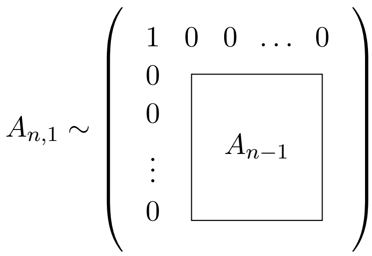


Figure 1.3: Illustration of the "freezing" process when a 1 at the top left corner occurs in an ASM of order n.

We also want to point out one last property of the triangular array Figure 1.2: it exhibits vertical symmetry. This simply stems from the fact that the vertical mirror image of an ASM is also an ASM.

All of these considerations about ASMs and the Pascal triangle aside, assuming Conjecture 1.3 holds implies a formula for the total number of $n \times n$ ASMs which is what we are primarily interested in.

Conjecture 1.4. (The ASM Conjecture) *The total number of ASMs of order n is given by*

$$A_n = \prod_{r=0}^{n-1} \frac{(3r+1)!}{(n+r)!}. \tag{1.7}$$

Our goal is to prove both conjectures separately. One can show that the refined ASM Conjecture 1.3 implies the formula for A_n but we do not take this path.

A detailed account of these conjectures (and many more related ones) along with some of their proofs and history can be found in [6]. The first appearance of these formulae is in [23], where Mills, Robbins and Rumsey prove yet another result about *plane partitions* called the *Macdonald Conjecture*. It was actually Richard Stanley, another mathematician, who pointed them towards the theory of plane partitions. In fact, he had read the work of George Andrews from just a couple of years earlier about descending plane partitions [3] where the sequence (1.4) also appeared! The three mathematicians then naturally began to search for a one-to-one correspondence between these purely combinatorial objects and ASMs. Their efforts were unsuccessful and it would take over a decade before the ASM Conjecture was finally proven.[4] The very first proof appeared in 1995 and is due to Zeilberger [34] who was also the first to prove the refined version a year later [35]. However, this first proof is very lengthy and technical and is not the focus here. The main purpose of this work is to reproduce in detail a much more simple and elegant proof found by Kuperberg [19] that uses statistical physics.

The text is structured as follows. In Chapter 2, we present the six-vertex model from statistical physics as the proof of the ASM Conjecture given by Kuperberg critically relies on a bijection between the configurations of this model and ASMs. Much of the heavy lifting will be done using the *Yang-Baxter equation*, a form of which we shall prove in this same chapter. Chapter 3 is dedicated to explaining and rigorously showing this bijection.

[4] A first bijective proof, albeit a very technical one, was only found recently [12].

As we shall see, the last piece of the puzzle lies in computing the partition function of the six-vertex model explicitly which we do in Chapter 4. Finally, in Chapter 5, we piece everything together and finish off the proof of the ASM Conjecture in two different ways: one using Cauchy determinants and the other Schur functions. To conclude this thesis, we turn our attention to the refined enumeration of ASMs. The proof we give of Conjecture 1.3 in Chapter 6 is made easier thanks to the introduction of Schur functions in the preceding chapter.

Chapter 2

The six-vertex model

Statistical physics is a branch of physics that aims at understanding macroscopic systems and their properties from the laws that govern its microscopic constituents by means of probabilistic methods. Historically, the first great achievement of this theory was being able to rederive the laws of thermodynamics [9]. Better yet, this theoretical framework gave much better predictive power than the more empirically driven field of thermodynamics. So much so in fact that it is now considered one of the main pillars of modern physics ([9] and [28] are prime examples of introductory textbooks on the matter). Indeed, far from being exclusively tied to thermodynamics, statistical physics is applied to subjects like phase transitions, superconductivity, soft matter physics, astrophysics, complex fluids, and many more. It has even found great applications well outside of physics, such as in biology, neuroscience and socio-economic networks for instance.[5]

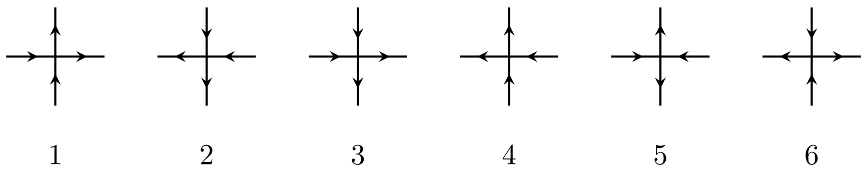


Figure 2.1: The six different vertex types of the six-vertex model.

In many areas of physics, and perhaps most so in statistical physics, we resort to the use of models in order to solve a given problem. Models are simplified versions of reality which allow for mathematical analysis, the results of which are usually surprisingly insightful. A paramount example in statistical physics is the *Ising model*, introduced by Wilhelm Lenz in 1920 [21] to study ferromagnetism and solved in the one-dimensional case by his PhD student Ernst Ising in 1925 [16].

2.1 Definition

The model we wish to study here originated from Linus Pauling's 1935 article [27] in which he made assumptions about the ways water molecules are arranged in 3D ice. Pauling was a chemist, the content and goal of his article are not our concern here but it laid the foundations for what would later be called *square ice*. Square ice is a mathematical

[5]For a detailed list of the many fields it is applied to in today's scientific research, see for example the "aims and scope" page of the *Physica A : Statistical Mechanics and its Applications* journal https://www.sciencedirect.com/journal/physica-a-statistical-mechanics-and-its-applications/about/aims-and-scope.

idealization of water that consists in H_2O molecules arranged in a two-dimensional square lattice. It was Elliott Lieb who first defined and studied this model [22].

Pauling found that, in ice crystals, the hydrogen atoms (which are tetrahedrally surrounded by four oxygen atoms) always lie on an oxygen-oxygen axis and are attached to either one of the two. Additionally, each oxygen atom has exactly two hydrogen atoms tied to it. The latter restriction came to be known as the *ice rule*. Lieb extrapolated these observations to hypothetical square ice. Figure 2.2 (left) gives an example of such a two-dimensional arrangement that respects Pauling's requirements.

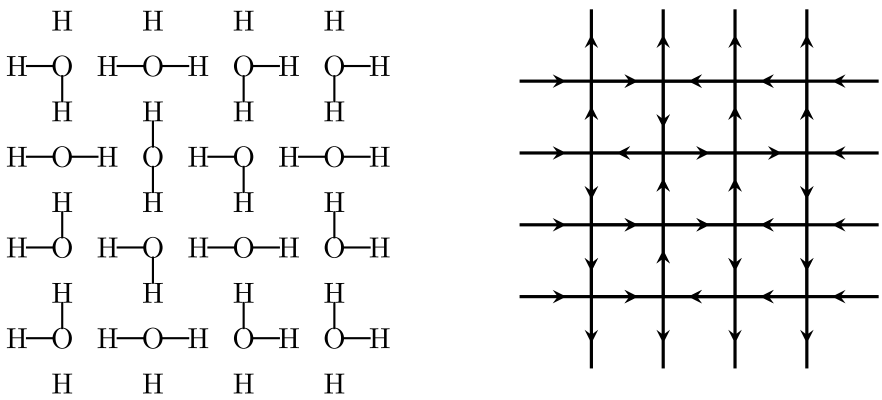


Figure 2.2: Four by four square ice configuration (left) and its equivalent six-vertex model state (right).

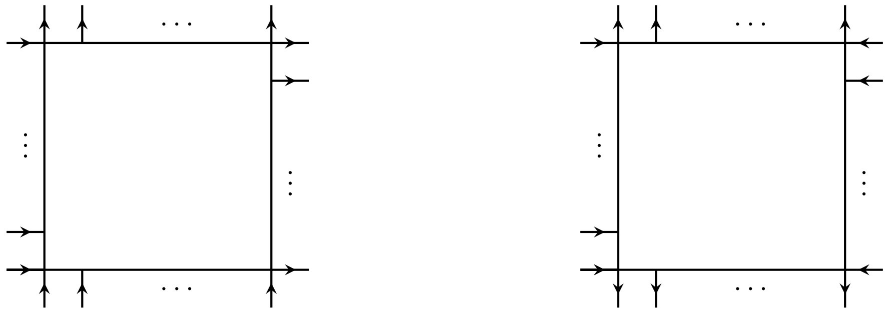
We now recall the convenient mathematical description of these configurations first given by Lieb. The way we proceed is by viewing the oxygen atoms as the vertices of a square lattice and the hydrogen atoms that are tied to it as arrows pointing towards it.[6] In Figure 2.2 (right), we see the configuration in the six-vertex model description that corresponds to the square ice arrangement on the left. In mathematical terms, we now have a directed graph for which Pauling's ice rule translates to each vertex having in-degree equal to 2 and out-degree equal to 2. That is, at each vertex, two arrows point inward and the two others outward. It is clear that this condition allows for $\binom{4}{2} = 6$ different types of vertices, hence the name of the model. These possibilities are labeled in Figure 2.1.
The word *configuration* has already been used several times throughout the previous paragraphs but it is worth clarifying what we really mean by it. A configuration—or *state*—of the model is defined as a given set of vertices assembled on a finite square grid of fixed size according to the previously established rules and with domain-wall boundary conditions, unless otherwise specified. This type of boundary conditions is described in the next subsection.

2.1.1 Boundary conditions

Another feature that arises when converting square ice configurations into six-vertex model ones is the specific behavior of the grid at its boundaries. Indeed, Pauling's rules force the left and right borders to be filled with hydrogen atoms that are tied to their oxygen atoms "inside" the grid whereas the top and bottom borders consist of arrays of hydrogen atoms that are *not* tied to the oxygen atoms inside the grid, such as in Figure 2.2.

[6]The degree of freedom thus lies on the edges of the lattice as compared to the Ising model for example where the degree of freedom simply lies on the vertices.

This property translates to inbound arrows on the left and right and outbound arrows at the top and bottom of the six-vertex model configurations. These specifications are known as *domain-wall boundary conditions* (DWBC) and are shown in Figure 2.3b. But other choices of boundary conditions exist for this model such as periodic (picture wrapping the grid around a torus) or ferroelectric for instance. The latter case is illustrated in Figure 2.3a. Clearly this is a trivial choice since it leads to all vertices being of type 1 because of the ice rule.



(a) Ferroelectric boundary conditions **(b)** Domain-wall boundary conditions (DWBC)

Figure 2.3: Examples of boundary conditions for the six-vertex model.

Generally speaking, an interesting property of the six-vertex model is its sensitivity to boundary conditions [31]. One can find a brief review of the six-vertex model with periodic boundary conditions in [32] for example. Lieb, who formally defined the model, was the first to solve[7] it with periodic boundary conditions [22].

As stated above, we are working on the square lattice but it is worth noting that the six-vertex model may be defined more generally on a *Baxter lattice*. A Baxter lattice consists in a set of straight-line segments on the two-dimensional plane which are allowed to intersect two by two (three or more lines are not allowed to intersect at the same point). The square grid is a simple example of a Baxter lattice.

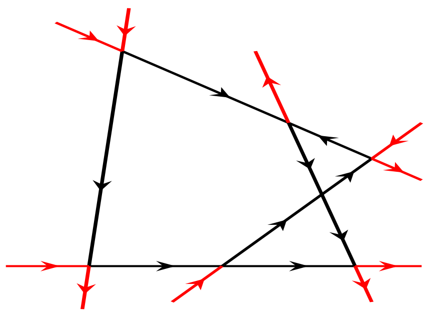


Figure 2.4: Example of a six-vertex model state defined on an arbitrary Baxter lattice. The boundary conditions depicted by red arrows are readily seen to verify Proposition 2.1.

Within this framework, we can prove the following interesting result, which restricts the possibilities for choosing boundary conditions:

[7] Solving a model in statistical physics conventionally means calculating the partition function since all the useful physical quantities derive from it (the free energy is usually of most interest). This point of view is somewhat controversial as a good number of researchers consider having solved a model once all correlation functions have been computed.

Proposition 2.1. *For any configuration of the six-vertex model on a Baxter lattice, the ice rule automatically implies that the number of outbound arrows must equal the number of inbound arrows on the border.*

Proof. Let us reason by contradiction. Suppose we have an arbitrary six-vertex model state defined on a given Baxter lattice such that the number of outbound arrows does not equal that of inbound arrows at the border.

We want to move our way through the grid by removing vertices one at a time while keeping track of the number of outbound and inbound arrows at the border. To be more precise we look at the difference Δ between *out*- and *in*-going arrows. We have $\Delta \neq 0$ by assumption. Clearly, the vertices none of whose 4 adjacent edges are on the border do not matter to us in the sense that removing them does not change Δ. We thus proceed to remove all such vertices from the grid.

All remaining vertices fall within one of three groups, they either have 1, 2 or 3 edges that belong on the border. We make the case that removing a vertex from whichever group always preserves the value of Δ. Here are the three cases:

1. The vertex has only 1 edge on the border. Without loss of generality, suppose its arrow is outbound. By virtue of the ice rule, the three other arrows of this vertex are such that two of them point towards the vertex and one points away from it. When we remove this vertex from the grid, the border arrow disappears, the two "inbound" arrows become two outbound arrows that belong to the border, and the last arrow becomes an inbound arrow on the border. In total, at the "new" border, we have lost one and created two outbound arrows, while creating one inbound arrow. This results in one extra arrow of each kind, meaning that Δ remains untouched.

2. The vertex has 2 edges belonging to the border. We face two subcases. Either the two arrows are the same kind, say outbound without loss of generality; or one is outbound and the other inbound.
   In the first subcase, removing the vertex simultaneously leads to the loss and creation of two outbound arrows, again due to the ice rule.
   In the second, we lose and create one of each kind. In any case, we thus preserve the respective number of outbound and inbound arrows. Hence Δ remains invariant.

3. The vertex has 3 edges on the border. Suppose without loss of generality that the remaining edge points towards the vertex. Removing this vertex results in the loss of three arrows (two outbound and one inbound) at the border and in the creation of one outbound arrow. In total we have thus lost one arrow of each kind and we conclude once more that Δ has not changed.

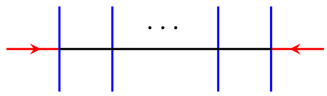
We keep removing vertices until we are inevitably left with a contradiction. Indeed, there must be a step at which at least one of the vertices violates the ice rule because of our assumption $\Delta \neq 0$. We conclude that our original hypothesis must be wrong. This means that $\Delta = 0$, hence the number of outbound arrows must equal the number of inbound arrows along the border. □

We did not state and prove Proposition 2.1 just for the sake of it. As we will see shortly, it will help us prove our first elementary result about the model on the square lattice with DWBC.

We call up/down arrows on the grid arrows that point North/South and similarly we assume right/left arrows are arrows that point East/West, written $n_\uparrow$, $n_\downarrow$, $n_\rightarrow$ and $n_\leftarrow$ respectively. These numbers obey the following simple rules:

Lemma 2.2. *For all configurations of the six-vertex model on an $n \times n$ square with DWBC, we have $n_\uparrow = n_\downarrow$ and $n_\leftarrow = n_\rightarrow$.*

Proof. Let us consider an arbitrary row of a given state of the model, along with its adjacent vertical edges



Let u and d be the respective number of up and down arrows in the lower row of vertical edges, and similarly u' and d' for the upper row of vertical edges. Clearly, we have

$$d + u = d' + u' = n. \tag{2.1}$$

Proposition 2.1 on the other hand permits us to write a second equation:

$$2 + u + d' = u' + d. \tag{2.2}$$

The difference of the previous two equations yields $d = d' + 1$ which in turn implies $u = u' - 1$. In other words, when moving from the top of the grid to the bottom, the number of up arrows decreases from n to 0 in increments of 1 at each step. Conversely, the number of down arrows increases from 0 to n in increments of 1 at each step. We conclude that, in the whole grid, the total number of down arrows $n_\downarrow$ equals the total number of up arrows $n_\uparrow$ and is given by

$$n_\uparrow = n_\downarrow = \sum_{m=0}^{n} m = n(n+1)/2. \tag{2.3}$$

The argument works exactly the same for a given column, from which it follows that $n_\rightarrow = n_\leftarrow$. □

2.1.2 Weights

To each vertex type in Figure 2.1 we associate an energy ϵ_i. As prescribed by elementary statistical mechanics theory, we define the *Boltzmann weight* of a vertex as $\omega_i = \exp(-\epsilon_i/k_B T)$, where k_B is Boltzmann's constant and T is temperature. In the six-vertex model, we define the *weight* ω of a configuration C as the product of all the respective weights of the vertices, that is

$$\omega(C) = \prod_{\text{vertices}} \omega(\text{vertex}). \tag{2.4}$$

The partition function Z in statistical physics represents the sum of the weights over the space of all configurations. In other words, we have

$$Z = \sum_{C} \prod_{\text{vertices}} \omega(\text{vertex}). \tag{2.5}$$

We assign a physical meaning to the ratio $\omega(C)/Z$: it corresponds to the probability of occurrence of configuration C in the space of all possible configurations.

Next, we define $n_i(C)$ to be the total number of vertices of type i in a given state C on an $n \times n$ lattice. With this convention, one writes the partition function (2.5) for the six-vertex model as[8]

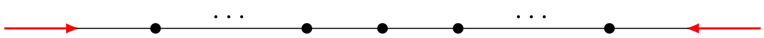
$$Z_n(\omega_1, \ldots, \omega_6) = \sum_C \omega_1^{n_1(C)} \omega_2^{n_2(C)} \omega_3^{n_3(C)} \omega_4^{n_4(C)} \omega_5^{n_5(C)} \omega_6^{n_6(C)}. \tag{2.6}$$

We dedicate the remainder of this subsection to proving the following fact: we only need three different weight types instead of six. The proof of this statement relies on two preliminary results about six-vertex model states in general. The first one will be crucial in exhibiting the link between the model and alternating sign matrices. We are thus killing two birds with one stone by properly proving this result.

It will be useful to introduce some terminology: we will call vertices of type 5 and 6 *collisions* because they are the only ones that have arrows pointing opposite each other. They will be referred to as horizontal/vertical collisions respectively.

Lemma 2.3. *In each row (and column) of the square lattice upon which the six-vertex model with DWBC is defined, the number of horizontal collisions is always one more than that of vertical collisions.*

Proof. Let us reason along a single row of the grid. The following argument will hold true for columns as well. It is clear that DWBCs (represented in red in the diagram below) force us to have at least one collision somewhere along the row. The dots on the line diagram below stand for collisions of either type.



It is impossible for two consecutive collisions to be vertical for they would create a zone which would imply a least one extra collision. Similarly, we cannot have two consecutive horizontal collisions for the zone they will have created would imply the existence of an additional collision between them.
We conclude that for any row on the grid, the collisions will appear in an alternating fashion, starting and ending with horizontal ones, thus proving there will always be one more horizontal collision relative to vertical-type collisions. □

This next lemma shows how the numbers of each type of vertices relate to one another in any given configuration.

Lemma 2.4. *Given a state C of the $n \times n$ six-vertex model, the following relations for the $n_i = n_i(C)$ always hold:*

$$n_1 = n_2; \;\; n_3 = n_4; \;\; n_5 = n_6 + n. \tag{2.7}$$

Proof. The first two equalities are consequences of Lemma 2.2. Indeed, the fact that $n_\uparrow = n_\downarrow$ directly implies that $n_1 + n_4 = n_2 + n_3$. Similarly, the left and right arrows equality translates to $n_1 + n_3 = n_2 + n_4$. These two equations combined yield the desired result, i.e $n_3 = n_4$ and $n_1 = n_2$.
Furthermore, the second relation follows trivially from Lemma 2.3 concluding the proof. □

[8] The subscript n in Z will be used throughout this manuscript to mean the $n \times n$ partition function of the model.

We are now in a position to prove the final proposition of this section.

Proposition 2.5. *The partition function for the choice of six different weights $\omega_1, \ldots, \omega_6$ is equal to the partition function for the three weights $a, b,$ and c (up to a trivial factor) by identifying $\omega_1\omega_2 \equiv a^2$, $\omega_3\omega_4 \equiv b^2$ and $\omega_5\omega_6 \equiv c^2$. The relation between the two reads*

$$Z_n(\omega_1, \ldots, \omega_6) = (\omega_5/\omega_6)^{n/2} \cdot Z_n(a, b, c). \tag{2.8}$$

Proof. By definition, and by using Lemma 2.4, we have

$$Z_n(\omega_1, \ldots, \omega_6) = \sum_C \omega_1^{n_1}\omega_2^{n_2}\omega_3^{n_3}\omega_4^{n_4}\omega_5^{n_5}\omega_6^{n_6} = \omega_5^n \sum_C (\omega_1\omega_2)^{n_1}(\omega_3\omega_4)^{n_3}(\omega_5\omega_6)^{n_6}. \tag{2.9}$$

We find, using the same lemma, that

$$Z_n(a, a, b, b, c, c) \equiv Z_n(a, b, c) = \sum_C a^{n_1+n_2}b^{n_3+n_4}c^{n_5+n_6} = c^n \sum_C a^{2n_1}b^{2n_3}c^{2n_6}. \tag{2.10}$$

Finally, applying the identification between the two sets of weights and noting that $\omega_5^n = c^n(\omega_5/\omega_6)^{n/2}$, we obtain the desired result by comparing the two previous expressions for Z_n:

$$Z_n(\omega_1, \ldots, \omega_6) = (\omega_5/\omega_6)^{n/2} \cdot Z_n(a, b, c). \tag{2.11}$$

□

We can give a physical interpretation to this result. Seeing as the arrows can only be in one of two states on each edge, we may view them as electric dipoles or spins.[9] With this point of view in mind, it seems reasonable to assume that in the absence of an external field, the energy of the system be invariant under charge reversal. In other words, a state and its counterpart built by flipping all of its arrows should be assigned an equal probability. This (physical) zero-field hypothesis thus backs our mathematical proof for the need of only three weight types. One easily computes the weight of any configuration simply by counting the numbers of vertices of each type. The one in Figure 2.2 for example contributes to the partition function $Z_5(a, b, c)$ with a weight of $\omega = a^6b^4c^6$.

2.2 The Yang-Baxter equation

For reasons that will become apparent in Chapter 3, we now focus our attention on the Yang-Baxter equation for the six-vertex model. The Yang-Baxter equation is undoubtedly one of the most fundamental equations in all of mathematical physics. One can find a good description of how broad its applications are in [25] for instance. The equation first came to life in Yang's 1967 paper [33] as way of solving a one-dimensional many body problem. It was independently (re)discovered a few years later by Rodney Baxter [4], this time to solve the eight-vertex model.[10] This equation plays a key role in *integrable systems*, i.e. systems or problems that are exactly solvable. In our case, we use it to derive an exact expression for the partition function (2.10) in Chapter 4.

[9] This justifies *a posteriori* the use of the word ferroelectric in subsection 2.1.1 about boundary conditions. In fact, this electrical jargon is quite common in statistical physics models.

[10] The eight-vertex model is a generalization of the six-vertex model we are studying here, it allows two extra vertex types: sinks and sources.

Notation: We define the *bracket* of $x \in \mathbb{C}$ (with $x \neq 0$) to be the following expression, which is used throughout this thesis:

$$[x] = x - x^{-1}. \tag{2.12}$$

First, we switch from the a,b,c parametrization to a ρ,z,q parametrization by defining

$$\begin{aligned} a(z) &= \rho\ [q^2/z], \\ b(z) &= \rho\ [z], \\ c(z) &= \rho\ [q^2]. \end{aligned} \tag{2.13}$$

We can see from this new parametrization that ρ plays no important role as it simply leads to a global factor of ρ^{n^2} in Z_n. Consequently, we choose to normalize the partition function by fixing $\rho = 1$. Furthermore, we do not make the dependence in q explicit in a, b and c for convenience of notation and because it is a simple $\mathbb{C}$-number whose value we shall assign later. The z parameter is just a $\mathbb{C}$-number as well but its importance becomes clear in the coming paragraph.

Terminology: We call q the *crossing parameter* and z the *spectral parameter*.

So far, we have considered that the weights only depend on the vertex type, which corresponds to the so-called *homogeneous* six-vertex model. We now introduce new parameters $z_1, \ldots, z_n$ and $w_1, \ldots, w_n$ associated respectively to the rows (starting from the bottom) and columns (starting from the left) of the square lattice, as can be seen in Figure 2.5. From now on, the weight of a vertex will not only depend on its type but also on its position (i, j) via the spectral parameter $z = z_i/w_j$. We call this slightly more sophisticated version of the model the *inhomogeneous* case.

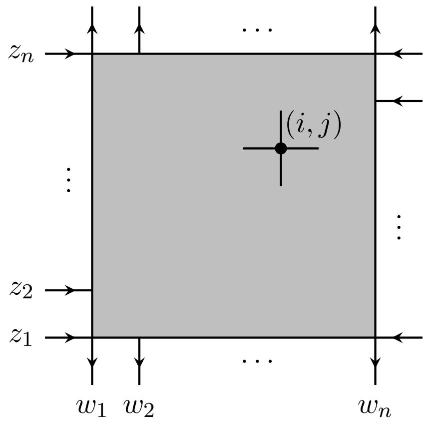
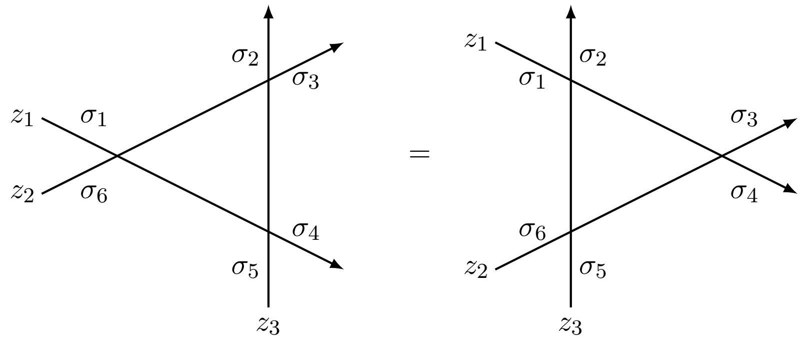


Figure 2.5: Inhomogeneous six-vertex model where vertex (i, j) has a weight with spectral parameter $z = z_i/w_j$.

The partition function is now characterized by these inhomogeneity parameters, which is to say that $Z_n = Z_n(z_1, \ldots, z_n; w_1, \ldots, w_n)$. The homogeneous case is then recovered by taking $z_1 = z_2 = \cdots = z_n = z$ and $w_1 = w_2 = \cdots = w_n = w$ so that all vertices share a common spectral parameter z/w.

The Yang-Baxter equation can take on many forms. Here we choose to present the six-vertex model version of it in the "graphical" formalism.

Theorem 2.6. (The Yang-Baxter equation for the six-vertex model) *Let z_1, z_2 and z_3 be three directed line segments.*[11] *For any choice of configuration $\{\sigma_1, \ldots, \sigma_6\}$ along the borders that respects the ice rule, the weights $a(z)$, $b(z)$ and $c(z)$ satisfy*



Note: The equal sign should be interpreted loosely, it is a shortcut for saying "the evaluation of the partition function on the left equals that of the partition function on the right".

Proof. Out of the six arrows on the border we only need to fix three since the ice rule freezes the remaining three arrows. This leads to $\binom{6}{3} = 20$ different possibilities. However, many of these are equivalent. The following observations will help us divide them into 5 distinct classes:

Observation 1 (180°) The weights of the vertices are invariant under 180° rotations. This means that if we have proven the theorem for a given set of orientations $\{\sigma_1, \ldots, \sigma_6\}$ and rotate the whole picture (both triangles) by 180°, the theorem will still hold. In other words, simultaneously applying transformations

$$\sigma_1 \leftrightarrow \sigma_4,\ \sigma_2 \leftrightarrow \sigma_5,\ \sigma_3 \leftrightarrow \sigma_6 \tag{2.14}$$

leads to the same total weight for both left and right triangles.

Observation 2 (flip) If we reflect the triangles with respect to the bisector that passes through the bottom vertex, the other two vertices get swapped and the directions of all axes get reversed. Under this transformation, all a and b-type vertices are fixed while c-type vertices get exchanged. This means the transformation

$$\sigma_1 \leftrightarrow \sigma_2,\ \sigma_3 \leftrightarrow \sigma_6,\ \sigma_4 \leftrightarrow \sigma_5 \tag{2.15}$$

is indeed weight-invariant.

Observation 3 (reverse) Finally, if we simply reverse the orientations of all line segments z_1, z_2 and z_3, then all the vertices get swapped within their respective weight group yielding yet another weight-invariant transformation. Unlike the two first ones, this transformation cannot be written properly, it simply consists in replacing each σ_i by its opposite.

[11]The black arrows on the axes represent a choice of orientations so as to unambiguously define each vertex type. One should not confuse these arrows with the degrees of freedom of the model, also represented by arrows. Note that in the square lattice these directions are implicit (pointing East for the rows and North for the columns).

We denote a set of orientations on the border as a triplet $\{\sigma_i, \sigma_j, \sigma_k\}$ which we choose to be the "inbound" arrows.
Pursuant to our observations, we find that the equivalence class of triplets that contains $\{\sigma_1, \sigma_2, \sigma_3\}$ also contains $\{\sigma_4, \sigma_5, \sigma_6\}$, $\{\sigma_2, \sigma_1, \sigma_6\}$ and $\{\sigma_3, \sigma_4, \sigma_5\}$. We obtained the first triplet by using observation 1 (or 3 equivalently), the second by using observation 2 and the third by applying observation 1 to the second triplet.
Reasoning this way, we find all other 16 triplets grouped in four distinct classes. In summary, the five classes of triplets are:

(I) $\{\sigma_1, \sigma_2, \sigma_3\}$, $\{\sigma_4, \sigma_5, \sigma_6\}$, $\{\sigma_2, \sigma_1, \sigma_6\}$, $\{\sigma_3, \sigma_4, \sigma_5\}$

(II) $\{\sigma_1, \sigma_2, \sigma_4\}$, $\{\sigma_4, \sigma_5, \sigma_1\}$, $\{\sigma_2, \sigma_1, \sigma_5\}$, $\{\sigma_3, \sigma_5, \sigma_6\}$
$\{\sigma_2, \sigma_3, \sigma_6\}$, $\{\sigma_1, \sigma_3, \sigma_6\}$, $\{\sigma_3, \sigma_4, \sigma_6\}$, $\{\sigma_2, \sigma_4, \sigma_5\}$

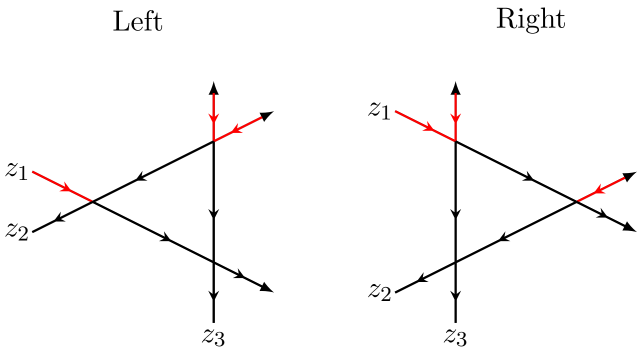
(III) $\{\sigma_2, \sigma_3, \sigma_4\}$, $\{\sigma_5, \sigma_6, \sigma_1\}$

(IV) $\{\sigma_2, \sigma_3, \sigma_5\}$, $\{\sigma_5, \sigma_6, \sigma_2\}$, $\{\sigma_1, \sigma_6, \sigma_4\}$, $\{\sigma_1, \sigma_3, \sigma_4\}$

(V) $\{\sigma_1, \sigma_3, \sigma_5\}$, $\{\sigma_2, \sigma_4, \sigma_6\}$

We now take the first triplet in each category and show that with that choice of inbound arrows on the border we do in fact get the same total weight for the left and right triangles. Cases (I), (III) and (V) turn out to be trivial in the sense that writing the total weights immediately results in the desired equality. To see through cases (II) and (IV) requires just a little more algebra. Also note that in some instances the three arrows inside the triangles are immediately set because of the ice rule and the border configuration, whereas in other cases there is some ambiguity and we must sum the two possible states.

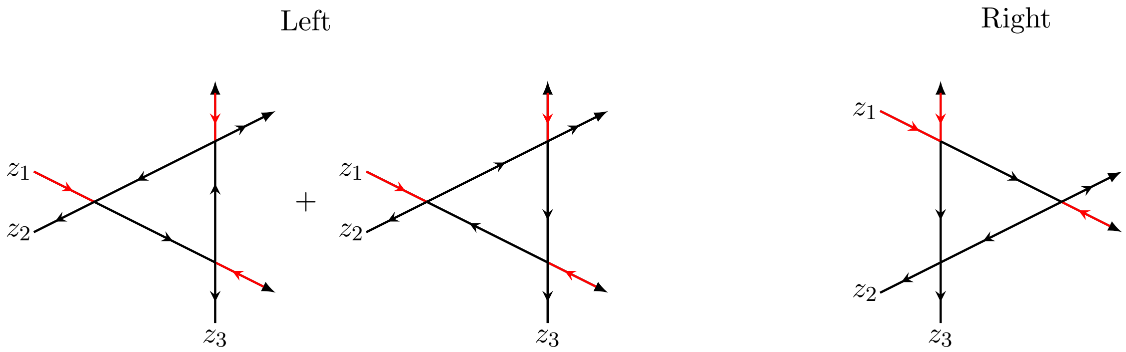
Inbound arrows $\{1, 2, 3\}$



$$\text{Total weight on the left:} \quad b(z_1/z_2)a(z_2/z_3)b(z_1/z_3).$$
$$\text{Total weight on the right:} \quad b(z_1/z_3)b(z_1/z_2)a(z_2/z_3).$$

Hence the equation is trivially satisfied in this case.

Inbound arrows $\{1,2,4\}$



$$\text{Total weight on the left:}\quad c(z_1/z_2)b(z_2/z_3)a(z_1/z_3) + b(z_1/z_2)c(z_2/z_3)c(z_1/z_3).$$
$$\text{Total weight on the right:}\quad b(z_1/z_3)c(z_1/z_2)a(z_2/z_3).$$

Note that there is a constant weight of c present in each term which we can thus dismiss. It remains to show that

$$\left[\frac{z_2}{z_3}\right]\left[\frac{q^2z_3}{z_1}\right] + \left[\frac{z_1}{z_2}\right]\left[q^2\right] = \left[\frac{z_1}{z_3}\right]\left[\frac{q^2z_3}{z_2}\right]. \tag{2.16}$$

Expanding and collecting all the terms, one finds for the left-hand side

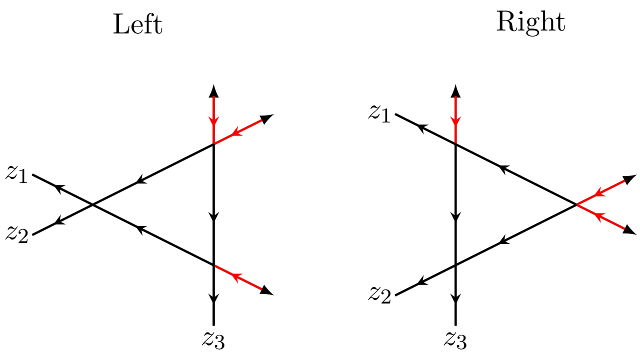
$$\frac{q^2z_2}{z_1} - \frac{z_1z_2}{q^2z_3^2} - \frac{q^2z_3^2}{z_1z_2} + \frac{z_1}{q^2z_2} + \frac{q^2z_1}{z_2} - \frac{z_1}{q^2z_2} - \frac{q^2z_2}{z_1} + \frac{z_2}{q^2z_1}. \tag{2.17}$$

After cancellation, we find exactly the four terms from the expansion of the right-hand side, namely

$$\frac{q^2z_1}{z_2} - \frac{z_1z_2}{q^2z_3^2} - \frac{q^2z_3^2}{z_1z_2} + \frac{z_2}{q^2z_1}. \tag{2.18}$$

This shows the equation holds for case (II).

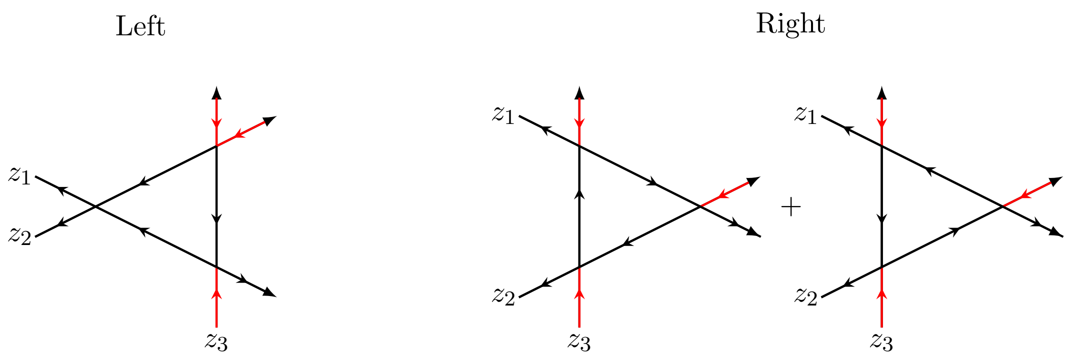
Inbound arrows $\{2,3,4\}$



$$\text{Total weight on the left:}\quad a(z_1/z_2)a(z_2/z_3)a(z_1/z_3).$$
$$\text{Total weight on the right:}\quad a(z_1/z_3)a(z_1/z_2)a(z_2/z_3).$$

Hence the equation is trivially satisfied in this case.

Inbound arrows $\{2,3,5\}$



$$\text{Total weight on the left:} \quad a(z_1/z_2)a(z_2/z_3)c(z_1/z_3).$$
$$\text{Total weight on the right:} \quad c(z_1/z_3)b(z_1/z_2)b(z_2/z_3) + a(z_1/z_2)c(z_2/z_3)c(z_1/z_3).$$

Note that there is a constant weight of c present in each term which we can thus dismiss. It remains to show that

$$\left[\frac{q^2 z_2}{z_1}\right]\left[\frac{q^2 z_3}{z_2}\right] = \left[\frac{z_1}{z_2}\right]\left[\frac{z_2}{z_3}\right] + \left[\frac{q^2 z_3}{z_1}\right]\left[q^2\right]. \tag{2.19}$$

Expanding and collecting all the terms, one finds for the right-hand side

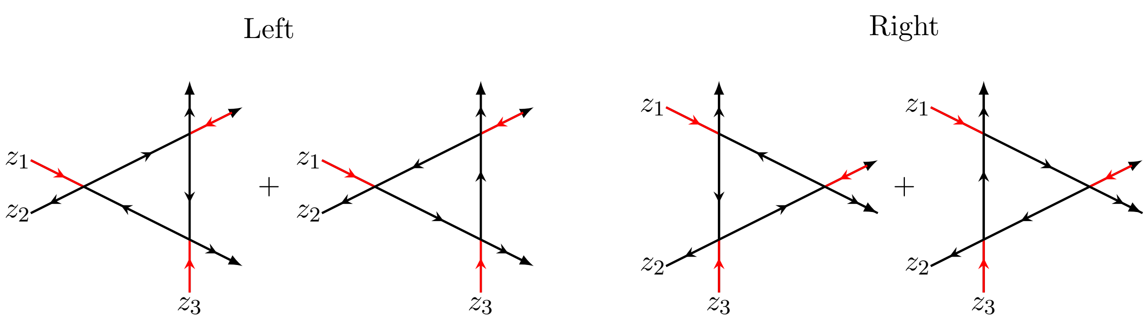
$$\frac{z_1}{z_3} - \frac{z_1 z_3}{z_2^2} - \frac{z_2^2}{z_1 z_3} + \frac{z_3}{z_1} + \frac{q^4 z_3}{z_1} - \frac{z_3}{z_1} - \frac{z_1}{z_3} + \frac{z_1}{q^4 z_3}. \tag{2.20}$$

After cancellation, we find exactly the four terms from the expansion of the left-hand side, namely

$$\frac{q^4 z_3}{z_1} - \frac{z_2^2}{z_1 z_3} - \frac{z_1 z_3}{z_2^2} + \frac{z_1}{q^4 z_3}. \tag{2.21}$$

This shows the equation holds for case (IV).

Inbound arrows $\{1,3,5\}$



$$\text{Total weight on the left:} \quad c(z_1/z_2)c(z_2/z_3)c(z_1/z_3) + b(z_1/z_2)b(z_2/z_3)a(z_1/z_3).$$
$$\text{Total weight on the right:} \quad c(z_1/z_3)c(z_1/z_2)c(z_2/z_3) + a(z_1/z_3)b(z_1/z_2)b(z_2/z_3).$$

The equation is again trivially satisfied in this case. This finishes the proof. □

Chapter 3

The bijection

The aim of this chapter is to show the link between the configurations of the six-vertex model on an $n \times n$ square with DWBC and ASMs of order n. This link is crucial, in fact it is a one-to-one correspondence between the two sets. Elkies *et al.* [11] were the first to observe this rather unexpected fact.

Once one establishes the bijection, the problem of counting ASMs becomes a problem of counting states of the six-vertex model with DWBC. To this end, we need to properly adjust the weights of the model and then reframe the problem as the computation of its associated partition function. We divide the proof of the bijection in two sections, first going from six-vertex model states with DWBC to ASMs and then the other way around.

3.1 6V-DWBC $\Longrightarrow$ ASMs

Before delving into the proof itself, we first build some intuition as to how the vertices of the model are related to the entries of ASMs. In other words, we would like to create a dictionary that links each of the six vertex types to a given entry $\{0, +1, -1\}$.

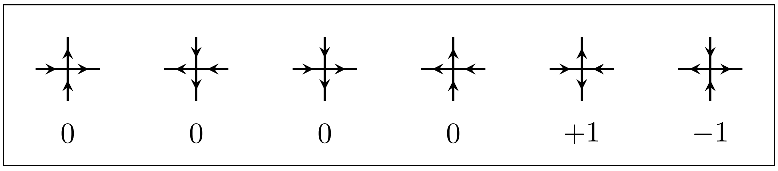
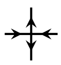


Figure 3.1: Translation between vertex types and ASM entries

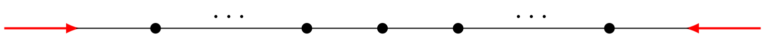
Observations:

1. With DWBC, it is impossible to place a vertical collision  on the boundaries of the square. This suggests it should be associated to the entry -1 in ASMs since -1's cannot occur on the "borders" of these matrices.

2. It is only possible to place a single horizontal collision  on each border of the square: once it is in place, it freezes the rest of its row or column because of the boundary conditions and all the other vertices are automatically of type a or b. We should thus assign these horizontal collisions to $+1$ entries in ASMs, again because there can only be a single $+1$ on each border of an ASM.

3. Finally, all the other vertices should be related to the 0's in the matrices.

Given the observations, we have good reason to believe that the dictionary in Figure 3.1 is indeed the correct one in attempting to prove the bijection. We want to show that we can associate each state of the model to a matrix and that this matrix is an ASM. It is evident from the dictionary we built that the entries of the associated matrices come from the set $\{0, +1, -1\}$. To prove the two other defining properties of ASMs—the sum rule and the alternating signs rule—we choose to reason along a single row of the model. The arguments we make easily generalize to columns. In the line graphs we use, the boundary arrows (which point inward since we are using DWBC) are represented in red and collisions (type-c vertices) are depicted as dots. Here is an example of such a row:

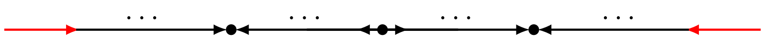


We have already proven the essence of what we are trying to show in Lemma 2.3. In that proof, we showed that collisions had to appear in alternating fashion starting and ending with horizontal collisions. Having one extra horizontal collision compared to the number of vertical collisions validates the sum rule for ASMs while the alternating property of these collisions directly translates to the alternating sign feature of ASMs.

3.2 ASMs $\implies$ 6V-DWBC

To prove this implication is to show that for each ASM of order n, there corresponds a unique $n \times n$ configuration of the model with DWBC. At first glance, this implication might seem trickier than the first because, according to the dictionary, the 0's in the matrices do not have a unique vertex to be paired to. We are about to see that this is not the case, and the reason is that the collisions alone uniquely determine the rest of the grid.

Given an ASM, we start by associating the -1's and $+1$'s to vertical and horizontal collisions respectively. Let us once again reason along a single row of the grid. In general, the picture will look like this once all of the collisions are in place:



All the arrows between the collisions are now forced to "follow the flow", meaning they should point in the direction imposed by the neighboring collisions. A very similar phenomenon happens along the columns, fixing all the arrows in the grid in a unique manner. The vertices we create by filling the grid this way are all of type a and b since these arrows never face each other.

3.3 Writing Z_n as a sum over ASMs

We first recall Lemma 2.4 from Section 2.1.2. This was a result about the number of vertices of each type in any configuration C, written $n_i(C) = n_i$ for $i = 1, \ldots, 6$. More specifically, it stated that $n_1 = n_2$, $n_3 = n_4$ and $n_5 = n_6 + n$. We know that on an $n \times n$ square, there are n^2 total vertices which yields the identity

$$\begin{aligned} &\textstyle\sum_{i=1}^{6} n_i = 2n_1 + 2n_3 + 2n_6 + n = n^2 \\ \iff &\; 2(n_1 + n_3) = n(n-1) - 2n_6. \end{aligned} \tag{3.1}$$

In light of the bijection we have just proven, the second equation of (3.1) relates the total number of 0's in a given ASM to the number of -1's.

We also recall formula (2.10) for the partition function in terms of the weights a, b and c

$$Z(a,b,c) = \sum_C a^{n_1+n_2} b^{n_3+n_4} c^{n_5+n_6} = c^n \sum_C a^{2n_1} b^{2n_3} c^{2n_6}. \tag{3.2}$$

We note that if $a = b$, Z_n becomes

$$Z_n(a,a,c) = c^n \sum_C a^{2(n_1+n_3)} c^{2n_6} = c^n a^{n(n-1)} \sum_C (c/a)^{2n_6} \tag{3.3}$$

Now assume the ratio $c/a = \sqrt{x}$ for some x. By explicitly using the bijection between the set of configurations of the six-vertex model with DWBC (on an $n \times n$ square) and the set of ASMs of order n, we express our partition function in terms of a generating function of ASMs

$$Z_n(a,a,a\sqrt{x}) = a^{n^2} x^{n/2} \sum_C x^{n_6} = a^{n^2} x^{n/2} \sum_{ASMs} x^{\#\{-1\}} \tag{3.4}$$

Finally, if we let $x = 1$ which amounts to equating all the weights, the sum over ASMs in (3.4) just ends up counting the total number of ASMs, denoted by the symbol A_n in the Introduction. We obtain this next identity

$$Z_n(a,a,a) = a^{n^2} A_n. \tag{3.5}$$

Equation (3.5) is crucial in attempting to prove the ASM Conjecture 1.4. It tells us that we need to compute the partition function explicitly and then set all the weights to be equal to one another. The next chapter deals with the computation of Z_n in the inhomogeneous case and Chapter 5 is concerned with appropriately taking the homogeneous limit in such a way that all weights become equal in order to finish off the proof.

Chapter 4

The partition function of the six-vertex model with DWBC

As we have seen in the previous chapter, it is now of interest to us to compute the partition function of the inhomogeneous six-vertex model with DWBC explicitly. To do this we first examine the main properties that this function should have according to the rules of the inhomogeneous model. Secondly, we show that the formula found by Izergin and Korepin [17] indeed possesses all of these properties and that they uniquely determine the function.

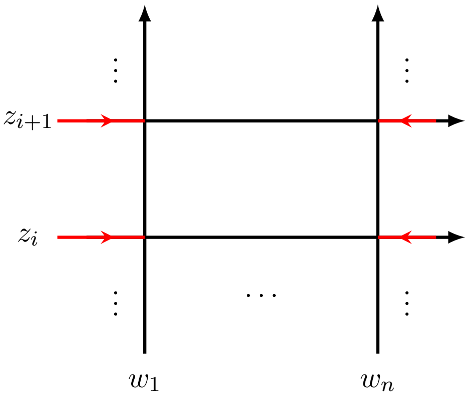
4.1 Main properties

4.1.1 Symmetry

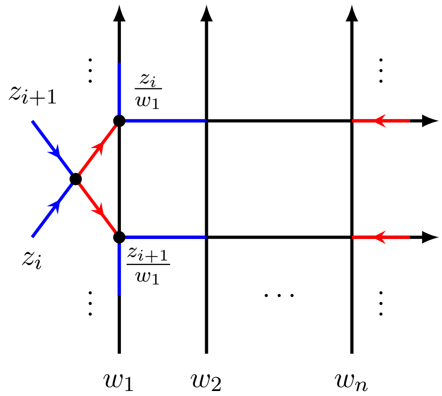
The first property we want to shed light on is one of symmetry. More specifically symmetry with respect to both sets of inhomogeneity parameters $\{z_1, \ldots, z_n\}$ and $\{w_1, \ldots, w_n\}$ separately.[12] This is where the Yang-Baxter equation comes into play. The idea is to add an extra vertex between two rows z_i and z_{i+1} at the left border of the lattice, move it across to the right using Yang-Baxter and then remove it. The partition function does not change in the process, which means we can freely exchange the two rows z_i and z_{i+1}. The calculation on the next page explains this process graphically. Exactly like in Section 2.2, one should read the equalities below as "the evaluation of the partition function in the left-hand side is the same as in the right-hand side". The red arrows represent the boundary conditions whereas the blue edges show the 6 adjacent edges to the triangles we are successively applying the Yang-Baxter equation to. Finally, some vertices have their weights explicitly labeled next to them so as to make it easier for the reader to follow the calculation step by step.

Note that in the first step, the new vertex we introduce must be of type "a" because of the boundary conditions and the ice rule. This is why we compensate by dividing by its respective weight $a(z_{i+1}/z_i)$.
Similarly, in the second to last step, the extra vertex is forced to be of type "a" as well, thereby justifying the multiplication by the same weight $a(z_{i+1}/z_i)$ at the end.

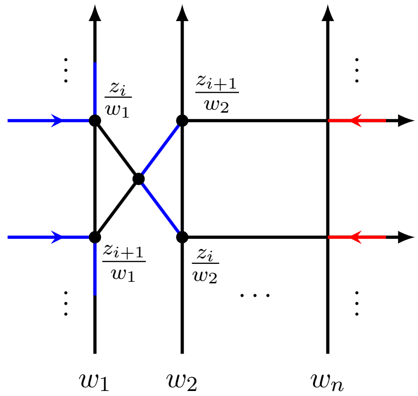
[12] With the specific value $q = e^{2i\pi/3}$ and a different choice of weights, Stroganov [30] showed that Z_n is actually symmetric with respect to *all* the parameters combined $\{z_1, \ldots, z_n\} \cup \{w_1, \ldots, w_n\}$. For now, we do not need to show this strong of a statement. We will nonetheless prove a similar fact (with $q = e^{i\pi/3}$) in Chapter 5 when using Schur functions.



$$= \frac{1}{a(z_{i+1}/z_i)}$$

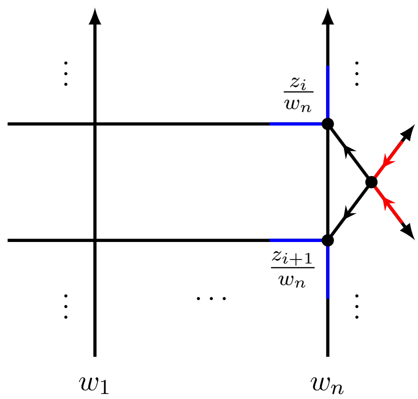


$$= \frac{1}{a(z_{i+1}/z_i)}$$

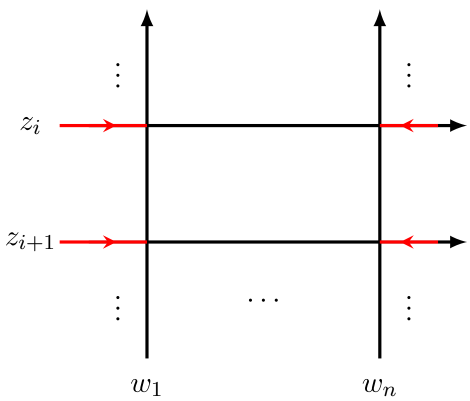


$$\vdots$$

$$= \frac{1}{a(z_{i+1}/z_i)}$$



$$=$$



In summary, this whole argument shows that we can exchange z_i and z_{i+1} without changing Z_n. A similar argument works for the columns by adding an extra vertex at the bottom and moving it across to the top and then removing it. Since any permutation can be produced from adjacent transpositions, we have successfully established the symmetry of Z_n with respect to both sets $\{z_1, \dots, z_n\}$ and $\{w_1, \dots, w_n\}$.

4.1.2 Laurent polynomials

We first recall the definition of a Laurent polynomial.

Definition 4.1. *A Laurent polynomial is a function $f : \mathbb{C} \to \mathbb{C}$ such that*

$$f(z) = \sum_{n=-k}^{l} a_n z^n, \quad \text{where } a_n \in \mathbb{C} \text{ and } a_{-k}, a_l \neq 0. \tag{4.1}$$

We refer to k and l as the *degrees* of the polynomial. If $k = l$, we say that f is a *centered* Laurent polynomial of *half width* l. Note that for $k > 1$, the polynomial is ill-defined at $z = 0$.

Our aim is to characterize $Z_n(z_1, \dots, z_n; w_1, \dots, w_n)$ as a Laurent polynomial in each of its variables z_i and w_j, with $i, j = 1, \dots, n$.

We know that each vertex has a weight which is a Laurent polynomial since they all take the form of the previously defined bracket $[x] = x - x^{-1}$. The partition function being a sum of products of these weights, it naturally inherits the property of being a Laurent polynomial in each of its variables.

The definition of the bracket also implies the "centeredness" of Z_n as a polynomial in each of its variables. Without loss of generality, let us look at Z_n as a (centered) Laurent polynomial in z_i and evaluate its half width. The row z_i, like all the others, has at least one collision which corresponds to a constant weight factor of $[q^2]$. In fact, the case of a single collision maximizes the power of z_i as one can see in Figure 4.1. Indeed, having only one collision in position j means there are $(j-1)$ type-a vertices and $(n-j)$ type-b vertices in that row, or the other way around. In both cases, the half width then equals the maximum power of z_1 in the product of $(n-1)$ brackets that all have z_i appearing to the first power, it is thus equal to[13] $(n-1)$.

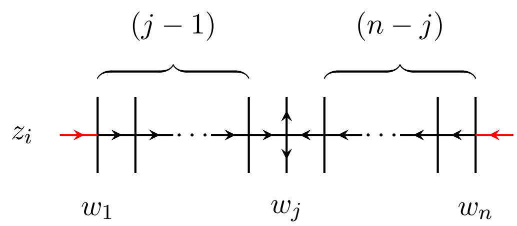


Figure 4.1: A row labeled z_i that admits a unique collision (in position j) maximizes the power of z_i in Z_n.

Although it is not necessary for our proof strictly speaking, we want to make an additional remark about parity.

[13]Notice that if z_i happens to admit more than one collision, the symmetry property allows us to exchange z_i with another row that does admit a single collision.

We observe that the parity transformation $z_i \to -z_i$ affects the weights according to:

$$\begin{cases} a(-z_i) = -a(z_i), \\ b(-z_i) = -b(z_i), \\ c(-z_i) = c(z_i). \end{cases} \tag{4.2}$$

Since there is always an odd number of collisions on any row, we conclude that the parity of Z_n with respect to the inversion of one of its components is $(n-1)$, that is

$$Z_n(z_1, \dots, -z_i, \dots, z_n; w_1, \dots, w_n) = (-1)^{n-1}\, Z_n(z_1, \dots, z_n; w_1, \dots, w_n). \tag{4.3}$$

4.1.3 Reduction relations

There is one last key observation we want to make: we can characterize Z_n inductively by looking at what happens in the bottom left corner of the lattice.[14] In fact, the DWBC only allow for 2 distinct vertex types in the corner whose weights are given by: $b(z_1/w_1) = [z_1/w_1]$ and $c(z_1/w_1) = [q^2]$.

Notation: The vector notation $\vec{x}/k$ stands for *all the components of* $\vec{x}$ *except the* k^{th} *one.*

If $z_1/w_1 = \pm 1$, then $b(z_1/w_1) = 0$ and we have a collision in the bottom left corner of the lattice. The ice rule (and DWBC) forces the rest of the first row and column to all be type-a vertices. The remainder of the lattice is simply an $(n-1)\times(n-1)$ six-vertex model state. The phenomenon is depicted in Figure 4.2. Notice that this is completely analogous to the freezing process in ASMs we described in the Introduction (see Figure 1.3). The partition function specialized to this case thus verifies

$$Z_n(z_1 = \pm w_1, z_2, \dots, z_n; w_1, \dots, w_n) = C(\vec{z}/1; \vec{w})\; Z_{n-1}(\vec{z}/1; \vec{w}/1). \tag{4.4}$$

We remark that symmetry allows us to generalize this equation to $z_i = \pm w_j$ for any $i, j = 1, \dots, n$. Expressions like (4.4) are what we call reduction relations.

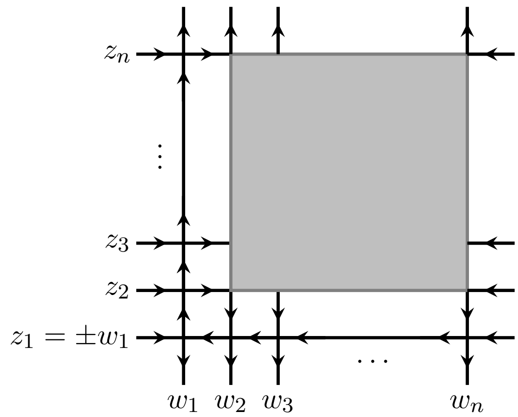


Figure 4.2: Illustration of the freezing process in the six-vertex model with DWBC when $z_1 = \pm w_1$.

One can find the expression for $C(\vec{z}/1; \vec{w})$, it corresponds to the product of weights of the vertices in the frozen row and column. To be exact, it equals

$$C(\vec{z}/1; \vec{w}) = \pm\, [q^2] \prod_{2 \le k \le n} \left[q^2 \tfrac{w_1}{w_k}\right] \left[q^2 \tfrac{z_k}{w_1}\right]. \tag{4.5}$$

[14] We could have taken any other corner but we choose the one we conventionally indexed by z_1 and w_1.

In Figure 4.2, the sublattice in gray corresponds to an $(n-1)\times(n-1)$ six-vertex model state with the appropriate boundary conditions. We might add for the keen reader that we inherit a second set of reduction relations when $z_i = q^2 w_j$, for $i, j = 1, \dots, n$.

4.2 The Izergin-Korepin formula

Izergin and Korepin proved [17] the following determinantal formula:

Theorem 4.2. *Assuming all the z_i's and all the w_i's are non-zero for $i = 1, \dots, n$, the partition function for the inhomogeneous six-vertex model with DWBC is given by*

$$Z_n(z_1, \dots, z_n; w_1, \dots, w_n) = \frac{\prod_{i,j=1}^{n} a(z_i/w_j) b(z_i/w_j)}{\prod_{1\le i<j\le n} [\frac{z_i}{z_j}][\frac{w_j}{w_i}]} \det_{i,j=1}^{n} \left(\frac{c(z_i/w_j)}{a(z_i/w_j) b(z_i/w_j)} \right). \quad (4.6)$$

Proof. We shall prove this in two steps. Namely, we need to show that:

- the three properties we studied in the previous section uniquely determine the function. As a reminder, these are (1) it is *symmetric* with respect to the z_i's and with respect to the w_j's, (2) it is a centered Laurent polynomial of half width $(n-1)$ (and well-defined parity), and finally (3) it satisfies the *reduction relations*.
- the Izergin-Korepin formula (4.6) satisfies these 3 properties. Uniqueness will then guarantee that it is in fact the partition function of the six-vertex model with DWBC.

For the first part, we shall proceed by induction. Let us suppose Z_{n-1} is known. The reduction relations for $z_1 = \pm w_j$[15] where $j \in \{1, \dots, n\}$ read

$$Z_n(z_1 = \pm w_j, \dots, z_n; w_1, \dots, w_n) = C(\vec{z}/1; \vec{w})\ Z_{n-1}(\vec{z}/1; \vec{w}/j). \quad (4.7)$$

By our induction hypothesis, we now know the value of Z_n at $2n$ different points ($z_1 = \pm w_j$, $j = 1, \dots, n$). Since $z_1^{n-1}\ Z_n$ is a polynomial of degree (at most) $2(n-1)$, it is completely determined by its value at $2n-1$ points. We are therefore done with this part of the proof.

We now turn to the problem of showing that (4.6) verifies all three properties. In what follows, let us call the numerator of the formula P, its denominator Q and the determinant part D.

1. Symmetry Let us consider the result of an exchange between some z_i and some z_j. Clearly P is symmetric under the exchange because it is simply a product across all possible i's and j's. Since $i < j$ in Q, the only term that will be affected is the one involving precisely the indices i and j being exchanged. It will turn a $[\frac{z_i}{z_j}]$ into a $[\frac{z_j}{z_i}]$ and yield a sign change. Finally, exchanging some i with some j in D amounts to switching two rows in the determinant. This brings about another sign change which cancels the other one out and proves Z_n is in fact symmetric with respect to the z_i's and with respect to the w_j's (the argument holds true for the columns as well).

[15] Note the use of symmetry in writing this.

2. <u>Laurent polynomial</u> We need to show that Z_n is a Laurent polynomial in all of its variables. Using the symmetry we just proved, we can show this for z_1 only. The same reasoning also holds for the w's.

   Clearly P and all terms containing z_1 in D are Laurent polynomials in z_1. Hence, the product PD is also a Laurent polynomial in z_1. To see that the ratio PD/Q is one as well, we only need to show that it has no pole other than at $z_1 = 0$ and $z_1 \to \infty$. We observe that it does have other poles when $z_1 = \pm z_i$, for $i = 2, \ldots, n$, and those are all simple zeroes of Q. But when $z_1 \to z_i$, two rows of D become equal and D also goes to 0 without any of its entries being singular. The determinant being a rational function of z_1, those zeroes are at least simple. We conclude that in the limit where $z_1 = \pm z_i$ $(i = 2, \ldots, n)$, the ratio PD/Q remains finite.
   We also have a potential issue when $z_1 = \pm w_j$ or when $z_1 = \pm q^2 w_j$ $(j = 1, \ldots, n)$. In these limits, D has poles, all of which are exactly compensated by the same zeroes in P. These are of the same nature, i.e single polynomial zeroes. The ratio PD/Q thus remains finite in these cases as well. This proves that Z_n is indeed a Laurent polynomial in z_1 (and in all of its other variables by symmetry).

   Furthermore, our Laurent polynomial is centered: the highest power of z_1 is the same as the highest power of z_1^{-1}, again because of the definition of the weights. To find its half width, we can reason by looking for the maximum power of z_1 in each part. For P we get a product of n terms of power z_1 for each bracket resulting in a maximum power of z_1^{2n}. For Q, there are $(n-1)$ terms of power z_1 (since $i < j$) multiplied together yielding a maximum power of $z_1^{-(n-1)}$. Finally, D only has one term in z_1 which is the product of 2 brackets giving us a maximum power of z_1^{-2}. Gathering these powers we find that the upper bound for the half width is given by $(n-1)$. We know that this bound will be attained by at least one of the z_i's. By symmetry we might as well attribute it to z_1.

   Additionally, we notice that (4.6) indeed has the right parity. The parity transformation (4.2) applied to z_1 results in P and D remaining the same. In Q however, we pick up a minus sign for each term that has z_1 in it. We just saw earlier that there are $(n-1)$ of them. Hence Z_n has the parity of $(-1)^{n-1}$, as it should.

3. <u>Reduction relations</u> Finally we want to establish the recurrence relation (4.7) for Z_n specialized to the case $z_1 = w_1$ (without loss of generality). In order to this, we need to extract all the $i = 1$ and $j = 1$ terms. Here is what we find

$$
\begin{aligned}
&Z_n(z_1 = w_1, \ldots, z_n; w_1, \ldots, w_n) = \\
&\quad = \frac{a\left(\frac{w_1}{w_2}\right)b\left(\frac{w_1}{w_2}\right)\ldots a\left(\frac{w_1}{w_n}\right)b\left(\frac{w_1}{w_n}\right)\cdot a\left(\frac{z_2}{w_1}\right)b\left(\frac{z_2}{w_1}\right)\ldots a\left(\frac{z_n}{w_1}\right)b\left(\frac{z_n}{w_1}\right)}{\left[\frac{w_1}{z_2}\right]\left[\frac{w_2}{w_1}\right]\ldots\left[\frac{w_1}{z_n}\right]\left[\frac{w_n}{w_1}\right]} \\
&\qquad \times \frac{\prod_{i,j=2}^n a\left(\frac{z_i}{w_j}\right)b\left(\frac{z_i}{w_j}\right)}{\prod_{2\le i<j\le n}\left[\frac{z_i}{z_j}\right]\left[\frac{w_j}{w_i}\right]} \times [q^2] \det_{i,j=2}^{n}\left(\frac{c\left(\frac{z_i}{w_j}\right)}{a\left(\frac{z_i}{w_j}\right)b\left(\frac{z_i}{w_j}\right)}\right) \qquad (4.8)\\
&\quad = [q^2]\left(\prod_{2\le k\le n}\left[q^2\frac{w_k}{w_1}\right]\left[q^2\frac{w_1}{z_k}\right]\right) Z_{n-1}(\vec{z}/1;\vec{w}/1).
\end{aligned}
$$

   We find the correct prefactor which we called $C(\vec{z}/1;\vec{w})$ in (4.5).

We should mention the simplifications that occurred to arrive at the final equation. The terms $b\left(\frac{w_1}{w_k}\right)$ cancel out the terms $\left[\frac{w_k}{w_1}\right]$, for all $k = 2, \ldots, n$. Similarly, the terms $b\left(\frac{z_k}{w_1}\right)$ cancel out the terms $\left[\frac{w_1}{z_k}\right]$, for all $k = 2, \ldots, n$. Together, these cancellations yield no sign change since they amount to a factor of $(-1)^{2(n-1)} = +1$. The remaining type-a terms along with the single factor of $[q^2]$ form the prefactor $C(\vec{z}/1; \vec{w})$, while the Z_{n-1} term comes from

$$\frac{\prod_{i,j=2}^{n} a\left(\frac{z_i}{w_j}\right) b\left(\frac{z_i}{w_j}\right)}{\prod_{2 \leq i < j \leq n} \left[\frac{z_i}{z_j}\right] \left[\frac{w_j}{w_i}\right]} \times \det_{i,j=2}^{n} \left(\frac{c\left(\frac{z_i}{w_j}\right)}{a\left(\frac{z_i}{w_j}\right) b\left(\frac{z_i}{w_j}\right)} \right) = Z_{n-1}(\vec{z}/1; \vec{w}/1). \tag{4.9}$$

The Izergin-Korepin formula thus satisfies all three properties. Uniqueness ensures it must be the partition function of the inhomogeneous six-vertex model with DWBC, concluding the proof.

□

Chapter 5

The proof of the ASM Conjecture

In Chapter 3, we established the bijection between ASMs and six-vertex model states with DWBC. As a reminder, the main takeaway from that section was (3.5). This equation relates the partition function of the $n \times n$ model when all of the vertex weights are equal to one another to the number of ASMs of order n via

$$Z_n|_{a=b=c} = [a]^{n^2} \; A_n. \tag{5.1}$$

Introducing inhomogeneity into the model, we were able to derive a formula for the partition function in Chapter 4, thanks to Izergin and Korepin. The natural step at this point is thus to balance all the vertex weights and analyze how it affects the partition function. In the inhomogeneous model, taking the limit

$$\begin{cases} z_1, \ldots, z_n & \to \; q, \\ w_1, \ldots, w_n & \to \; 1, \end{cases} \tag{5.2}$$

affects the weights as follows:

$$\begin{cases} a(z_i/w_j) & \to \; [q], \\ b(z_i/w_j) & \to \; [q], \\ c(z_i/w_j) & \to \; [q^2]. \end{cases} \tag{5.3}$$

The limit (5.2) takes us back to a homogeneous model where the weights only depend on the vertex type and not on the position in the lattice. Moreover, this so-called *homogeneous limit* leads to the weights a and b being equal to each other. From now on we choose to work with the crossing parameter value of $q = e^{i\pi/3}$. With this specific value of q, the following relation holds:

$$[q] = [q^2] = \sqrt{3}i. \tag{5.4}$$

Consequently, the homogeneous limit makes all the weights equal to one another which is precisely what we need. Let us recall the expression for the partition function (4.6):

$$Z_n(z_1, \ldots, z_n; w_1, \ldots, w_n) = \frac{\prod_{i,j=1}^{n} a(z_i/w_j) b(z_i/w_j)}{\prod_{1 \le i < j \le n} [\frac{z_i}{z_j}][\frac{w_j}{w_i}]} \; \det_{i,j=1}^{n} \left(\frac{c(z_i/w_j)}{a(z_i/w_j) b(z_i/w_j)} \right). \tag{5.5}$$

Notice that the homogeneous limit of (5.5) is singular. Indeed, we pick up $[1] = 0$ terms in the denominator while the determinant gets applied to a constant matrix resulting in an indeterminate form.

We therefore need to proceed with care. In this chapter, we present two distinct ways of finishing off the proof. The first follows Kuperberg's original article [19] and uses the Cauchy determinant. The second relies on the separate works of Okada [26], Stroganov [30] as well as Gorin and Panova [15], and utilizes Schur functions.

5.1 The Kuperberg approach

The way that Kuperberg [19] sidestepped the singular homogeneous limit of Z_n was to make the following substitution (with $q = e^{i\pi/3}$)

$$\begin{cases} z_i = q\ s^{i+1}, \\ w_j = s^{-j}. \end{cases} \quad \text{for all } i, j = 0, \dots, n-1, \tag{5.6}$$

The idea is that the determinant is computable with this parametrization. Once this computation is finished, we take the limit as $s \to 1$ to recover the homogeneous limit. It is important to note that we have relabeled the z and w parameters so that they run from 0 to $n-1$ instead of 1 to n, meaning we also change the indices of the products and determinant in (5.5) accordingly.[16]

Applying (5.6) to the partition function (5.5) (with the new indices) yields

$$\begin{aligned} Z_n = 3^{n/2}\,(-1)^{n/2}\, &\frac{\prod_{i,j=0}^{n-1}\left(1 + s^{2(i+j+1)} + s^{-2(i+j+1)}\right)}{\prod_{0\le i<j\le n-1}\left[s^{i-j}\right]^2} \\ &\times \det_{i,j=0}^{n-1}\left(\frac{1}{1 + s^{2(i+j+1)} + s^{-2(i+j+1)}}\right). \end{aligned} \tag{5.7}$$

We notice that

$$1 + s^{2(i+j+1)} + s^{-2(i+j+1)} = \frac{s^{3(i+j+1)} - s^{-3(i+j+1)}}{s^{i+j+1} - s^{-(i+j+1)}} = \frac{\left[s^{3(i+j+1)}\right]}{\left[s^{i+j+1}\right]}. \tag{5.8}$$

In order to compute the determinant in (5.7) in its new form (thanks to (5.8))

$$\det_{i,j=0}^{n-1}\left(\frac{1}{1 + s^{2(i+j+1)} + s^{-2(i+j+1)}}\right) = \det_{i,j=0}^{n-1}\left(\frac{\left[s^{i+j+1}\right]}{\left[s^{3(i+j+1)}\right]}\right), \tag{5.9}$$

we first need to recall the *Cauchy determinant formula* and a generalization thereof.

5.1.1 The Cauchy determinant

In the 1841 book [8], Auguste-Louis Cauchy set out to compute determinants of matrices of the form $\left(\frac{1}{x_i - y_j}\right)_{i,j=0}^{n-1}$. We refer to such matrices as *Cauchy matrices*. Cauchy himself found the following formula for which we provide a proof:

Theorem 5.1. *Let D_{n-1} denote the determinant of an $n \times n$ Cauchy matrix, where the x_i's and y_j's are all distinct. Then, for any $n \ge 1$,*

$$D_{n-1} = \det_{i,j=0}^{n-1}\left(\frac{1}{x_i - y_j}\right) = \frac{\prod_{0\le i<j\le n-1}(x_j - x_i)(y_i - y_j)}{\prod_{i,j=0}^{n-1}(x_i - y_j)}. \tag{5.10}$$

[16] This change allows to more comfortably use and prove the results in this section.

Proof. Let us first define

$$\tilde{D}_{n-1} = \prod_{j=0}^{n-1} (x_{n-1} - y_j)\ D_{n-1}. \tag{5.11}$$

We want to show that $\tilde{D}_{n-1}$ is a polynomial of degree $(n-1)$ in x_{n-1}. Let us reason by induction. The base case $n = 1$ checks out because $\tilde{D}_0 = 1$ is indeed a polynomial of degree 0 in x_0. We assume that $\tilde{D}_{k-1}$ is a polynomial of degree $(k-1)$ in x_{k-1} for some $k \in \{2, \dots, n-1\}$. For the k^{th} case, we then have

$$\tilde{D}_k = \prod_{j=0}^{k} (x_k - y_j)\ D_k. \tag{5.12}$$

We develop the determinant D_k with respect to the last row

$$D_k = \frac{1}{x_k - y_k}\ D_{k-1} - \frac{1}{x_k - y_{k-1}}\ \hat{D}_{k,k-1} + \frac{1}{x_k - y_{k-2}}\ \hat{D}_{k,k-2} - \dots\ , \tag{5.13}$$

where D_{k-1}, $\hat{D}_{k,k-1}$, $\hat{D}_{k,k-2}, \dots$ are the first minors of the matrix $\left(\frac{1}{x_i - y_j}\right)_{i,j=0}^{k}$.
We note that all of these minors are polynomials of degree 0 in x_k since we developed D_k with respect to the k^{th} row. Injecting (5.13) into (5.12), we find

$$\begin{aligned} \tilde{D}_k = \prod_{j=0}^{k-1} (x_k - y_j)\ D_{k-1}\ -\ (x_k - y_k) \prod_{j=0}^{k-2} (x_k - y_j)\ \hat{D}_{k,k-1} \\ +\ (x_k - y_k)(x_k - y_{k-1}) \prod_{j=0}^{k-3} (x_k - y_j)\ \hat{D}_{k,k-2} - \dots \end{aligned} \tag{5.14}$$

This shows that $\tilde{D}_k$ is in fact a polynomial of degree k in x_k and proves by induction that $\tilde{D}_{n-1}$ is a polynomial of degree $(n-1)$ in z_{n-1}.

We know that any determinant vanishes when two of its rows are equal. In particular, we have $\tilde{D}_{n-1}\Big|_{x_{n-1}=x_i} = 0$ for $i = 0, \dots, n-2$. Hence

$$\tilde{D}_{n-1} = \prod_{i=0}^{n-2} (x_{n-1} - x_i)\ C_{n-1}, \tag{5.15}$$

where C_{n-1} is independent of x_{n-1} because we have shown that $\tilde{D}_{n-1}$ is a polynomial of degree $(n-1)$ in z_{n-1}.
This means we may write the determinant D_{n-1} as

$$D_{n-1} = \frac{\prod_{i=0}^{n-2}\ (x_{n-1} - x_i)}{\prod_{j=0}^{n-1}\ (x_{n-1} - y_j)}\ C_{n-1}. \tag{5.16}$$

We have yet to compute the value of C_{n-1}. To this end, we consider the limit $x_{n-1} \to y_{n-1}$. We once again develop with respect to the last row, so that

$$D_{n-1} = \frac{1}{x_{n-1} - y_{n-1}}\ D_{n-2} + \sum_{i=0}^{n-2} \hat{D}_{n-1,i}. \tag{5.17}$$

Here, all of the minors $\hat{D}_{n-1,i}$ for $i = 0, \dots, n-2$ are non-singular in the limit $x_{n-1} \to y_{n-1}$, from which it follows that

$$\lim_{x_{n-1} \to y_{n-1}} (x_{n-1} - y_{n-1}) \, D_{n-1} = D_{n-2}. \tag{5.18}$$

Identities (5.16) and (5.18) imply a recurrence relation on C_{n-1} which yields (after some algebra)

$$C_{n-1} = \prod_{0 \le j < i \le n-1} \left(\frac{y_i - y_j}{x_{i-1} - y_j} \right) \prod_{0 \le i < j \le n-2} \left(\frac{x_j - x_i}{y_{j+1} - x_i} \right) \prod_{i=0}^{n-2} \left(\frac{1}{y_{i+1} - x_i} \right). \tag{5.19}$$

Substituting this value of C_{n-1} into (5.16) and combining like products, one arrives at the desired result

$$D_{n-1} = \frac{\prod_{0 \le i < j \le n-1} (x_j - x_i)(y_i - y_j)}{\prod_{i,j=0}^{n-1} (x_i - y_j)}. \tag{5.20}$$

□

It will prove relevant in upcoming calculations to compute the determinant of the matrix $\left(\frac{1}{[z_i/w_j]} \right)_{i,j=0}^{n-1}$. Cauchy's theorem 5.1 handily generalizes to such problems.

Lemma 5.2. *Let all z_i's and w_j's be distinct for $i, j = 0, \dots, n-1$. Then,*

$$\det_{i,j=0}^{n-1} \left(\frac{1}{[z_i/w_j]} \right) = \frac{\prod_{0 \le i < j \le n-1} [z_j/z_i][w_i/w_j]}{\prod_{i,j=0}^{n-1} [z_i/w_j]}. \tag{5.21}$$

Proof. First, we observe that

$$\frac{1}{z_i \, [z_i/w_j] \, w_j} = \frac{1}{z_i^2 - w_j^2}. \tag{5.22}$$

On one hand, the left-hand side of (5.22) can be seen as a product of three matrices, two of which are diagonal, so that

$$\det_{i,j=0}^{n-1} \left(\frac{1}{z_i \, [z_i/w_j] \, w_j} \right) = \frac{1}{\prod_{i=0}^{n-1} \, z_i w_i} \det_{i,j=0}^{n-1} \left(\frac{1}{[z_i/w_j]} \right). \tag{5.23}$$

On the other hand, using Cauchy's theorem 5.1 with $x_i = z_i^2$, $y_j = w_j^2$, the right-hand side of (5.22) becomes

$$\begin{aligned} \det_{i,j=0}^{n-1} \left(\frac{1}{z_i^2 - w_j^2} \right) &= \frac{\prod_{0 \le i < j \le n-1} \, z_i z_j [z_j/z_i] w_i w_j [w_i/w_j]}{\prod_{i,j=0}^{n-1} \, z_i w_j [z_i/w_j]} \\ &= \frac{1}{\prod_{i=0}^{n-1} \, z_i w_i} \frac{\prod_{0 \le i < j \le n-1} [z_j/z_i][w_i/w_j]}{\prod_{i,j=0}^{n-1} [z_i/w_j]}. \end{aligned} \tag{5.24}$$

Comparing (5.23) with (5.24), we find the desired result. □

5.1.2 Kuperberg's Lemma

Recall that our objective is to compute the determinant (5.9). The reason we spent time on Cauchy determinants is because they are useful in proving the following result, due to Kuperberg (see Lemma 10 from [19]):

Lemma 5.3. *Let*

$$\left(S_n(s,t)\right)_{i,j=0}^{n-1} = \frac{\left[s^{(i+j+1)/2}\right]}{\left[t^{(i+j+1)/2}\right]}. \tag{5.25}$$

Then,

$$\det_{i,j=0}^{n-1} S_n(s,t) = s^{-n^2/2} \prod_{k=0}^{n-1}(s-t^k)^{n-k} \prod_{k=1}^{n-1}(s-t^{-k})^{n-k} \frac{\prod_{0\le i<j\le n-1}\left[t^{-(i-j)/2}\right]^2}{\prod_{i,j=0}^{n-1}\left[t^{(i+j+1)/2}\right]}. \tag{5.26}$$

Proof. Let us look at $\det_{i,j=0}^{n-1} S_n(s,t)$ as a function of s. We observe that the "diagonal" term in the expansion of the determinant is the one that maximizes the power of s. This maximum power equals $n^2/2$.[17] We conclude that $s^{n^2/2}\ \det_{i,j=0}^{n-1} S_n(s,t)$ is a polynomial of degree n^2 in s. Furthermore, we have

$$S_n(t^k,t)_{i,j} = \frac{\left[t^{k(i+j+1)/2}\right]}{\left[t^{(i+j+1)/2}\right]}, \text{ for } k=1,\dots,n-1. \tag{5.27}$$

Moreover, for $x \neq 1$ and $k = 0,\dots,n-1$,

$$\begin{aligned} \frac{\left[t^{kx/2}\right]}{\left[t^{x/2}\right]} &= \frac{\left[t^{x/2}\right]\left(t^{(k-1)x/2}+t^{-(k-1)x/2}+t^{(k-3)x/2}+t^{-(k-3)x/2}+\dots\right)}{\left[t^{x/2}\right]} \\ &= \sum_{l=0}^{k-1} t^{x\left(l-\frac{k-1}{2}\right)}. \end{aligned} \tag{5.28}$$

Let us define the matrix $A(z)_{i,j} = z^{i+j+1}$, for $i,j = 0,\dots,n-1$. We then have

$$S_n(t^k,t)_{i,j} = \sum_{k=0}^{n-1} A\left(t^{l-\frac{k-1}{2}}\right)_{i,j}. \tag{5.29}$$

Notice that A is a matrix of rank 1 since each of its rows is a multiple of the first row ($a_{i,*} = z^i a_{0,*}$). It follows that the matrix $S_n(t^k,t)$ is at most[18] of rank k. Hence, at the points $s = t^k$, the matrix S_n does not have maximal rank and the determinant vanishes. These zeroes are of order $(n-k)$. The same argument applies to the points $s = t^{-k}$ (for $k = 1,\dots,n-1$) and yields a second set of zeroes whose order is also $(n-k)$. We deduce that $\det S_n(s,t)$ is divisible by $(s-t^k)^{n-k}$ and by $(s-t^{-k})$, hence

$$s^{n^2}\ \det_{i,j=0}^{n-1} S_n(s,t) = \prod_{k=0}^{n-1}(s-t^k)^{n-k} \prod_{k=1}^{n-1}(s-t^{-k})^{n-k}\ C_n(t). \tag{5.30}$$

The factor $C_n(t)$ is independent of s because $\prod_{k=0}^{n-1}(s-t^k)^{n-k}\prod_{k=1}^{n-1}(s-t^{-k})^{n-k}$ is already a polynomial of degree n^2 in s. We still have to compute the value of $C_n(t)$. To this end,

[17] This is because the product of all diagonal terms reads $s^{1/2+3/2+\dots+(2n-1)/2} = s^{n^2/2}$.
[18] Because of the well-known inequality $\mathrm{rank}(A+B) \le \mathrm{rank}(A) + \mathrm{rank}(B)$.

we take the limit $s \to 0$ on both sides of (5.30). Indeed, in this limit, the right-hand side reads

$$\lim_{s\to 0} \prod_{k=0}^{n-1}(s-t^k)^{n-k} \prod_{k=1}^{n-1}(s-t^{-k})^{n-k}\ C_n(t) = (-1)^n\ C_n(t). \tag{5.31}$$

As for the left-hand side, we may write

$$\lim_{s\to 0}\ s^{n^2/2} \det_{i,j=0}^{n-1} S_n(s,t) = \lim_{s\to 0}\ s^{n^2/2} \det_{i,j=0}^{n-1}\left(s^{-(i+j+1)/2}\ \tilde{S}_{ij}\right), \tag{5.32}$$

where

$$\tilde{S}_{ij} = \frac{s^{i+j+1}-1}{[t^{(i+j+1)/2}]} \tag{5.33}$$

is the non-singular part of the determinant of $S_n(s,t)$ in the limit $s \to 0$. We further decompose the determinant in (5.32) as follows:

$$\lim_{s\to 0}\ s^{n^2/2} \det_{i,j=0}^{n-1}\left(s^{-(i+j+1)/2}\ \tilde{S}_{ij}\right) = \lim_{s\to 0}\ s^{n^2/2} \det_{i,j=0}^{n-1}(U\ \tilde{S}\ V), \tag{5.34}$$

where U and V are diagonal matrices defined by $(U_{ij}) = (u_i\delta_{ij})$, and $(V_{ij}) = (v_j\delta_{ij})$ for $u_i = s^{-i/2}$ and $v_j = s^{-(j+1)/2}$. We then find

$$\begin{aligned}\lim_{s\to 0}\ s^{n^2/2} \det_{i,j=0}^{n-1}(U\ \tilde{S}\ V) &= \lim_{s\to 0}\ s^{n^2/2} \overbrace{\prod_{i=0}^{n-1} s^{-i/2}}^{=\ s^{-n(n-1)/4}} \overbrace{\prod_{j=0}^{n-1} s^{-(j+1)/2}}^{=\ s^{-n(n+1)/4}} \det_{i,j=0}^{n-1} \tilde{S}_{ij} \\ &= (-1)^n \det_{i,j=0}^{n-1}\left(\frac{1}{[t^{(i+j+1)/2}]}\right).\end{aligned} \tag{5.35}$$

We compute this last determinant in (5.35) using the "generalized" Cauchy Lemma 5.2 with $z_i = t^{i/2}$ and $w_j = t^{-(j+1)/2}$ to find

$$\det_{i,j=0}^{n-1}\left(\frac{1}{[t^{(i+j+1)/2}]}\right) = \frac{\prod_{0\le i<j\le n-1}\left[t^{(j-i)/2}\right]^2}{\prod_{i,j=0}^{n-1}\left[t^{(i+j+1)/2}\right]}. \tag{5.36}$$

In summary, the limit as $s \to 0$ in the left-hand side of (5.30) reads

$$\lim_{s\to 0}\ s^{n^2/2} \det_{i,j=0}^{n-1} S_n(s,t) = (-1)^n \frac{\prod_{0\le i<j\le n-1}\left[t^{(j-i)/2}\right]^2}{\prod_{i,j=0}^{n-1}\left[t^{(i+j+1)/2}\right]}. \tag{5.37}$$

Comparing this expression with that of the limit on the right-hand side (5.31) yields the value of $C_n(t)$

$$C_n(t) = \frac{\prod_{0\le i<j\le n-1}\left[t^{(j-i)/2}\right]^2}{\prod_{i,j=0}^{n-1}\left[t^{(i+j+1)/2}\right]}. \tag{5.38}$$

Plugging (5.38) into (5.30) and dividing both sides by $s^{n^2/2}$ produces the final result (5.26). □

5.1.3 Computations

Kuperberg' Lemma 5.3 puts us in the right direction but we now wish to reformulate part of the right-hand side of (5.26) in a more convenient way using the following result:

Proposition 5.4. *We have*

$$s^{-n^2/2}\prod_{k=0}^{n-1}(s-t^k)^{n-k}\prod_{k=1}^{n-1}(s-t^{-k})^{n-k}=\prod_{i,j=0}^{n-1}[s^{1/2}t^{(i-j)/2}]. \tag{5.39}$$

Proof. Let $k=j-i$, then

$$\begin{aligned}\prod_{i,j=0}^{n-1}[s^{1/2}t^{(i-j)/2}] &= \prod_{i=0}^{n-1}\prod_{k=-i}^{n-i-1}[s^{1/2}t^{-k/2}]\\ &= \prod_{i=0}^{n-1}\prod_{k=-i}^{n-i-1} s^{-1/2}\ t^{-k/2}\ (s-t^k)\\ &= s^{-n^2/2}\prod_{i=0}^{n-1} t^{-\frac{1}{2}\left(\frac{n(n-1)}{2}-in\right)}\prod_{k=-i}^{n-i-1}(s-t^k)\\ &= s^{-n^2/2}\ t^{-\frac{n^2(n-1)}{4}+\frac{n^2(n-1)}{4}}\prod_{i=0}^{n-1}\left(\prod_{k=-i}^{n-i-1}(s-t^k)\right)\\ &= s^{-n^2/2}\prod_{i=0}^{n-1}\left(\prod_{k=1}^{i}(s-t^{-k})\prod_{k=0}^{n-i-1}(s-t^k)\right).\end{aligned} \tag{5.40}$$

We reindex using $i'=n-i-1$ in the second product inside the parentheses of the last line in (5.40) to finally find

$$\begin{aligned}\prod_{i,j=0}^{n-1}[s^{1/2}t^{(i-j)/2}] &= s^{-n^2/2}\prod_{k=1}^{n-1}\prod_{i=k}^{n-1}(s-t^{-k})\prod_{k=0}^{n-1}\prod_{i'=k}^{n-1}(s-t^k)\\ &= s^{-n^2/2}\prod_{k=0}^{n-1}(s-t^k)^{n-k}\prod_{k=1}^{n-1}(s-t^{-k})^{n-k}.\end{aligned} \tag{5.41}$$

□

This last result allows us to write the determinant of $S_n(s,t)$ in Kuperberg's Lemma 5.3 as

$$\det_{i,j=0}^{n-1} S_n(s,t)=\det_{i,j=0}^{n-1}\frac{[s^{(i+j+1)/2}]}{[t^{(i+j+1)/2}]}=\prod_{0\le i<j\le n-1}\left[t^{-(i-j)/2}\right]^2\prod_{i,j=0}^{n-1}\frac{[s^{1/2}t^{(i-j)/2}]}{[t^{(i+j+1)/2}]}. \tag{5.42}$$

Recall from (5.7) and (5.9) that our aim is to compute

$$\det_{i,j=0}^{n-1}\left(\frac{[s^{i+j+1}]}{[s^{3(i+j+1)}]}\right). \tag{5.43}$$

We therefore use (5.42) and replace s with s^2 and t with s^6 which yields

$$\det_{i,j=0}^{n-1}\left(\frac{[s^{i+j+1}]}{[s^{3(i+j+1)}]}\right)=\prod_{0\le i<j\le n-1}\left[s^{-3(i-j)}\right]^2\prod_{i,j=0}^{n-1}\frac{[s^{3(i-j)+1}]}{[s^{3(i+j+1)}]}. \tag{5.44}$$

Injecting this last result into (5.7), we obtain the following expression for partition function when $z_i = q\ s^{i+1},\ w_j = s^{-j}$, with $i, j = 0, \dots, n-1$, and $q = e^{i\pi/3}$:

$$Z_n(q\ s^{i+1};\ s^{-j}) = (-1)^{n/2}\ 3^{n/2} \prod_{i,j=0}^{n-1} \frac{[s^{3(i-j)+1}]}{[s^{i+j+1}]} \prod_{0\le i<j\le n-1} \left(\frac{[s^{-3(i-j)}]}{[s^{i-j}]} \right)^2 . \tag{5.45}$$

5.1.4 The homogeneous limit

The final step towards the proof of the ASM Conjecture consists in taking the limit $s \to 1$ in (5.45). Indeed, this limit is equivalent to the homogeneous limit of the six-vertex model (with DWBC) where all the weights are equal to one another.

One needs the following result in order to compute the value of Z_n in the limit $s \to 1$:

Proposition 5.5. *For any x and $y \neq 0$, we have*

$$\lim_{s\to 1} \frac{[s^x]}{[s^y]} = \frac{x}{y}. \tag{5.46}$$

Proof. Using the definition of the bracket and L'Hospital's rule, one directly finds

$$\lim_{s\to 1} \frac{[s^x]}{[s^y]} = \lim_{s\to 1} \frac{s^x - s^{-x}}{s^y - s^{-y}} = \lim_{s\to 1} \frac{x s^{x-1} + x s^{-x-1}}{y s^{y-1} + y s^{-y-1}} = \frac{x}{y}. \tag{5.47}$$

□

Taking the limit $s \to 1$ in (5.45), we find using Proposition 5.5

$$Z_n(q, \dots, q; 1, \dots, 1) = (-1)^{n/2}\ 3^{n/2+n(n-1)} \prod_{i,j=0}^{n-1} \frac{3i - 3j + 1}{i + j + 1}. \tag{5.48}$$

Let us compute the numerator N and denominator D of the double product that appears in (5.48) separately.

Numerator:

We have

$$\begin{aligned} N = \prod_{i,j=0}^{n-1} (3i - 3j + 1) &= \prod_{0\le i<j\le n-1} (3i - 3j + 1) \prod_{0\le j<i\le n-1} (3i - 3j + 1) \\ &= (-1)^{n(n-1)/2} \prod_{0\le j<i\le n-1} \big(3(i-j) - 1\big) \big(3(i-j) + 1\big) . \end{aligned} \tag{5.49}$$

Next, we define the index $k = i - j$, which allows us to write

$$\begin{aligned} N &= (-1)^{n(n-1)/2} \prod_{j=0}^{n-2} \prod_{k=1}^{n-1-j} \frac{(3k-1)\ 3k\ (3k+1)}{3k} \\ &= (-1)^{n(n-1)/2} \prod_{j=0}^{n-2} \frac{1}{3^{n-1-j}(n-1-j)!} \prod_{k=1}^{n-1-j} (3k-1)\ 3k\ (3k+1). \end{aligned} \tag{5.50}$$

We now make the transformation $j \to n-1-j$ to find

$$\begin{aligned} N &= (-1)^{n(n-1)/2} \prod_{j=1}^{n-1} \frac{1}{3^j j!} \prod_{k=1}^{j} (3k-1)\ 3k\ (3k+1) \\ &= (-1)^{n(n-1)/2}\ 3^{-n(n-1)/2} \prod_{j=1}^{n-1} \frac{(3j+1)!}{j!}. \end{aligned} \tag{5.51}$$

Notice we may include the $j = 0$ term in this last product since it corresponds to a factor of 1.

Denominator:

We are able to compute D directly

$$D = \prod_{i,j=0}^{n-1} (i+j+1) = \prod_{j=0}^{n-1} (j+1)(j+2)\dots(j+n) = \prod_{j=0}^{n-1} \frac{(n+j)!}{j!}. \tag{5.52}$$

Combining what we found for N (5.51) and D (5.52) and injecting these formulas into (5.48) yields

$$Z_n(q,\dots,q;1,\dots,1) = (-1)^{n^2/2}\ 3^{n^2/2} \prod_{j=0}^{n-1} \frac{(3j+1)!}{(n+j)!}. \tag{5.53}$$

Finally, we recall the combinatorial relation (3.5) between the partition function and A_n, the number of ASMs of order n, in the homogeneous limit:

$$Z_n(q,\dots,q;1,\dots,1) = [q]^{n^2} A_n. \tag{5.54}$$

When combining (5.53) with (5.54), the powers of (-1) and 3 cancel out and we are left with the desired result

$$A_n = \prod_{j=0}^{n-1} \frac{(3j+1)!}{(n+j)!}. \tag{5.55}$$

This concludes our first proof of the ASM Conjecture 1.4.

5.2 The Okada/Stroganov approach

5.2.1 Okada's determinant

We choose to present a second approach to solving the ASM Conjecture that links the determinant in (5.5) with $q = e^{i\pi/3}$ to *Schur functions*. This way of proving the conjecture and its refined version was first introduced by Stroganov [30].

First, we replicate part of Theorem 3.3 from Okada's paper [26] as it is the main result we need. We do not provide proof of this statement as it is well beyond the scope of this work.

Theorem 5.6. *For* $\vec{u} = (u_1, \dots, u_n)$*, and* $\vec{\alpha} = (\alpha_1, \dots, \alpha_n)$*, let* $V^{n,n}(\vec{u}, \vec{\alpha})$ *be a* $2n \times 2n$ *matrix whose* i^{th} *row is given by*

$$(1, u_i, u_i^2, \dots, u_i^{n-1}, \alpha_i, \alpha_i u_i, \alpha_i u_i^2, \dots, \alpha_i u_i^{n-1}).$$

For $\vec{x} = (x_1, \dots, x_n)$*,* $\vec{y} = (y_1, \dots, y_n)$*,* $\vec{a} = (a_1, \dots, a_n)$ *and* $\vec{b} = (b_1, \dots, b_n)$*, we have*

$$\det_{i,j=1}^{n} \left(\frac{b_j - a_i}{y_j - x_i} \right) = \frac{(-1)^{n(n-1)/2}}{\prod_{i,j=1}^{n} (y_j - x_i)} \ \det V^{n,n}(\vec{x}, \vec{y}; \vec{a}, \vec{b}). \tag{5.56}$$

We now take the Izergin-Korepin formula (5.5) and rewrite the determinant in such a way that we can use Theorem 5.6. The reason we do this is because the $2n \times 2n$ determinant that we get can be related to a Schur function.

We rewrite the product $a(z_i/w_j)b(z_i/w_j)$ that appears in the determinant and in the numerator of its prefactor in (5.5) as (recall that $q = e^{i\pi/3}$)

$$\begin{aligned} a(z)b(z) &= \frac{(q^2 z^{-1} - q^{-2} z)(z - z^{-1})(q^2 z - q^{-2} z^{-1})}{(q^2 z - q^{-2} z^{-1})} \\ &= \frac{q^3 z^3 - (q^3 z^3)^{-1} + z \overbrace{(1 - q - q^{-1})}^{=0} + z^{-1} \overbrace{(q + q^{-1} - 1)}^{=0}}{q^2 z - (q^2 z)^{-1}} \\ &= \frac{[(qz)^3]}{[q^2 z]}. \end{aligned} \tag{5.57}$$

Therefore, the determinant in (5.5) takes the form

$$\begin{aligned} \det_{i,j=1}^{n} \left(\frac{c(z_i/w_j)}{a(z_i/w_j) b(z_i/w_j)} \right) &= \left[q^2\right]^n \det_{i,j=1}^{n} \left(\frac{(q^2 z_i/w_j) - (q^2 z_i/w_j)^{-1}}{(q z_i/w_j)^3 - (q z_i/w_j)^{-3}} \right) \\ &= (-1)^{n/2}\ 3^{n/2} \det_{i,j=1}^{n} \left(\frac{\frac{-1}{q^2 z_i w_j} \left(w_j^2 - (q^2 z_i)^2 \right)}{\frac{1}{(q^2 z_i w_j)^3} \left(w_j^6 - (q^2 z_i)^6 \right)} \right) \\ &= (-1)^{3n/2}\ 3^{n/2} \prod_{i=1}^{n} (q^2 z_i w_i)^2 \det_{i,j=1}^{n} \left(\frac{w_j^2 - (q^2 z_i)^2}{w_j^6 - (q^2 z_i)^6} \right). \end{aligned} \tag{5.58}$$

Using Theorem 5.6 with $\vec{x} = \left((q^2 z_1)^6, \dots, (q^2 z_n)^6\right)$, $\vec{y} = (w_1^6, \dots, w_n^6)$, $\vec{a} = (q^2 z_1, \dots, q^2 z_n)$ and $\vec{b} = (w_1^2, \dots, w_n^2)$, we find

$$\det_{i,j=1}^{n} \left(\frac{c(z_i/w_j)}{a(z_i/w_j) b(z_i/w_j)} \right) = (-1)^{n(n+2)/2}\ 3^{n/2} \frac{\prod_{i=1}^{n} \left(\frac{z_i w_i}{q} \right)^2}{\prod_{i,j=1}^{n} (w_j^6 - z_i^6)} \ \det V^{n,n}(\vec{x}, \vec{y}; \vec{a}, \vec{b}). \tag{5.59}$$

Next, we define $\mathbf{X}$ as the concatenation of $\vec{x}$ and $\vec{y}$ and $\mathbf{A}$ as the concatenation of $\vec{a}$ and $\vec{b}$

$$\begin{aligned} \mathbf{X} &= \left((\tfrac{z_1}{q})^6, \dots, (\tfrac{z_n}{q})^6, w_1^6, \dots, w_n^6 \right), \\ \mathbf{A} &= \left((\tfrac{z_1}{q})^2, \dots, (\tfrac{z_1}{q})^2, w_1^2, \dots, w_n^2 \right). \end{aligned} \tag{5.60}$$

The i^{th} row of the matrix $V^{n,n}(\mathbf{X},\mathbf{A})$ is

$$\begin{aligned}&\left(1, (\tfrac{z_i}{q})^6, (\tfrac{z_i}{q})^{12}, \dots, (\tfrac{z_i}{q})^{6(n-1)}, (\tfrac{z_i}{q})^2, (\tfrac{z_i}{q})^8, \dots, (\tfrac{z_i}{q})^{6n-4}\right), \quad \text{for } i = 1, \dots, n,\\ &\left(1, w_{i-n}^6, w_{i-n}^{12}, \dots, w_{i-n}^{6(n-1)}, w_{i-n}^2, w_{i-n}^8, \dots, w_{i-n}^{6n-4}\right), \qquad \text{for } i = n+1, \dots, 2n.\end{aligned} \tag{5.61}$$

If we define

$$Y_i = \begin{cases} \left(\frac{z_i}{q}\right)^2, & \text{for } i = 1, \dots, n, \\ w_{i-n}^2, & \text{for } i = n+1, \dots, 2n, \end{cases} \tag{5.62}$$

then the i^{th} row of $V^{n,n}(\mathbf{X},\mathbf{A})$ becomes

$$\left(1, Y_i^3, Y_i^6, \dots, Y_i^{3(n-1)}, Y_i, Y_i^4, Y_i^7, \dots, Y_i^{3n-2}\right). \tag{5.63}$$

We can summarize our findings as follows:

$$\left(V^{n,n}(\mathbf{X},\mathbf{A})\right)_{i,j=1}^{2n} = \begin{cases} Y_i^{3(j-1)}, & \text{for } j = 1, \dots, n, \\ Y_i^{3(j-n)-2}, & \text{for } j = n+1, \dots, 2n. \end{cases} \tag{5.64}$$

5.2.2 Schur functions

We want to recall the concept of Schur functions. The reader can find a detailed account of Schur functions including their definition, main properties, as well as their link to combinatorial objects such as plane partitions and Young tableaux in Chapter 4 of Bressouds's book [6]. Schur functions admit several different definitions, we choose to present the one which most closely relates to our problem. These functions are indexed by integer partitions. It thus seems worthwhile to recall what these are: an integer partition $\lambda = (\lambda_1, \dots, \lambda_n)$ of size n is a list of integers such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and whose sum equals a certain integer. For example, $(4, 4, 1)$ is a partition of the number 9. With integer partitions in mind, we may recall the definition of Schur functions.

Definition 5.7. *Let $\lambda = (\lambda_1, \dots, \lambda_m)$ be an integer partition of size m and let δ be the "staircase" partition of size m defined by $\delta_i = m - i$, for $i = 1, \dots, m$. A Schur function s_λ is the ratio of two determinants:*

$$s_\lambda(x_1, \dots, x_m) = \frac{\det_{i,j=1}^m \left(x_i^{\lambda_j + \delta_j}\right)}{\det_{i,j=1}^m \left(x_i^{\delta_j}\right)}. \tag{5.65}$$

The determinant in the denominator is a *Vandermonde determinant*, it is well known that they are equal to the following product:

$$\det_{i,j=1}^{m} \left(x_i^{m-j}\right) = \prod_{1 \leq i < j \leq m} (x_i - x_j). \tag{5.66}$$

Notice that Schur functions are polynomials because the numerator is an alternating polynomial and is thus divisible by a Vandermonde determinant. Also notice that these polynomials are always symmetric since the sign changes in the numerator cancel those in the denominator for any transposition.

Consider the partition $\lambda = (n-1, n-1, n-2, n-2, \ldots, 1, 1, 0, 0)$, defined by $\lambda_i = \lfloor n - i/2 \rfloor$ for $i = 1, \ldots, 2n$, where $\lfloor x \rfloor$ is the floor function. We refer to this partition as the *double staircase partition*. We give the first three Schur polynomials for this particular partition in the following example:[19]

Example 5.8. *For λ the double staircase partition, we find*

$$\begin{aligned} &s_\lambda(x_1, x_2) = 1, \\ &s_\lambda(x_1, x_2, x_3, x_4) = x_1 x_2 + (x_1 + x_2)\ x_3 + (x_1 + x_2 + x_3)\ x_4, \\ &s_\lambda(x_1, \ldots, x_6) = x_1^2 x_2^2 x_3 x_4 + x_1^2 x_2^2 x_3 x_5 + x_1^2 x_2^2 x_3 x_6 + x_1^2 x_2^2 x_4 x_5 + x_1^2 x_2^2 x_4 x_6 + \ldots \end{aligned} \tag{5.67}$$

We observe that the double staircase partition yields the exact powers of Y in $V^{n,n}$ (5.64) when combined with the (simple) staircase partition $\delta = (2n-1, 2n-2, \ldots, 1, 0)$ since

$$\lambda + \delta = (3n-2, 3n-3, 3n-5, 3n-6, \ldots, 4, 3, 1, 0). \tag{5.68}$$

We thus rearrange the columns of $V^{n,n}$ so that its i^{th} row looks like

$$(Y_i^{3n-2}, Y_i^{3n-3}, Y_i^{3n-5}, \ldots, Y_i^4, Y_i^3, Y_i, 1). \tag{5.69}$$

The number of permutations required to go from (5.63) to (5.69) yields a sign factor in the determinant equal to

$$\begin{gathered} (-1)^{2n-1}\ (-1)^{n-1}\ (-1)^{2n-3}\ (-1)^{n-2}\ \ldots\ (-1)^{2n-(2n-1)}\ (-1)^{n-n} \\ = \\ (-1)^{\sum_{j=1}^{n} 2j-1}\ (-1)^{\sum_{k=0}^{n-1} k} = (-1)^{n+n(n-1)/2}, \end{gathered} \tag{5.70}$$

We can finally express the determinant of $V^{n,n}$ as a Schur polynomial multiplied by a Vandermonde determinant. Gathering all the factors, (5.59) becomes

$$\det_{i,j=1}^{n} \left(\frac{c(z_i/w_j)}{a(z_i/w_j) b(z_i/w_j)} \right) = (-1)^{n(2n+3)}\ 3^{n/2}\ \frac{\prod_{i=1}^{n} \left(\frac{z_i w_i}{q}\right)^2}{\prod_{i,j=1}^{n} (w_j^6 - z_i^6)}\ \det_{i,j=1}^{2n} \left(Y_i^{\delta_j}\right)\ s_\lambda(Y_1, \ldots, Y_{2n}). \tag{5.71}$$

5.2.3 Computations

The numerator of the partition function Z_n (5.5) also has a factor involving the product of the weights a and b. Using (5.57) and (5.58), we find

$$\prod_{i,j=1}^{n} a(z_i/w_j) b(z_i/w_j) = \prod_{i,j=1}^{n} \frac{(w_j^6 - z_j^6)}{\left(\frac{z_i w_j}{q}\right)^2 \left(\left(\frac{z_i}{q}\right)^2 - w_j^2\right)}. \tag{5.72}$$

Next, we compute the Vandermonde determinant, which we decompose in three distinct "regions": $i, j \in \{1, \ldots, n\}$, $i, j \in \{n+1, \ldots, 2n\}$ and $i \in \{1, \ldots, n\}$, $j \in \{n+1, \ldots, 2n\}$.

$$\begin{aligned} \det_{i,j=1}^{2n} \left(x_i^{\delta_j}\right) &= \prod_{1 \le i < j \le 2n} (Y_i - Y_j) \\ &= \prod_{1 \le i < j \le n} \left(\left(\frac{z_i}{q}\right)^2 - \left(\frac{z_j}{q}\right)^2 \right)\ (w_i^2 - w_j^2)\ \prod_{i,j=1}^{n} \left(\left(\frac{z_i}{q}\right)^2 - w_j^2 \right). \end{aligned} \tag{5.73}$$

[19] The third Schur polynomial $s_\lambda(x_1, \ldots, x_6)$ for the double staircase partition already possesses 121 terms in its expansion. We only show the first five terms for lack of space!

Putting (5.71)—(5.73) together, we find for the partition function[20] Z_n (5.5)

$$Z_n = (-1)^{n(2n+3)}\, 3^{n/2}\, \frac{\prod_{i=1}^{n}\left(\frac{z_i w_i}{q}\right)^2}{\prod_{i,j=1}^{n}\left(\frac{z_i w_j}{q}\right)^2} \prod_{1\le i<j\le n} \frac{(w_i^2 - w_j^2)\left((z_i/q))^2 - (z_j/q)^2\right)}{\left[\frac{z_i}{z_j}\right]\left[\frac{w_j}{w_i}\right]} \times s_\lambda(Y_1,\ldots,Y_{2n}). \tag{5.74}$$

Let us simplify this result by reducing the first fraction to

$$\begin{aligned}\frac{\prod_{i=1}^{n}\left(\frac{z_i w_i}{q^2}\right)^2}{\prod_{i,j=1}^{n}\left(\frac{z_i w_j}{q}\right)^2} &= \frac{1}{\prod_{1\le i<j\le n}\left(\frac{z_i w_j}{q}\right)^2\left(\frac{z_j w_i}{q}\right)^2} \\ &= \frac{1}{\prod_{1\le i<j\le n}\left(\frac{z_i z_j w_i w_j}{q^2}\right)\ \prod_{i=1}^{n}\left(\frac{z_i w_i}{q}\right)^{n-1}}.\end{aligned} \tag{5.75}$$

Then, we inject the first product of the denominator in (5.75) into the appropriate product in (5.74). A helpful cancellation then occurs. On one hand we have for the numerator of this product

$$(w_i^2 - w_j^2)\left((z_i/q)^2 - (z_j/q)^2\right) = \left(\frac{z_j w_j}{q}\right)^2 - \left(\frac{z_i w_j}{q}\right)^2 - \left(\frac{z_j w_i}{q}\right)^2 + \left(\frac{z_i w_i}{q}\right)^2. \tag{5.76}$$

On the other hand, the denominator yields the same factors but each with an opposite sign

$$\left(\frac{z_i z_j w_i w_j}{q^2}\right)\left[\frac{z_i}{z_j}\right]\left[\frac{w_j}{w_i}\right] = \left(\frac{z_i w_j}{q}\right)^2 - \left(\frac{z_i w_i}{q}\right)^2 - \left(\frac{z_j w_j}{q}\right)^2 + \left(\frac{z_j w_i}{q}\right)^2. \tag{5.77}$$

The product over $1 \le i < j \le n$ thus reduces to[21] a simple factor of $(-1)^{n(n-1)/2}$ and we get

$$Z_n = (-1)^{n(3n+2)/2}\, 3^{n/2} \prod_{i=1}^{n}\left(\frac{q}{z_i w_i}\right)^{n-1} s_\lambda(Y_1,\ldots,Y_{2n}). \tag{5.78}$$

We make a remark about symmetry: the product above being diagonal and Schur polynomials being symmetric, we conclude that Z_n is in fact symmetric with respect to the whole set $\{z_1/q,\ldots,z_n/q,w_1,\ldots,w_n\}$ when $q = e^{i\pi/3}$. This surprising fact does not hold for generic q.

5.2.4 The homogeneous limit

Now that we have successfully written Z_n in terms of a Schur function, we intend to take the homogeneous limit where all the z's are equal to q and all the w's are equal to 1, just like in the first proof of Section 5.1. Recall that the act of taking this limit is what allows us to reconnect with ASMs via (3.5):

$$Z_n(q,\ldots,q;1,\ldots,1) = [q]^{n^2} A_n. \tag{5.79}$$

On the other hand, the homogeneous limit makes the product in (5.78) disappear and turns all the arguments of the Schur function into 1's so that

$$Z_n(q,\ldots,q;1,\ldots,1) = (-1)^{n(3n+2)/2}\, 3^{n/2}\, s_\lambda(1,\ldots,1). \tag{5.80}$$

[20] We do not explicitly write the dependence in $z_1,\ldots,z_n;w_1,\ldots,w_n$ in Z_n for simplicity.
[21] The exponent comes from the number of (-1) factors in this product.

Comparing (5.79) with (5.80) yields

$$A_n = \overbrace{(-1)^{n(3n+2)/2-n^2/2}}^{=(-1)^{n(n+1)}=1} \, 3^{n/2-n^2/2} \, s_\lambda(1,\dots,1) \tag{5.81}$$
$$= 3^{-n(n-1)/2} \, s_\lambda(1,\dots,1).$$

We have yet to calculate $s_\lambda(1,\dots,1)$ where λ is the double staircase partition. The next result, which is important in representation theory, helps us get there. The interested reader may find more information in [14].

Theorem 5.9. (Weyl's formula) *For a Schur polynomial s indexed by an integer partition $\lambda = (\lambda_1,\dots,\lambda_m)$ for $m \geq 1$, we have:*

$$s_\lambda(1,\dots,1) = \prod_{1\leq i<j\leq m} \frac{\lambda_i - i - (\lambda_j - j)}{i-j}. \tag{5.82}$$

Proof. Let us consider the following Schur function:

$$s_\lambda(1,t,t^2,\dots,t^m) = \frac{\det_{i,j=1}^m \left(t^{(j-1)(\lambda_i+m-i)}\right)}{\det_{i,j=1}^m \left(t^{(j-i)(m-i)}\right)}. \tag{5.83}$$

Since both of these determinants are Vandermonde's, it follows that

$$s_\lambda(1,t,t^2,\dots,t^m) = \prod_{1\leq i<j\leq m} \frac{t^{\lambda_i+m-i} - t^{\lambda_j+m-j}}{t^{j-1}-t^{i-1}}. \tag{5.84}$$

We now take the limit as t goes to 1 and conclude using L'Hospital's rule

$$\begin{aligned} s_\lambda(1,\dots,1) &= \lim_{t\to 1} s_\lambda(1,t,\dots,t^m) \\ &= \lim_{t\to 1} \prod_{1\leq i<j\leq m} \frac{(\lambda_i+m-i)\; t^{\lambda_i+m-i} - (\lambda_j+m-j)\; t^{\lambda_j+m-j}}{(j-1)\; t^{j-2} - (i-1)\; t^{i-2}} \\ &= \prod_{1\leq i<j\leq m} \frac{\lambda_i - i - (\lambda_j - j)}{i-j}. \end{aligned} \tag{5.85}$$

□

The final step consists in applying Weyl's formula (5.82) to the double staircase partition with $m = 2n$. Let us investigate the numerator and denominator separately.

Numerator:

We observe that the inside of the product in the numerator of Weyl's formula (with the double staircase partition) can take on four distinct forms depending on whether i and j are even or odd. Let us thus compute the term $\lambda_i - i - (\lambda_j - j)$ in each of these cases. Notice that the even entries of λ are $\lambda_{2p} = \lfloor n-p \rfloor = n-p$ and its odd entries are $\lambda_{2p+1} = \lfloor n-p-1/2 \rfloor = n-p-1$, for $p \in \mathbb{N}$. If we define $\mathcal{E} = \{1,\dots,n\}$ and $\mathcal{O} = \{0,\dots,n-1\}$, we have

$$\lambda_i - i - (\lambda_j - j) = \begin{cases} 3(l-k), & \text{if } i=2k \text{ and } j=2l, \quad k,l\in\mathcal{E}, \\ 3(l-k)+2, & \text{if } i=2k \text{ and } j=2l+1, \quad k\in\mathcal{E},\ l\in\mathcal{O}, \\ 3(l-k)-2, & \text{if } i=2k+1 \text{ and } j=2l, \quad k\in\mathcal{O},\ l\in\mathcal{E}, \\ 3(l-k), & \text{if } i=2k+1 \text{ and } j=2l+1, \quad k,l\in\mathcal{O}. \end{cases} \tag{5.86}$$

We are now in a position to compute the numerator of Weyl's formula which we call N from now on. We are dealing with a "triangular" product which we divide into two pieces: when i is even and when i is odd. Thus,

$$\begin{aligned} N = \prod_{1\le i<j\le 2n} \lambda_i - i - (\lambda_j - j) &= \prod_{k=1}^{n-1} \left(\prod_{l=k+1}^{n} 3(l-k) \prod_{l=k}^{n-1} \{3(l-k)+2\} \right) \\ &\times \prod_{k=0}^{n-2} \left(\prod_{l=k+1}^{n} \{3(l-k)-2\} \prod_{l=k+1}^{n-1} 3(l-k) \right). \end{aligned} \tag{5.87}$$

We extract the factors of 3 from the first and last terms and we let $p = l - k$ to find

$$N = 3^{n(n-1)} \prod_{k=1}^{n-1} \left(\prod_{p=1}^{n-k} p \prod_{p=0}^{n-1-k} (3p+2) \right) \prod_{k=0}^{n-2} \left(\prod_{p=1}^{n-k} (3p-2) \prod_{p=1}^{n-1-k} p \right). \tag{5.88}$$

The first and last terms are products over k of certain factorials which we pull out in the front. We group the two "middle" terms as products over $k = 1, \ldots, n-2$. Extracting the $k = n-1$ and $k = 0$ terms from their respective products yields a factor of 2 and a factor of $\prod_{p=1}^{n}(3p-2)$ such that

$$\begin{aligned} N = 3^{n(n-1)} & \prod_{k=1}^{n-1} (n-k)! \prod_{k=0}^{n-2} (n-1-k)! \times 2 \prod_{p=1}^{n} (3p-2) \\ & \times \prod_{k=1}^{n-2} \left(\prod_{p=0}^{n-1-k} (3p+2) \prod_{p=1}^{n-k} (3p-2) \right). \end{aligned} \tag{5.89}$$

By reindexing, one can see that the first two terms are both equal to $\prod_{p=1}^{n-1} p!$. We also reindex the third and fourth products by defining $r = p + 1$. We then find

$$N = 2\; 3^{n(n-1)} \prod_{p=1}^{n-1} (p!)^2 \prod_{r=0}^{n-1} (3r+1) \prod_{k=1}^{n-2} \prod_{r=1}^{n-k} \frac{(3r-2)(3r-1)(3r)}{(3r)}. \tag{5.90}$$

The reason we multiplied and divided by $3r$ is to recover a factorial of $3(n-k)$ in the numerator of the last term. We thus reindex one last time with $s = n - k$. Including the $s = 0$ and $s = 1$ terms yields a factor of $1/2$ which cancels out the 2 in front. In total, we have

$$\begin{aligned} N &= 3^{n(n-1)} \prod_{p=1}^{n-1} (p!)^{\cancel{2}} \prod_{r=0}^{n-1} (3r+1) \prod_{s=0}^{n-1} \frac{(3s)!}{3^p\, \cancel{p!}} \\ &= 3^{n(n-1)/2} \left(\prod_{r=0}^{n-1} r!\; (3r+1)! \right). \end{aligned} \tag{5.91}$$

Denominator:

The denominator D demands less work. In the following computation, we use two new indices: $p = j - i$ and $r = n - i$. We calculate that

$$\begin{aligned} D = \prod_{1 \le i < j \le 2n} (j-i) = \prod_{i=1}^{2n-1} \prod_{j=i+1}^{2n} (j-i) &= \prod_{i=1}^{2n-1} \prod_{p=1}^{2n-i} p \\ &= \prod_{i=1}^{2n-1} (2n-i)! \\ &= \prod_{r=0}^{1-n} (n+r)! \prod_{r=1}^{n-1} (n+r)! \\ &= \left(\prod_{r=1}^{n} r! \right) \left(\prod_{r=1}^{n-1} (n+r)! \right). \end{aligned} \tag{5.92}$$

We combine these two products to run over $r = 0$ to $r = n - 1$ (the factors $n!$ that arise cancel each other out). We thus find

$$D = \prod_{r=0}^{n-1} r!\, (n+r)! \tag{5.93}$$

Combining (5.91) and (5.93) we get the following result for λ the double staircase partition:

$$s_\lambda(1, \dots, 1) = 3^{n(n-1)/2} \prod_{r=0}^{n-1} \frac{(3r+1)!}{(n+r)!}. \tag{5.94}$$

The powers of 3 from Weyl's formula cancel with the ones in (5.81) which concludes our second proof of the ASM Conjecture 1.4.

Chapter 6

The refined ASM Conjecture

To conclude this thesis, we examine the refined enumeration of ASMs. Recall from the Introduction that ASMs only have a single +1 in their first row. We aim at calculating $A_{n,k}$, the number of $n \times n$ ASMs that have their +1 in the k^{th} column of the first row. In Conjecture 1.3, Robbins, Rumsey and Mills proposed the following formula:

$$\frac{A_{n,k}}{A_{n,1}} = \frac{(n+k-2)!\ (2n-k-1)!}{(2n-2)!\ (k-1)!\ (n-k)!}. \tag{6.1}$$

Doron Zeilberger, who first proved the ASM Conjecture, was also the first to prove this refined ASM Conjecture [35].

6.1 The combinatorial link between $A_{n,k}$ and Z_n

It is no surprise that we once more resort to the bijection between ASMs and the six-vertex model. We have to ask, however, what we need to apply to the parameters $\{z_1, \ldots, z_n, w_1, \ldots, w_n\}$ so as to make the connection with $A_{n,k}$. The answer for the "simple" enumeration was to take the so-called homogeneous limit where $z_i = q$ and $w_j = 1$ for all $i, j = 1, \ldots, n$. In the present case, we wish to fix the first row parameter z_1 and let all the other z's equal to q, keeping all the w's equal to 1. We refer to the following limit as the *refined limit*:

$$\begin{cases} z_1 = z,\ z_i = q, & \text{for } i = 2, \ldots, n, \\ w_j = 1 & \text{for } j = 1, \ldots, n. \end{cases} \tag{6.2}$$

To be clear, we are still working with $q = e^{i\pi/3}$. All the weights are then still equal to one another except in the first row of the lattice where $a(z) = [q^2/z]$, $b(z) = [z]$ and $c(z) = [q^2]$, with $z = z_1$.

We know that there is a single collision in the first row of an $n \times n$ six-vertex model state with DWBC.[22] We have represented the situation graphically in Figure 6.1. Also notice that we have chosen to label the z parameters from top to bottom which is allowed because of the symmetry of Z_n. The reason we have done this is to more easily compare the model configurations with ASMs and their top-row +1.

[22] One can view this as a consequence of the DWBC or equivalently as a consequence of the bijection with ASMs that only have a single +1 in their first row.

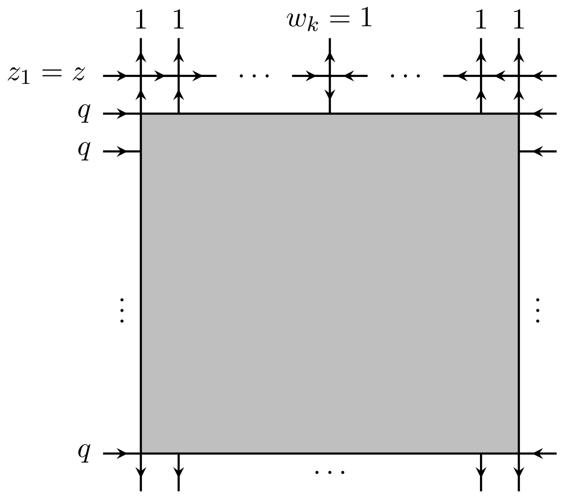


Figure 6.1: Six-vertex model state with DWBC in the refined limit with a collision in the k^{th} column of the first row.

The computation of the partition function Z_n in the refined limit, carried out below, exhibits the connection to the generating function of $A_{n,k}$. The "freezing" of the first row means it has $(k-1)$ vertices of type b, 1 of type c and $(n-k)$ of type a. Furthermore, in the remainder of the lattice, all the weights have the same value of $[q] = [q^2] = \sqrt{3}i$ and there are $n^2 - n$ of these vertices. These observations allow us to write the partition function as follows:

$$Z_n(z, q, \ldots, q; 1, \ldots, 1) = \sum_{k=1}^{n} [q^2/z]^{k-1} \, [q^2] \, [z]^{n-k} \sum_{\substack{\text{ASMs} \\ 1 \text{ in pos. } k}} [q]^{n^2-n}. \tag{6.3}$$

The first sum in (6.3) runs over the space of all possible configurations and is therefore indexed by k, the position of the top-row collision. The second sum counts the number of leftover configurations once the top row is fixed. By virtue of the bijection between six-vertex model configurations and ASMs and since $[q]$ is just a constant, this last sum simply counts the number of ASMs with a 1 in position k of the first row, or $A_{n,k}$. We now have

$$Z_n(z, q, \ldots, q; 1, \ldots, 1) = [q]^{n^2-n+1} \sum_{k=1}^{n} [z]^{k-1} \, [q^2/z]^{n-k} \, A_{n,k}. \tag{6.4}$$

Let us now introduce the generating function of $A_{n,k}$ which we call $A(t)$:

$$A(t) = \sum_{k=1}^{n} A_{n,k} \, t^{k-1}. \tag{6.5}$$

Relating the variable t to z with

$$t = \frac{[q^2/z]}{[z]}, \tag{6.6}$$

turns (6.4) into

$$Z_n(z, q, \ldots, q; 1, \ldots, 1) = [q]^{n^2-n+1} \, [z]^{n-1} \, A(t). \tag{6.7}$$

We recall (5.78) and (5.62) from Chapter 5 in which the partition function from the inhomogeneous six-vertex model is expressed in terms of a Schur function indexed by the double staircase partition as such (where q has already been set to $e^{i\pi/3}$ in the prefactors):

$$Z_n = (-1)^{n(3n+2)/2}\ 3^{n/2} \prod_{i=1}^{n} \left(\tfrac{q}{z_i w_i}\right)^{n-1}\ s_\lambda(Y_1,\ldots,Y_{2n}),$$
$$\text{with } Y_i = \begin{cases} \left(\tfrac{z_i}{q}\right)^2, & \text{for } i=1,\ldots,n, \\ w_{i-n}^2, & \text{for } i=n+1,\ldots,2n. \end{cases} \tag{6.8}$$

Taking the refined limit in this last equation turns all the Y variables into 1's except for the first one, hence the following relation:

$$Z_n(z,q,\ldots,q;1,\ldots,1) = (-1)^{n(3n+2)/2}\ 3^{n/2} \left(\frac{q}{z}\right)^{n-1}\ s_\lambda\left(\left(\tfrac{z}{q}\right)^2,1,\ldots,1\right). \tag{6.9}$$

Equating (6.7) with (6.9) and solving for $A(t)$, we find

$$A(t) = 3^{-n(n-1)/2} \left(\frac{q[q]}{z[z]}\right)^{n-1}\ s_\lambda\left(\left(\tfrac{z}{q}\right)^2,1,\ldots,1\right). \tag{6.10}$$

Before moving further, we need to express $\left(\frac{q[q]}{z[z]}\right)^{n-1}$ and $\left(\frac{z}{q}\right)^2$ as functions of t. It requires little algebra to show that we end up with the following expression for $A(t)$:

$$A(t) = 3^{-n(n-1)/2}\ (q^{-1}+qt)^{n-1}\ s_\lambda\left(\tfrac{q+q^{-1}t}{q^{-1}+qt},1,\ldots,1\right). \tag{6.11}$$

6.2 Finding a differential equation for $A(t)$

The end goal is to obtain the following differential equation for $A(t)$ which was first discovered (but given without proof) by Stroganov [30]:

$$t(1-t)\ A''(t) + 2(1-n-t)\ A'(t) + n(n-1)\ A(t) = 0. \tag{6.12}$$

However, we shall obtain Stroganov's differential equation using a more recent result by Gorin and Panova [15]. Their article is useful to us as it investigates the following normalized Schur function (for $\lambda = (\lambda_i)_{i=1}^N$ an integer partition of size N):

$$S_\lambda(x) = \frac{s_\lambda(x,1,\ldots,1)}{s_\lambda(1,\ldots,1)}. \tag{6.13}$$

Gorin and Panova showed (see Theorem 3.8 from [15]) that $S_\lambda(x)$ admits a contour integral representation. We reproduce their theorem below but do not provide a proof as it lies well beyond the scope of this work.

Theorem 6.1. *Let $\lambda = (\lambda_i)_{i=1}^N$ be an integer partition of size N and $\mu_i = \lambda_i + N - i$ for $i = 1,\ldots,N$. For any $x \in \mathbb{C}$ such that $x \neq 0,1$, the normalized Schur polynomial $S_\lambda(x)$ (6.13) may be written as*

$$S_\lambda(x) = \frac{(N-1)!}{(x-1)^{N-1}} \oint_{\mathscr{C}} \frac{d\xi}{2\pi i}\ \frac{x^\xi}{\prod_{i=1}^N(\xi-\mu_i)}, \tag{6.14}$$

where $\mathscr{C}$ is a simple contour which encloses all the poles μ_i of the integrand.

We now fix $N = 2n$ and choose $\lambda = (2n-1, 2n-1, \ldots, 1, 1, 0, 0)$ to be the double staircase partition. We also define a differential operator T.

Definition 6.2. *Let $f : \mathbb{C} \to \mathbb{C}$ be a function of class $\mathcal{C}^2(\mathbb{C})$. For any $x \in \mathbb{C}$, we define the operator T as*

$$(Tf)(x) = (1 - x^3)\ \frac{df}{dx}(x) + 6(n-1)\ x^2 \frac{df}{dx}(x) - 3(n-1)(3n-2)\ xf(x). \tag{6.15}$$

We prove a first intermediary result.

Lemma 6.3. *For $\lambda = (\lambda_i)_{i=1}^{2n}$ the double staircase partition and $\mu_i = \lambda_i + 2n - i$, we have*

$$T(x^\xi) = \xi(\xi-1)\ x^{\xi-2} - (\xi - \mu_1)(\xi - \mu_2)\ x^{\xi+1}. \tag{6.16}$$

Proof. First we observe that $6(n-1) = \mu_1 + \mu_2 - 1$ and $3(n-1)(3n-2) = \mu_1\mu_2$. We then compute

$$\begin{aligned} T(z^\xi) &= (1-x^3)\xi(\xi-1)\ x^{\xi-2} + (\mu_1+\mu_2-1)\xi\ x^{\xi+1} - \mu_1\mu_2\ x^{\xi+1} \\ &= \xi(\xi-1)\ x^{\xi-2} - \big(\xi(\xi-1) - (\mu_1+\mu_2-1)\xi + \mu_1\mu_2\big)\ x^{\xi+1} \\ &= \xi(\xi-1)\ x^{\xi-2} - (\xi-\mu_1)(\xi-\mu_2)\ x^{\xi+1}. \end{aligned} \tag{6.17}$$

□

Next, we want to show that the normalized Schur polynomial S_λ, where λ is the double staircase partition, satisfies a certain operator equation when T is applied to it.

Lemma 6.4. *For $x \in \mathbb{C}$ with $x \neq 0, 1$, we have*

$$T\Big((x-1)^{2n-1}\ S_\lambda(x)\Big) = 0. \tag{6.18}$$

Proof. The linearity of T as a differential operator along with Lemma 6.3 allow us to calculate

$$\begin{aligned} T\Big((x-1)^{2n-1}\ S_\lambda(x)\Big) &= \frac{(2n-1)!}{2\pi i} \oint_{\mathscr{C}} d\xi\ \frac{T(x^\xi)}{\prod_{i=1}^{2n}(\xi-\mu_i)} \\ &= \frac{(2n-1)!}{2\pi i} \oint_{\mathscr{C}} d\xi\ \left(\frac{\xi(\xi-1)\ x^{\xi-2}}{\prod_{i=1}^{2n}(\xi-\mu_i)} - \frac{(\xi-\mu_1)(\xi-\mu_2)\ x^{\xi+1}}{\prod_{i=1}^{2n}(\xi-\mu_i)}\right) \\ &= \frac{(2n-1)!}{2\pi i} \left[\oint_{\mathscr{C}'} d\xi\ \frac{x^{\xi-2}}{\prod_{i=1}^{2n-2}(\xi-\mu_i)} - \oint_{\mathscr{C}''} d\xi\ \frac{x^{\xi+1}}{\prod_{i=3}^{2n}(\xi-\mu_i)}\right]. \end{aligned} \tag{6.19}$$

Notice that we used $\mu_{2n} = 0$ and $\mu_{2n-1} = 1$ to get from the second to the third equation.[23] If we make the substitution $\xi' = \xi - 3$ in the first integral, the factors in the denominator become $(\xi' - (\mu_i - 3))$. Furthermore, for $i = 1, \ldots, 2n-2$, we have $\mu_{i+2} = \mu_i - 3$. The change of variables $\xi \to \xi'$ thus deforms the contour $\mathscr{C}'$ into $\mathscr{C}''$ such that we may write

$$T\Big((x-1)^{2n-1}\ S_\lambda(x)\Big) = \frac{(2n-1)!}{2\pi i} \left[\oint_{\mathscr{C}''} d\xi'\ \frac{x^{\xi'+1}}{\prod_{i=1}^{2n-2}(\xi'-\mu_{i+2})} - \oint_{\mathscr{C}''} d\xi\ \frac{x^{\xi+1}}{\prod_{i=3}^{2n}(\xi-\mu_i)}\right]. \tag{6.20}$$

[23] Also notice that the contours $\mathscr{C}'$ and $\mathscr{C}''$ now enclose the poles of their respective integrals, namely $\mu_1, \ldots, \mu_{2n-2}$ and $\mu_3, \ldots, \mu_{2n}$

Reindexing the product in the first integral with $i' = i+2$, we conclude that both integrals are exactly the same, yielding the desired result. □

The final step towards Stroganov's differential equation (6.12) is to use Lemma 6.4 and replace x with $\frac{q+q^{-1}t}{q^{-1}+qt}$, where $q = e^{i\pi/3}$. In this case, the function $(x-1)^{2n-1}S_\lambda(x)$ upon which the operator T acts, reads

$$\left(\frac{[q]\ (1-t)}{q^{-1}+qt}\right)^{2n-1} \ \frac{s_\lambda(q^{-1}+qt,1,\dots,1)}{s_\lambda(1,\dots,1)}. \tag{6.21}$$

Substituting relation (6.11) for $A(t)$, we find that (6.21) is given by the following function which we call h:

$$h(t) \equiv \frac{(1-t)^{2n-1}}{(q^{-1}+qt)^{3n-2}}\ A(t). \tag{6.22}$$

Note that we have omitted all the constants $[q]^{2n-1}$, $3^{n(n-1)}$ and $s_\lambda(1,\dots,1)$ as they play no role since T is linear and (6.18) is homogeneous. Below, we evaluate the changes in the first and second order derivatives in going from the variable x to the variable t. The former transforms as

$$\frac{d}{dx} = \frac{dt}{dx}\ \frac{d}{dt} = -\frac{(q^{-1}+qt)^2}{[q]}\ \frac{d}{dt}, \tag{6.23}$$

and the latter becomes

$$\frac{d^2}{dx^2} = \frac{dt}{dx}\ \frac{d^2t}{dtdx}\ \frac{d}{dt} + \left(\frac{dt}{dx}\right)^2 \frac{d^2}{dt^2} = \frac{2q}{[q]^2}\ (q^{-1}+qt)^3\ \frac{d}{dt} + \frac{(q^{-1}+qt)^4}{[q]^2}\ \frac{d}{dt^2}. \tag{6.24}$$

Substituting (6.22)—(6.24) modifies operator T, let us call its new version $\tilde{T}$. Furthermore, our specific value for q allows us to express the coefficient of the second order derivative in (6.15) as

$$1 - \left(\frac{q+q^{-1}t}{q^{-1}+qt}\right)^3 = \frac{3[q]t(t-1)}{(q^{-1}+qt)^3}. \tag{6.25}$$

Notation: Let us now use the prime (') notation for the derivatives with respect to t.

Our change of variables has modified the differential equation (6.18) into $\tilde{T}(h)(t) = 0$, which, after simplifications, explicitly reads

$$\begin{aligned} t(1-t)(q^{-1}+qt)\ h''(t) + 2\left(qt(1-t) + (n-1)(q+q^{-1}t)^2\right)\ h'(t) \\ + \frac{[q](n-1)(3n-2)(q+q^{-1}t)}{q^{-1}+qt}\ h(t) = 0. \end{aligned} \tag{6.26}$$

The computations of $h'(t)$ and $h''(t)$ result in

$$h'(t) = -\frac{(1-t)^{2n-2}}{(q^{-1}+qt)^{3n-2}}\left[(2n-1) + \frac{q(3n-2)(1-t)}{q^{-1}+qt}\right]\ A(t) + \frac{(1-t)^{2n-1}}{(q^{-1}+qt)^{3n-2}}\ A'(t), \tag{6.27}$$

and

$$h''(t) = \frac{(1-t)^{2n-3}}{(q^{-1}+qt)^{3n-2}} \left[(2n-2)(2n-1) + \frac{q(3n-2)(4n-3)(1-t)}{q^{-1}+qt} \right.$$
$$\left. + \left(\frac{q(3n-2)(1-t)}{q^{-1}+qt} \right)^2 + \frac{q(3n-2)(1-t)}{q^{-1}+qt} + \frac{q^2(3n-2)(1-t)^2}{(q^{-1}+qt)^2} \right] A(t)$$
$$- 2\frac{(1-t)^{2n-2}}{(q^{-1}+qt)^{3n-2}} \left[(2n-1) + \frac{q(3n-2)(1-t)}{q^{-1}+qt} \right] A'(t)$$
$$+ \frac{(1-t)^{2n-1}}{(q^{-1}+qt)^{3n-2}} A''(t). \tag{6.28}$$

We then plug in the values of $h'(t)$ and $h''(t)$ into the differential equation (6.26) to finally (tediously) find the desired result (6.12).

6.3 Solving the differential equation

Now that we have successfully derived Stroganov's differential equation, we seek a solution thereof. Though there are solutions to this equation in the form of infinite series[24], we know for certain that there must be a polynomial solution that corresponds to the previously defined (finite) series (6.5)

$$A(t) = \sum_{k=1}^{n} A_{n,k} t^{k-1}, \tag{6.29}$$

which corresponds to the generating function for the refined enumeration of ASMs. Plugging this "ansatz" into (6.12) we get

$$\sum_{k=1}^{n} (k-1)(k-2) A_{n,k}\, t^{k-2} - \sum_{k=1}^{n} (k-1)(k-2) A_{n,k}\, t^{k-1} + 2\sum_{k=1}^{n} (k-1) A_{n,k}\, t^{k-2}$$
$$- 2n \sum_{k=1}^{n} k(k-1) A_{n,k}\, t^{k-2} - 2\sum_{k=1}^{n} (k-1) A_{n,k}\, t^{k-1} + n(n-1) \sum_{k=1}^{n} A_{n,k}\, t^{k-1} = 0. \tag{6.30}$$

We reindex the first, third and fourth sums using $k' = k-1$. We are then able to combine all the sums into one. We observe that the coefficient of t^{n-1}, which comes from the second, fifth and last sums, vanishes. Furthermore, we omit the $k' = 0$ terms as they all happen to be zero. We thus have

$$\sum_{k=1}^{n-1} \left[k(k-1) A_{n,k+1} - (k-1)(k-2) A_{n,k} + 2k A_{n,k+1} - 2nk A_{n,k+1} \right.$$
$$\left. -2(k-1) A_{n,k} + n(n-1) A_{n,k} \right] t^{k-1} = 0. \tag{6.31}$$

The left-hand side only vanishes if the coefficient of t^{k-1} is zero for all $k = 1, \ldots, n-1$. After a few steps, one finds the following recursion:

$$A_{n,k} = \frac{(n-k+1)(n+k-1)}{(k-1)(2n-k)} A_{n,k-1}. \tag{6.32}$$

[24] One can show that Stroganov's equation may be written as the hypergeometric ODE. It thus admits solutions that involve hypergeometric functions.

Building on the recursion all the way down to $A_{n,1}$ and grouping like terms into factorials, one finds the desired result (6.1):

$$\frac{A_{n,k}}{A_{n,1}} = \frac{(n+k-2)!\,(2n-k-1)!}{(2n-2)!\,(k-1)!\,(n-k)!}. \tag{6.33}$$

This concludes the proof of the refined ASM Conjecture.

Chapter 7

Conclusion and Perspectives

In this thesis, we examined two conjectures regarding alternating sign matrices, both of which were stated in the 1980's and proven in the 1990's. First, we investigated the ASM Conjecture for which we gave two different proofs. The first followed Greg Kuperberg's work using Cauchy determinants and the second combined results from Okada, Gorin and Panova and Stroganov relating to Schur polynomials. Secondly, we studied the refined ASM Conjecture and gave a proof thereof that built upon our previous findings with respect to Schur polynomials.

All three proofs relied on one crucial observation: there is a one-to-one correspondence between ASMs and states of the six-vertex model on the square lattice with domain-wall boundary conditions. Once this bijection was established, it became clear that we needed to compute the partition function of the model. We have thus dedicated a whole chapter to convince ourselves, given the properties of this function, that the formula found by Izergin and Korepin was in fact the right one. It is at this stage that the Yang-Baxter equation for the six-vertex model was decisively used.

It is worth pointing out that many mathematicians and physicists got interested in the problem over the 14 years it took to prove the ASM Conjecture. The field of enumerative combinatorics significantly grew as a consequence and many ASM-related conjectures have since been stated and proven. We can mention the enumerations of ASMs that present any combination of symmetries of the square, a summary of which can be found in [20]. The case of totally symmetric ASMs remains unsolved as no formula has even been conjectured to this day. This is an incentive to dig deeper and could be an interesting topic for further investigations.
Just as we have studied refined ASMs with the top row fixed, one can also study other statistics on ASMs, for instance by freezing the first and last rows and columns or even the first two rows, etc. Studies like [5, 18] examine such problems while others have gone further in investigating refined enumerations of symmetry classes of ASMs [13].
Alternating sign matrices, along with other combinatorial objects such as descending plane partitions and monotone triangles continue to capture the attention of many scientists in the field where much work remains to be accomplished.

As for the six-vertex model itself, one might ask if there are any experimental ties with real-world implications. The answer is yes, in fact Algara-Siller *et al.* [2] managed to trap ice water between two sheets of graphene less than one millimeter apart and access its two-dimensional crystal structure exhibiting for the first time the existence of square ice.

Bibliography

[1] F. F. Abeles. "Dodgson condensation: The historical and mathematical development of an experimental method". *Linear Algebra Appl.* **429,** 429–438 (2008).

[2] G. Algara-Siller *et al.* "Square ice in graphene nanocapillaries". *Nature* **519,** 443–445 (2015).

[3] G. E. Andrews. "Plane partitions (III): The weak Macdonald conjecture". *Invent. Math.* **53,** 193–225 (1979).

[4] R. J. Baxter. "Partition Function of the Eight-Vertex Lattice Model". *Ann. Phys.* **70,** 193–228 (2000).

[5] R. E. Behrend, P. Di Francesco & P. Zinn-Justin. "A doubly-refined enumeration of alternating sign matrices and descending plane partitions". *J. Comb. Theory Ser. A* **120,** 409–432 (2013).

[6] D. Bressoud. "Proofs and Confirmations: The Story of the Alternating-Sign Matrix Conjecture" (Cambridge University Press, 1999).

[7] D. Bressoud & J. Propp. "How the Alternating Sign Matrix Conjecture Was Solved". *Not. Am. Math. Soc.* **46,** 10 (1999).

[8] A.-L. Cauchy. "Exercices d'analyse et de physique mathématique. Tome 2" (Bachelier, Paris, 1841).

[9] B. Diu, D. Lederer, B. Roulet & C. Guthmann. "Éléments de Physique Statistique" (Hermann, Paris, 2001).

[10] C. L. Dodgson. "Condensation of Determinants, Being a New and Brief Method for Computing their Arithmetical Values". *Proc. R. Soc. London* **15,** 150–155 (1886).

[11] N. Elkies, G. Kuperberg, M. Larsen & J. Propp. "Alternating-Sign Matrices and Domino Tilings (Part I)". *J. Algebr. Comb.* **1,** 111–132 (1992).

[12] I. Fischer & M. Konvalinka. "A bijective proof of the ASM theorem, Part II: ASM enumeration and ASM-DPP relation". *Int. Math. Res. Not.* **2022,** 7203–7230 (2022).

[13] I. Fischer & M. P. Saikia. "Refined enumeration of symmetry classes of alternating sign matrices". *J. Comb. Theory Ser. A* **178,** 105350 (2021).

[14] W. Fulton & J. Harris. "Representation theory: a first course". *Graduate Texts in Mathematics* **129** (Springer, New York, 2004).

[15] V. Gorin & G. Panova. "Asymptotics of symmetric polynomials with applications to statistical mechanics and representation theory". *Ann. Prob.* **43** (2015).

[16] E. Ising. "Beitrag zur Theorie des Ferromagnetismus". *Z. Phys.* **31,** 253–258 (1925).

[17] A. G. Izergin, D. A. Coker & V. E. Korepin. "Determinant formula for the six-vertex model". *J. Phys. A: Math. Gen.* **25,** 4315–4334 (1992).

[18] M. Karklinsky & D. Romik. "A formula for a doubly refined enumeration of alternating sign matrices". *Adv. Appl. Math.* **45,** 28–35 (2010).

[19] G. Kuperberg. "Another proof of the alternative-sign matrix conjecture". *Int. Math. Res. Not.* **1996,** 139–150 (1996).

[20] G. Kuperberg. "Symmetry Classes of Alternating-Sign Matrices under One Roof". *Ann. Math.* **156,** 835–866 (2002).

[21] W. Lenz. "Beitrag zum Verständnis der magnetischen Erscheinungen in festen Körpern". *Z. Phys.* **21,** 613–615 (1920).

[22] E. Lieb. "Exact Solution of the Problem of the Entropy of Two-Dimensional Ice". *Phys. Rev. Lett.* **18,** 692–694 (1967).

[23] W. H. Mills, D. P. Robbins & H. Rumsey. "Proof of the Macdonald conjecture". *Invent. Math.* **66,** 73–87 (1982).

[24] W. H. Mills, D. P. Robbins & H. Rumsey. "Alternating sign matrices and descending plane partitions". *J. Comb. Theory Ser. A* **34,** 340–359 (1983).

[25] F. Nichita. "Introduction to the Yang-Baxter Equation with Open Problems". *Axioms* **1,** 33–37 (2012).

[26] S. Okada. "Enumeration of symmetry classes of alternating sign matrices and characters of classical groups". *J. Algebr. Comb.* **23,** 43–69 (2006).

[27] L. Pauling. "The Structure and Entropy of Ice and of Other Crystals with Some Randomness of Atomic Arrangement". *J. Am. Chem. Soc.* **57,** 2680–2684 (1935).

[28] F. Reif. "Fundamentals of statistical and thermal physics" (Waveland Press, Long Grove, IL, 2009).

[29] D. P. Robbins & H. Rumsey. "Determinants and alternating sign matrices". *Adv. Math.* **62,** 169–184 (1986).

[30] Y. G. Stroganov. "Izergin-Korepin Determinant at a Third Root of Unity". *Theor. Math. Phys.* **146,** 53–62 (2006).

[31] T. S. Tavares, G. A. P. Ribeiro & V. E. Korepin. "Influence of boundary conditions on bulk properties of six-vertex model". *J. Phys. A: Math. Theor.* **48,** 454004 (2015).

[32] S. Vijay. "Studying the Six-Vertex Model with the Yang-Baxter Equation". *Department of Physics, MIT* (2009).

[33] C. N. Yang. "Some Exact Results for the Many-Body Problem in one Dimension with Repulsive Delta-Function Interaction". *Phys. Rev. Lett.* **19,** 1312–1315 (1967).

[34] D. Zeilberger. "Proof of the Alternating Sign Matrix Conjecture". *Electron. J. Comb.* **3** (1996).

[35] D. Zeilberger. "Proof of the Refined Alternating Sign Matrix Conjecture". *New York J. Math.* **2,** 59–68 (1996).