On Combinatorics, Integrability and Puzzles

by

Timothy Miller

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2020

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In the last decade, many old and new results in combinatorics have been shown using the theory of quantum integrable systems from particle physics. The key to solving such problems is the derivation of an underlying Yang-Baxter equation. In this thesis, we explore some of the results in this area, focusing on two proofs due to Zinn-Justin in [37]. The first is a proof of Knutson, Tao and Woodward's puzzle rule [18] which states that Littlewood-Richardson coefficients count the number of tilings of an equilateral triangle with three different types of tiles. The second result concerns Knutson and Tao's product rule for two factorial Schur functions [17]. We present an extension of Zinn-Justin's constructions to Grothendieck polynomials [35] and close with an overview of integrable vertex models. The purpose of this thesis is to make "combinatorics and integrability" more accessible to the general mathematician and illustrate the power and elegance of these ideas.

Acknowledgements

I first and foremost thank my supervisor Kevin Purbhoo for all his help and incredible guidance. Over the last two years, he has shown me some of the most satisfying and beautiful math. I am very happy to be continuing with a PhD under his supervision. I thank my parents for supporting me through this entire process. Going back to Carleton for undergrad, moving to Waterloo for a Masters, chasing a dream: they have been there every step of the way. I thank Karen Yeats and David Wagner for agreeing to read this thesis and their kind words about it. I was amazed that Karen was able to find some very subtle errors in the diagrams. I want to thank all the friends I made along the way who made the last two years some of the most memorable of my life. In order that may or may not be randomized: Soffia, James, Caelan, Tina, Martin, Josh, Andrew, Ronen, Jamie, Ali, Rose, Miriam, Kazu, Tao, Weston, Adina, Madeleine, Max, Matt, Sabrina, Lukas, Bethany, Kel, Christopher, Mariia, Nick, Naomi, Shannon and Ben.

Table of Contents

Chapter 1

Introduction

In this thesis, we explore combinatorial proofs inspired by quantum integrability. The idea is to use combinatorial objects as models of physical systems which have an "integrability condition" known as the Yang-Baxter equation. Applying the Yang-Baxter equation reduces the complexity of many proofs involving unwieldy combinatorial objects. Integrability has been applied to many problems involving polynomials from combinatorics, representation theory and algebraic geometry such as Schur polynomials, factorial or double Schur polynomials and Grothendieck polynomials. Many of these results are from the last decade and combinatorics and integrability appears to be an area of growing interest [6, 7, 11, 12, 28, 35, 33].

We focus mainly on puzzles. In 2003, Knutson, Tao and Woodward introduced puzzles (KTW-puzzles) as a tool to compute the Schubert calculus of the Grassmannian [18]. A KTW-puzzle is an equilateral triangle filled with three different tiles. Three Young diagrams (λ, μ, ν) may be read off the boundary and the number of puzzles (of a particular size) with that boundary is given by the Littlewood-Richardson coefficient $c_{\mu,\nu}^{\lambda}$. This fact is known as the puzzle rule. In the same year, Knutson and Tao showed a product rule for factorial Schur functions in [17] using MS-puzzles, named after Molev and Sagan who showed an earlier formulation in [27]. Zinn-Justin was able to reprove both of these results from integrability [37].

Proving the puzzle rule involves a simple two-dimensional model of a system of fermions. Particles are allowed to move to different energy levels in discrete time steps according to certain transfer matrices. Repeatedly applying a transfer matrix reveals an implicit triangle representing a KTW-puzzle (see Figure 4.4). Applying a different pair of transfer matrices returns supersymmetric Schur functions. The Yang-Baxter equation of Theorem 4.7.2

1

shows these transfer matrices commute and this allows us to exploit a supersymmetric Schur function identity to prove the result.

Constructions from the proof of the puzzle rule are then repurposed to address factorial Schur functions. A large puzzle represents the multiplication of two factorial Schur functions and applying the same Yang-Baxter equation as before transforms this puzzle into the desired expansion; this is succinctly summarized by the picture in Figure 5.4. As an original contribution, we use these constructions to show a neat identity for factorial Schur functions in Proposition 5.4.1.

We show the above proofs in full with some key differences. We take a more direct approach and only cover what is needed to prove the desired theorems whereas Zinn-Justin proves larger generalizations. For example, Zinn-Justin proves 81 identities for general transfer matrices; we cut down on the number of transfer matrices and try to streamline their definitions. We also minimize use of the shift operator. In [37], transfer matrices lay down rows of tiles and then shift them a half unit left or right to ensure elements read off of rows are in the same space. We avoid this where possible and instead alternate between two spaces $\mathcal{F}_0$ and $\mathcal{F}_1$ which correspond to even and odd-numbered rows. In the case of the product of factorial Schur functions, Zinn-Justin gives a second proof which we omit. We also define factorial Schur functions where the second set of variables is infinite as opposed to finite (a more standard definition) as we find it simplifies certain statements.

The structure of this thesis is as follows. In Chapter 2, we introduce basic terminology and Schur polynomials along with two generalizations: the supersymmetric and factorial Schur polynomials. In Chapter 3, we give an overview of the physical theory and language used in subsequent chapters. We explain our use of Dirac notation, fermions, the Yang-Baxter equation and Fock space. In Chapter 4, we present Zinn-Justin's proof of the puzzle rule in detail. We give a tiling model, define transfer matrices and prove the Yang-Baxter equation. In Chapter 5, we give Zinn-Justin's proof of the product rule for factorial Schur functions and prove our identity. Finally, in Chapter 6 we explore more recent research in integrability and combinatorics. Littlewood-Richardson coefficients for Grothendieck polynomials are found by modifying the previous constructions. We also discuss the relationship between puzzles and integrable vertex models, closing with an overview of some results in this area.

Chapter 2

Schur Polynomials and Generalizations

In this chapter, we introduce Schur polynomials and two of their generalizations. Schur polynomials are an important $\mathbb{Z}$-basis for the ring of symmetric polynomials. They have well known applications to representation theory where they are characters of irreducible polynomial representations of $\mathrm{GL}_n(\mathbb{C})$. They are also relevant to the representation theory of the symmetric group S_n; irreducible characters of S_n appear as coefficients in the expansion of Schur polynomials in terms of the power sum basis. Schur polynomials are also important in algebraic geometry where they represent Schubert classes.

2.1 Preliminaries

We denote an **integer partition** λ in multiple ways. We may say $\lambda = (\lambda_1, \ldots, \lambda_k)$ for positive integers $\lambda_1 \geq \cdots \geq \lambda_k$ and we may use exponential notation where $\lambda = n_1^{e_1} \cdots n_k^{e_k}$ for positive integers $e_1, \ldots, e_k$ and $n_1 < \cdots < n_k$. If λ is a partition of n, we say $\lambda \vdash n$ and say it has **size** $|\lambda| = n$. Each integer λ_i in $\lambda = (\lambda_1, \ldots, \lambda_k)$ is a **part** of λ and the number of parts in λ is its **length**, denoted $\ell(\lambda) = k$. The **width** of partition is the size of its largest part, λ_1. We say a partition μ is **contained** within λ and write $\mu \subseteq \lambda$ if $\mu_i \leq \lambda_i$ for all i. We allow an empty partition of size 0, denoted by $\varnothing$.

A **Young diagram** is a visual representation of a partition using rows of boxes to represent parts. We use the English convention where rows weakly decrease from top to bottom. We identify partitions and their Young diagrams, treating them interchangeably.

For concrete examples, we may draw the Young diagram inline: for example, $(3, 3, 1) = 13^2 = $ ⊞. If $\mu \subseteq \lambda$, the **skew diagram** λ/μ is obtained by removing the boxes of μ from the diagram of λ: for example, if $\lambda = $ ⊞ and $\mu = $ ⊞, then $\lambda/\mu = $ ⊞. A Young diagram with no boxes removed is said to have **straight shape**. A **horizontal strip** is a skew diagram with at most one box in each column. Similarly, a **vertical strip** is a skew diagram with at most one box in each row. In the previous example λ/μ is a horizontal strip, but not a vertical strip.

For a Young diagram λ, the **conjugate** diagram $\tilde{\lambda}$ is obtained by reflecting along the main diagonal: for example, if $\lambda = $ ⊞, then $\tilde{\lambda} = $ ⊞. The $\boldsymbol{n \times m}$ **rectangle** is the Young diagram with n rows of size m, which we denote by $\square = m^n$. If $\lambda \subseteq \square$, its **complement** within $\square$ is the diagram of $\square/\lambda$ rotated 180 degrees, denoted $\bar{\lambda}$: for example, if $\lambda = $ ⊞ and $\square = $ ⊞, then $\bar{\lambda} = $ ⊞.

A **semistandard Young tableau** is a Young diagram where each box is assigned a positive integer **entry**. The entries must weakly increase across rows from left to right and strictly increase down columns. We denote the set of all semistandard Young tableau of **shape** λ by $\text{SSYT}(\lambda)$.

Example 2.1.1. The following tableau is in $\text{SSYT}($⊞$)$:

$$
T =
\begin{array}{|c|c|c|c|}
\hline
1 & 2 & 2 & 4 \\
\hline
2 & 3 & 6 \\
\cline{1-3}
4 \\
\cline{1-1}
\end{array}
$$

We denote a box in a Young diagram or tableau with an ordered pair $\alpha = (i, j)$ where i is the row number and j is the column number. Rows and columns start at 1 read from top to bottom and left to right. If α is a box in a Young diagram λ and T is a Young tableau of shape λ, we say $\alpha \in \lambda$, $\alpha \in T$ and denote the entry of box α in T by $T(\alpha)$. In Example 2.1.1, $T(2, 3) = 6$. Let $\mathbf{x} = (x_1, \ldots, x_n)$. Each semistandard Young tableau is assigned a monomial

$$
\mathbf{x}^T = \prod_{\alpha \in T} x_{T(\alpha)}
$$

where $x_{T(\alpha)} = 0$ if $T(\alpha) > n$. In words, $\mathbf{x}^T = 0$ if the maximum entry in T is larger than n, otherwise it is the product of each variable x_i to the power of the number of times i appears as an entry in T. In Example 2.1.1, $\mathbf{x}^T = x_1 x_2^3 x_3 x_4^2 x_6$. The **Schur polynomial**

4

$s_\lambda(\mathbf{x})$ is the sum of all monomials $\mathbf{x}^T$ where $T \in \mathrm{SSYT}(\lambda)$. We write

$$s_\lambda(\mathbf{x}) = s_\lambda(x_1, \ldots, x_n) = \sum_{T \in \mathrm{SSYT}(\lambda)} \mathbf{x}^T.$$

This definition can be extended to infinitely many variables $\mathbf{x} = (x_1, x_2, \ldots)$ in which case $s_\lambda(\mathbf{x})$ is called the **Schur function**. We only consider Schur polynomials and their generalizations in finitely many variables, but all definitions and facts proved or stated in this thesis extend to the infinite variable case.

A polynomial $f(\mathbf{x})$ is is **symmetric** if $f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ for all $\sigma \in S_n$. The set of all such polynomials form a ring Λ_n, called the **ring of symmetric polynomials** in n independent variables. This is also denoted by $\mathbb{Z}[x_1, \ldots, x_n]^{S_n}$, meaning the set of all polynomials over $\mathbb{Z}$ fixed by S_n. The set of Schur polynomials in n variables is one of several known bases for Λ_n and hence the product of two Schur polynomials is expressible as a linear combination of Schur polynomials. The **Littlewood-Richardson coefficients** are nonnegative integers $c^\nu_{\lambda,\mu}$ depending only on λ, μ and ν such that

$$s_\mu(\mathbf{x})s_\nu(\mathbf{x}) = \sum_\nu c^\nu_{\lambda,\mu} s_\nu(\mathbf{x}).$$

Lastly, we give an algorithm to generate Young tableaux which will reappear in a different form in Section 4.5. Let $T \in \mathrm{SSYT}(\lambda)$ have maximum entry n. A box α with entry n in T cannot have boxes below it and no box in the same column as α can have entry n. If a box is to the right of α, it must contain entry n. Thus, all the boxes $\alpha \in T$ for which $T(\alpha) = n$ are contained in a horizontal strip $\lambda/\lambda^{(n)}$ for some $\lambda^{(n)} \subseteq \lambda$. With this observation, we can recursively compute all tableau in $\mathrm{SSYT}(\lambda)$ with maximum entry n. First, let $i = n$. In every possible way, fill boxes in the diagram λ with entry i so that there exists $\lambda^{(i)}$ such that $\lambda/\lambda^{(i)}$ is a horizontal strip and recursively apply this algorithm to $\lambda^{(i)}$, filling with entry $i - 1$. Terminate after filling with entry 1 and return all the tableau that were entirely filled. An example is given below.

Example 2.1.2. We compute all semistandard Young tableaux on $\lambda = \boxplus$ with maximum entry $n = 3$. Each row in step i represents all possible ways of filling a tableau from step $i + 1$ with entry i according to the algorithm. Step $i = 3$ initializes by filling a diagram that has no entries. The algorithm should return the highlighted tableaux.

$i = 3:$ (tableaux)

$i = 2:$ (tableaux)
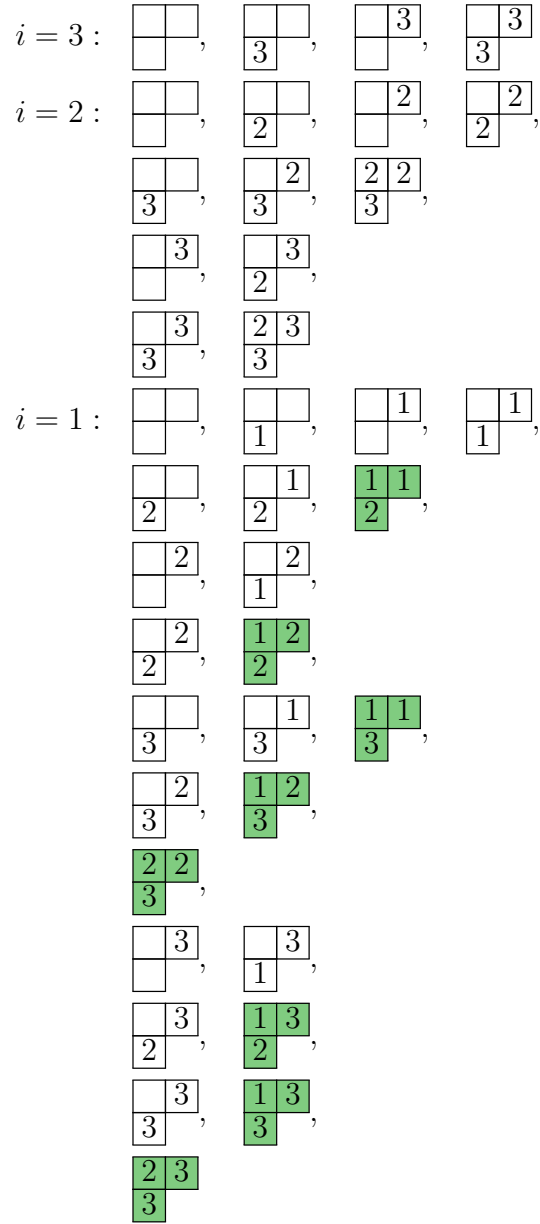
$i = 1:$ (tableaux)

Note that if we fill λ with vertical strips rather than horizontal strips, keeping the rest of the algorithm the same, it will generate tableaux which weakly increase down columns

and strictly increase across rows. Reflecting along the main diagonal, we see these are all semistandard Young tableau on the conjugate partition $\tilde{\lambda}$.

2.2   Supersymmetric Schur polynomials

Supersymmetric polynomials were first introduced by Metropolis, Nicoletti and Rota in [22] where they were called bisymmetric. Supersymmetric Schur polynomials are generalizations of Schur polynomials to two sets of variables. They were introduced by Berel and Regev in [1] where they called them Hook Schur functions due to their definition in terms of what are known as hook tableaux. Supersymmetric Schur functions have applications to Lie superalgebras which are $\mathbb{Z}_2$-graded Lie algebras satisfying certain Lie bracket axioms; a characterization can be found in [15]. In the early 1970s, physicists began to show interest in Lie superalgebras for their applications to supersymmetric gauge theory. A review of this history can be found in [10]. Supersymmetric Schur polynomials are irreducible characters of certain polynomial representations of the Lie superalgebra $\mathfrak{gl}(m/n)$, which is a generalization of the Lie group $\mathfrak{gl}(n)$. This is analogous to the relationship between Schur polynomials and $\mathfrak{gl}(n)$; more on this topic is in [1, 13, 31]. A thorough account of supersymmetric Schur polynomials can be found in a PhD thesis by Moens in [24].

A **doubly symmetric** polynomial is symmetric in two sets of variables. If $\mathbf{x} = (x_1, \ldots, x_m)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, then the doubly symmetric polynomial $f(\mathbf{x}/\mathbf{y}) = f(x_1, \ldots, x_m/y_1, \ldots, y_n)$ is in the ring $\mathbb{Z}[x_1, \ldots, x_m, y_1, \ldots, y_n]^{S_m \times S_n}$ which is to say it satisfies $f(\mathbf{x}/\mathbf{y}) = f(x_{\sigma(1)}, \ldots, x_{\sigma(m)}/y_{\tau(1)}, \ldots, y_{\tau(n)})$ for all permutations $\sigma \in S_m$ and $\tau \in S_n$. A polynomial $f(\mathbf{x}/\mathbf{y})$ satisfies the **cancellation property** if setting $x_m = t$ and $y_n = -t$ is the same as setting $x_m = y_n = 0$. The **ring of supersymmetric polynomials** on x and y, denoted $\Lambda_{m/n}$, is the set of doubly symmetric polynomials in $\mathbb{Z}[x_1, \ldots, x_m, y_1, \ldots, y_n]^{S_m \times S_n}$ that satisfy the cancellation property. As one might expect, supersymmetric Schur polynomials are a $\mathbb{Z}$-basis for $\Lambda_{m/n}$.

Example 2.2.1. The following function is in $\Lambda_{2,1}$:

$$f(x_1, x_2/y_1) = x_1^2 + x_1 x_2 + x_2^2 + x_1 y_1 + x_2 y_1$$

Note that setting $x_2 = t$ and $y_1 = -t$ yields x_1^2, which is the same as setting $x_2 = y_1 = 0$. This function is actually the supersymmetric Schur function $s_2(x_1, x_2/y_1)$.

A **supertableau**, also called a bitableau, is a filling of a Young diagram with two sets of positive integers; the integers in one set are marked with the prime symbol to

distinguish them. If the supertableau has shape λ, the non-primed integers are arranged in a semistandard Young tableau in a subdiagram $\mu \subseteq \lambda$ and the primed integers give a semistandard Young tableau on $\tilde{\lambda}/\tilde{\mu}$. We denote the set of all supertableau of shape λ by SSSYT(λ). A supertableau S is assigned a weight $(\mathbf{x}/\mathbf{y})^S$ where the $\mathbf{x}$ variables count the non-primed integers and the $\mathbf{y}$ variables count the primed integers. As before, $(\mathbf{x}/\mathbf{y})^S = 0$ if it has a non-primed entry above m or primed entry over n.

Example 2.2.2. The following supertableau has weight $(\mathbf{x}/\mathbf{y})^S = x_1 x_2^3 x_3 x_4 y_1^2 y_2^3 y_3$.

$$
S = \begin{array}{|c|c|c|c|c|}
\hline
1 & 2 & 2 & 1' & 2' \\
\hline
2 & 3 & 2' & 3' \\
\cline{1-4}
4 & 1' & 2' \\
\cline{1-3}
\end{array}
$$

The **supersymmetric Schur polynomial** $s_\lambda(\mathbf{x}/\mathbf{y})$ is defined as

$$
s_\lambda(\mathbf{x}/\mathbf{y}) = s_\lambda(x_1, \ldots, x_m/y_1, \ldots, y_n) = \sum_{S \in \text{SSSYT}(\lambda)} (\mathbf{x}/\mathbf{y})^S.
$$

These are clearly generalizations of Schur polynomials since, setting $y_1 = \cdots = y_n = 0$, the above formula reduces to a sum over semistandard Young tableau and we get the Schur polynomial $s_\lambda(\mathbf{x})$. There are many equivalent definitions of supersymmetric Schur polynomials; from [1, p.152], they may also be written as a sum of involving Littlewood-Richardson coefficients, which we state as a theorem.

Theorem 2.2.3. *The supersymmetric Schur polynomial $s_\lambda(\mathbf{x}/\mathbf{y})$ may be written as the following summation*

$$
s_\lambda(\mathbf{x}/\mathbf{y}) = \sum_{\mu,\nu} c_{\mu,\nu}^\lambda s_\nu(\mathbf{x}) s_{\tilde{\mu}}(\mathbf{y})
$$

where $c_{\mu,\nu}^\lambda$ are the Littlewood-Richardson coefficients. Moreover, these are the unique constants satisfying this equation.

We emphasize that the Littlewood-Richardson coefficients are the unique constants satisfying this equation, as this is key to the proof of the puzzle rule in Chapter 4. Lastly, note that all supertableaux of shape λ can be found with an algorithm like in Example 2.1.2 where, for each $\mu \subseteq \lambda$, we fill the boxes in λ/μ with vertical strips of primed integers and then fill the remaining boxes in μ with horizontal strips of regular integers.

$$T = \begin{array}{|c|c|c|c|} \hline 1 & 3 & 4 & 4 \\ \hline 2 & 6 \\ \cline{1-2} 4 \\ \cline{1-1} \end{array}$$

Figure 2.1: A semistandard Young tableau of weight $(x_1-y_1)(x_2-y_1)(x_3-y_4)(x_4-y_2)(x_4-y_6)(x_4-y_7)(x_6-y_6)$.

2.3 Factorial Schur polynomials

Factorial Schur polynomials are another generalization of Schur polynomials to two sets of variables. A less general form was first described by Biedenharn and Louck in [4] which was then generalized to the modern definition by Macdonald in [21]. Factorial Schur polynomials represent what are known as equivariant Schubert classes as shown in [23].

Let $\mathbf{x} = (x_1, \ldots, x_k)$ and $\mathbf{y} = (y_1, y_2, \ldots)$. For a box $\alpha = (i, j)$ in a Young diagram λ, the **content** of α is its column minus row number, denoted $c(\alpha) = j - i$. The **factorial Schur polynomial** is defined as follows:

$$s_\lambda(\mathbf{x}|\mathbf{y}) = s_\lambda(x_1, \ldots, x_k|y_1, y_2 \ldots) = \sum_{T \in \mathrm{SSYT}(\lambda)} \prod_{\alpha \in \lambda} (x_{T(\alpha)} - y_{T(\alpha)+c(\alpha)}).$$

where each tableau T has maximum entry k. An example of a tableau and its weight assigned in this summation is given in Figure 2.1. Note that despite $\mathbf{y}$ having infinite size, these are still polynomials since the tableaux in the summation have maximum entry k.

Factorial Schur polynomials are not doubly symmetric like supersymmetric Schur polynomials, but they are symmetric in the first set of variables. As before, setting $y_1 = y_2 = \cdots = 0$ reduces the above formula to the Schur polynomial $s_\lambda(\mathbf{x})$.

Remark 2.3.1. The factorial Schur polynomials get their name from the original less general case. As explained in [25], the modern polynomial has the following determinantal formula:

$$s_\lambda(\mathbf{x}|\mathbf{y}) = \frac{\det[(x_j|\mathbf{y})_{\lambda_i+k-i}]_{1\le i,j\le k}}{\prod_{i<j}(x_i - x_j)}$$

where $(z|\mathbf{y})_m = (z - y_1)\cdots(z - y_m)$. The original polynomial studied by Biedenharn and Louck in [4] had one set of variables and was denoted $t_\lambda(\mathbf{x})$. Specifically, $t_\lambda = s_\lambda(\mathbf{x}|\mathbf{y})$ where $y_1 = 0, y_2 = 1, \ldots, y_n = n - 1$, so in terms of the above formula, we have

$$t_\lambda(\mathbf{x}) = \frac{\det[(x_j)_{\lambda_i+k-i}]_{1\le i,j\le k}}{\prod_{i<j}(x_i - x_j)}$$

9

where $(z)_m$ is the *falling factorial* $z(z-1)\cdots(z-m+1)$.

Remark 2.3.2. Factorial Schur polynomials are closely related to the double Schur polynomials which have a similar definition. As before, we let $\mathbf{x} = (x_1, \ldots, x_k)$, but $\mathbf{y} = (y_i)_{i \in \mathbb{Z}}$ is a doubly infinite set of variables. A **reverse semistandard Young tableau** is a Young diagram filled with positive integers which decrease weakly across rows and strictly decrease down columns. Denoting the set of all reverse semistandard Young tableaux of shape λ as RSSYT(λ), the **double Schur polynomial** on λ is

$$s_\lambda(\mathbf{x}||\mathbf{y}) = \sum_{T \in \text{RSSYT}(\lambda)} \prod_{\alpha \in \lambda} (x_{T(\alpha)} - y_{T(\alpha)-c(\alpha)})$$

where the maximum entry in each tableau is k.

Double Schur polynomials can be obtained from factorial Schur polynomials by substituting the second set of variables with reversed and shifted indices. Specifically, if $z_i = y_{k-i+1}$ for $i \in \mathbb{Z}$, then $s_\lambda(\mathbf{x}||\mathbf{y}) = s_\lambda(\mathbf{x}|\mathbf{z})$ [26]. For this reason, some authors address double Schur polynomials, which have some geometric significance, through factorial Schur polynomials since the latter have a more natural combinatorial definition. Zinn-Justin takes this approach in the paper central to this thesis [37].

Let $\mathbf{x} = (x_1, \ldots, x_k)$, $\mathbf{y} = (y_1, y_2, \ldots)$ and $\mathbf{z} = (z_1, z_2, \ldots)$. The product of two factorial Schur polynomials with the same first set of variables $s_\lambda(\mathbf{x}|\mathbf{y})s_\mu(\mathbf{x}|\mathbf{z})$ can be written as a summation of factorial Schur polynomials on $\mathbf{x}$ and $\mathbf{y}$ with (polynomial) coefficients $c^\nu_{\lambda,\mu}(\mathbf{y};\mathbf{z}) \in \mathbb{Z}[y_1, y_2, \ldots, z_1, z_2, \ldots]$. These coefficients are computed with combinatorial objects known as MS-puzzles which we define in Section 5.2; we may expand to product as

$$s_\lambda(\mathbf{x}|\mathbf{z})s_\mu(\mathbf{x}|\mathbf{y}) = \sum_\nu c^\nu_{\lambda,\mu}(\mathbf{y};\mathbf{z})s_\nu(\mathbf{x}|\mathbf{y}).$$

This was first expressed with "unbarred" tableaux in [27] which was then rephrased by Knutson and Tao in terms of MS-puzzles [18]. This equation is restated as Theorem 5.3.1 in Chapter 5 where we ultimately present Zinn-Justin's proof from [37].

Chapter 3

Physics Background

In this chapter, we explain the physical language, notation and theory used in the proceeding chapters. We avoid delving into too much detail as we are mostly interested in the origins of the pictures in Chapter 4.

3.1 Dirac notation

Keeping with the physics literature, we use **Dirac notation**, also called **bra-ket notation**. Let $\mathcal{H}$ be a complex Hilbert space with a countable orthonormal basis $\mathcal{B}$ and an inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ defined by $\langle g, f \rangle = \delta_{f,g}$ for $f, g \in \mathcal{B}$. Let $f, g \in \mathcal{B}$ throughout. We define the "bra-ket" $\langle g|f \rangle = \langle g, f \rangle$. We index a "bra" $\langle \cdot |$ or "ket" $| \cdot \rangle$ with a basis element like $\langle g|$ or $|f\rangle$. We can have linear combinations of kets over $\mathbb{C}$ and a linear operator $\mathbf{A} : \mathcal{H} \to \mathcal{H}$ can act on an element $|f\rangle$ where $\mathbf{A}f = \sum_{h \in \mathcal{B}} a_h h$ if and only if $\mathbf{A}|f\rangle = \sum_{h \in \mathcal{B}} a_h |h\rangle$ for $a_h \in \mathbb{C}$. We think of "closing the bra-ket" with a bra $\langle g|$ as coefficient extraction, which is to say $\langle g|\mathbf{A}|f\rangle = a_g$ where $\mathbf{A}$ is as before. This perspective is all that is needed for our purposes.

3.2 Fermions

A **fermion** is a particle with a half-odd integer spin. Elementary particles such as quarks and electrons are fermions, but they may also be composite particles made of multiple elementary particles. Fermions satisfy the **Pauli exclusion principle**, meaning no two

can have the exact same state at the same time in a quantum system. The state of a fermion can consist of various numbers such as energy level, angular momentum and spin, but we will only concern ourselves with energy level and assume other values are equal.

In particle physics, solutions to the Dirac equation permit the existence of fermions at arbitrary negative energy levels. Particles tend to move down energy levels, so having infinite negative positions allows them to move down arbitrarily, which conflicts with the observed behaviour of fermions. A possible solution is to assume that there is a vacuum state in which every negative energy level is occupied by a fermion; this prevents indefinite descents since no two fermions can occupy the same state. This infinite set of fermions at negative energy levels is known as the **Dirac sea**.

We work with a simple two-dimensional fermionic model of energy level over time which assumes a Dirac sea. We represent basis states in the system pictorially as number lines with dots on every half-odd integer. A black dot at position $k \in \mathbb{Z} + \frac{1}{2}$ represents a particle with energy level k and white dots represent holes. The basis state with all negative energy levels occupied and no other particles is called the **vacuum state**, denoted $|\varnothing\rangle$:

$$|\varnothing\rangle = \ \cdots \ \bullet\!-\!\!\bullet\!-\!\!\bullet\!-\!\!\bullet\!-\!\!\bullet\ |\ \circ\!-\!\circ\!-\!\circ\!-\!\circ\!-\!\circ\ \cdots$$

Each basis state must have all sites to the left of some point occupied by particles and all sites to the right of some point occupied by holes; in the diagrams, all dots have the same colour beyond an ellipsis. To transition between states, it is conventional to define **creation and annihilation** operators $\psi_k^\star$ and ψ_k for each $k \in \mathbb{Z} + \frac{1}{2}$. Informally, one can think of these operators as attempting to insert or remove a particle from position k in the diagram of a basis state. We may associate every basis state to a composition of finitely many creation and annihilation operators with indices increasing left to right applied to the vacuum state as in the following example:

$$\psi_{-\frac{7}{2}}\psi_{-\frac{3}{2}}\psi_{\frac{1}{2}}^\star\psi_{\frac{5}{2}}^\star\psi_{\frac{7}{2}}^\star|\varnothing\rangle = \ \cdots \ \bullet\!-\!\circ\!-\!\bullet\!-\!\circ\!-\!\bullet\ |\ \bullet\!-\!\circ\!-\!\bullet\!-\!\bullet\!-\!\circ\ \cdots$$

The operators must also satisfy the following **anti-commutation relations**:

$$[\psi_i^\star, \psi_j]_+ = \delta_{i,j}, \qquad [\psi_i^\star, \psi_j^\star]_+ = [\psi_i, \psi_j]_+ = 0$$

where $[a, b]_+ = ab + ba$. One fact implied by the anti-commutation relations is that, if a basis state $|f\rangle$ has a particle with energy level k, then $\psi_k^\star|f\rangle = 0$; similarly, $\psi_k|f\rangle = 0$ if $|f\rangle$ has a hole at position k. Otherwise, a creation or annihilation operator applied to $|f\rangle$ returns a new basis state with a particle inserted or removed from position k, multiplied by a sign ± 1 determined by the anti-commutation relations.

Let $\mathcal{F}$ be the Hilbert space over $\mathbb{C}$ with the above basis state and inner product given by $\langle g|f \rangle = \delta_{g,f}$ for basis states $|f\rangle, |g\rangle \in \mathcal{F}$. To model changes over time, we evolve states of $\mathcal{F}$ in *discrete* time steps using transfer matrices. A **transfer matrix** $\mathbf{T} : \mathcal{F} \to \mathcal{F}$ is just an operator returning a weighted sum of basis states. They may also be given an argument x called a **spectral parameter** which affects the weights assigned to the basis states. That is, for a basis state $|f\rangle \in \mathcal{F}$,

$$\mathbf{T}(x)|f\rangle = \sum_{|g\rangle \in \mathcal{F}} w_g(x)|g\rangle$$

where $w_g(x)$ may, for instance, be a polynomial in x. In Chapter 4, we define the specific rules our transfer matrices use. A physical interpretation of the weights $w_g(x)$ is that they are the probability amplitude (essentially square root of probability) of state $|f\rangle$ collapsing into state $|g\rangle$ after a discrete time step.

Creation and annihilation operators are not needed beyond this point and we only consider the number lines going forward. In general, the operator defining the time evolution of a physical system is called the Hamiltonian. In the physics literature, the models we consider are called "free fermionic" because they have a Hamiltonian that is quadratic in the creation and annihilation operators. In [14], Schur functions are derived from a system with a state space like $\mathcal{F}$ using a Hamiltonian operator. A more modern approach, which our notation above is taken from is found in [36]. This is the inspiration for the proof in Chapter 4 which uses transfer matrices that decompose this Hamiltonian; specifically, chaining n transfer matrices together generates Schur polynomials in n variables.

3.3 Fock space

To model infinite identical particles, physicists use a construction called **Fock space**. Formally, if $\mathcal{H}$ is a complex Hilbert space representing the state space of a system with one particle in it, then the Fock space $\mathcal{F}(\mathcal{H})$ is the following completion:

$$\mathcal{F}(\mathcal{H}) = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes n} = \mathbb{C} \oplus \mathcal{H} \oplus (\mathcal{H} \otimes \mathcal{H}) \cdots.$$

In the literature, symmetric tensors are used to represent bosons and antisymmetric tensors are used to represent fermions. The symmetrization of the tensors makes the particles indistinguishable so as to simulate identical particles. This construction allows one to talk about infinite particle states, but we will only be concerned with particles in a finite region.
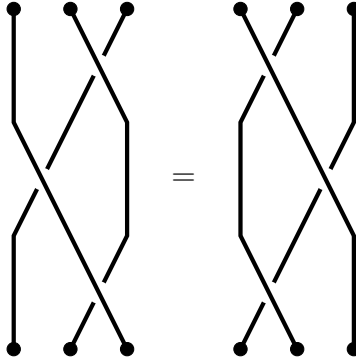
Figure 3.1: Illustration of the braid relation.

The number lines in the previous section can be thought of as elements in a Fock space $\mathcal{F}(\mathcal{H})$ where $\mathcal{H}$ is spanned by the two elements $\{\bullet, \circ\}$ where $\bullet$ is a fermion and $\circ$ is a hole. Basis states in this model can be thought of as "infinite" tensors $\bigotimes_{k \in \mathbb{Z} + \frac{1}{2}} x_k$ where $x_k \in \{\bullet, \circ\}$.

3.4 The Yang-Baxter equation and quantum integrability

The Yang-Baxter equation appears frequently in physics as a condition which allows a system to be solved; several examples from physics can be found in [16, 30]. Let V be a vector space. We say a matrix $\mathbf{R} : V \otimes V \to V \otimes V$ satisfies the **Yang-Baxter equation** if

$$(\mathbf{I} \otimes \mathbf{R})(\mathbf{R} \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{R}) = (\mathbf{R} \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{R})(\mathbf{R} \otimes \mathbf{I}).$$

The Yang-Baxter equation is illustrated by the braid relation as in Figure 3.1. After twisting two strings at a time as shown, the two "braids" can be deformed into each other without any twists and so they are equivalent.

An **R**-matrix can also be given a spectral parameter. In this case, the **parameter-dependent Yang-Baxter equations** are given by

$$(\mathbf{I} \otimes \mathbf{R}(x))(\mathbf{R}(z) \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{R}(y)) = (\mathbf{R}(y) \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{R}(z))(\mathbf{R}(x) \otimes \mathbf{I}).$$

14

The parameter z is often dependent on $x + y$ and for this reason it is common to write z as $x + y$ in the equation above (possibly redefining $\mathbf{R}$). This is the case in the Yang-Baxter equation of Theorem 4.7.2.

Lastly, the systems we consider are said to be **quantum integrable**. There is not a universally accepted definition of quantum integrability; see the discussion in [8, 9]. However, a sufficient condition for quantum integrability is that the system satisfies the Yang-Baxter equation. This is also the notion of quantum integrability that appears to be common in combinatorics research. We defer to the opinion of Karanth, Richmond and Schmidt in [16]: "we emphasize that the Yang-Baxter equation is an integrability condition. If it is satisfied, the model in question is integrable in the sense understood by all physicists."

Chapter 4

The Puzzle Rule

In this chapter, we present Zinn-Justin's proof [37] of Knutson, Tao and Woodward's puzzle rule [18]. This is the largest chapter where we introduce most of the constructions. We build a two-dimensional fermionic model of energy level over time. Tiles are decorated so that they trace out the paths fermions take as they jump to different energy levels in discrete time steps. We encode different types of behaviour with transfer matrices so that they give us Schur functions and supersymmetric Schur functions in Lemma 4.6.2. Applying the Yang-Baxter equation in Theorem 4.7.2 to two composed transfer matrices "unzips" them, showing they commute, which gives the proof.

4.1   Knutson-Tao-Woodward puzzles

A **Knutson-Tao-Woodward puzzle (KTW-puzzle)** is a filling of an equilateral triangle with the tiles in Figure 4.1. The triangle has integer side length n and each tile has unit side length. Tiles may be rotated, but not reflected and the edges of adjacent tiles must have matching labels. In the literature, 0 is often used in place of $-$ and 1 in place of $+$, but we use the notation of Zinn-Justin in [37].

We call a string of $+$ and $-$ signs a **±-string**. Every ±-string S has an associated Young diagram which we denote by $\lambda(S)$. To produce $\lambda(S)$, simply draw a lattice path P from the origin in $\mathbb{Z}^2$ taking one step North for every $-$ and one step East for every $+$ in the order given by S. If P contains k North steps, enclose it within $\square = (n-k)^k$. Then the top left region enclosed by P and $\square$ is the desired Young diagram in the English convention. An example is shown in Figure 4.2.

Figure 4.1: The tiles and a valid KTW-puzzle.

Figure 4.2: Computing $\lambda(--++-+-++) = \boxplus$.

Given a KTW-puzzle, the labels along each side are read off (each from left to right) to give a $\pm$-string. Let the **side strings** (B, L, R) be the triple of strings read off the bottom, left and right sides respectively. The triple $(\lambda(B), \lambda(L), \lambda(R))$ is called the **boundary**.

We say the side strings are **homogeneous** if they all have the same number of $+$ and $-$ signs. We can see the following lemma in the upcoming proof of the puzzle rule, but give a short, elegant argument from [18] here.

Lemma 4.1.1. *The side strings of a KTW-puzzle must be homogeneous.*

Proof. For each tile, draw a unit vector pointing outwards perpendicular to each edge containing a $+$ as follows:

The sum of the vectors around each tile is zero and hence the sum of all the vectors in a KTW-puzzle must be zero. If two adjacent tiles meet at edges labelled with a $+$, the vectors attached to those edges point in opposite directions and sum to zero, so all internal

17

vectors cancel. The only remaining vectors on the sides of the puzzle must also sum to zero, which is only possible if each side string has the same number of $+$ signs. $\qquad\square$

We can now state the **puzzle rule** for computing Littlewood-Richardson coefficients, originally proved by Knutson, Tao and Woodward.

Theorem 4.1.2 (Knutson, Tao, Woodward [18]). *The number of KTW-puzzles of side length n with homogeneous side strings and boundary (λ, μ, ν) is the Littlewood-Richardson coefficient $c^{\lambda}_{\mu,\nu}$.*

Example 4.1.3. Taking the filling in Figure 4.1, we have the boundary $(\lambda, \mu, \nu) = (\boxplus, \boxminus, \square)$. The Littlewood-Richardson coefficient $c^{\lambda}_{\mu,\nu}$ is equal to 1, so this is the only KTW-puzzle with this boundary and side length 4:
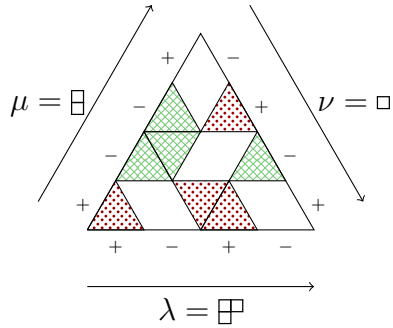
4.2 Tiling the plane

We now turn Zinn-Justin's proof of the puzzle rule from [37]. We introduce the new tiles in Figure 4.3 and note that the α and β tiles give all the rotations of the original tiles in Figure 4.1 if the β tiles are joined along the dotted lines. We address the γ tiles later. We say the tile is of **upper type** if it is in the top row and **lower type** if it is in the bottom row. The two new symbols 0 and $\tilde{0}$ ensure that β and γ tiles form a rhombus with one possible matching tile along a dotted edge. Throughout, we hide the dashed edges that meet between adjacent tiles.

In Figure 4.4, we show how KTW-puzzles emerge. The picture is a filling of the upper half-plane using only the α and β tiles with no β_0 sitting on the horizontal axis. We also assume that all coloured lines to the left of some integer N_- are green and all coloured lines to the right of some integer N_+ are red. The green lines can only move left and the
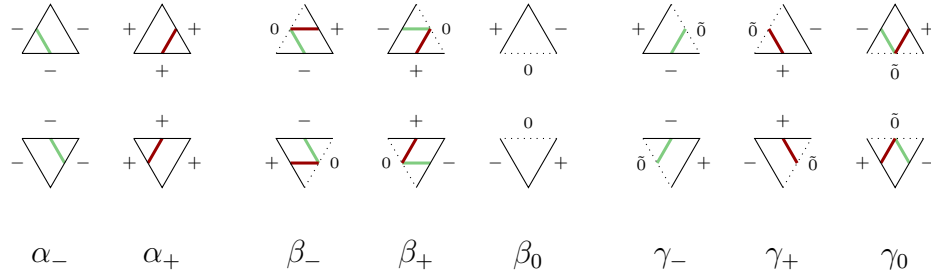
Figure 4.3: The new tiles.
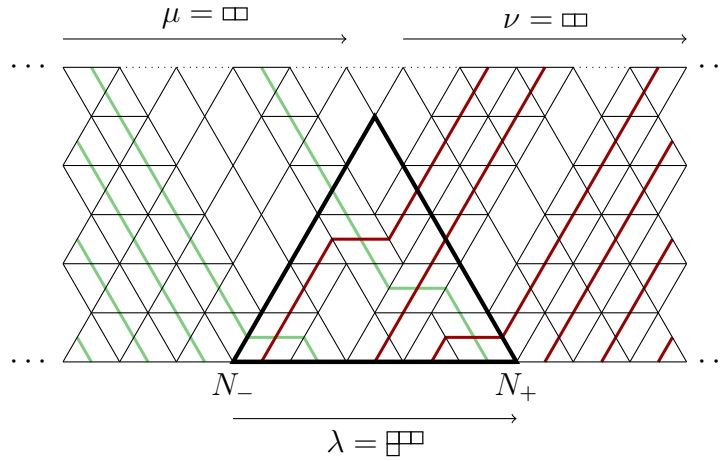
Figure 4.4: Tiling the upper-half plane with the α and β tiles.

red lines can only move right, so each green line will eventually be to the left of each red line. After all crossings take place, there is only one way to continue tiling upward and coloured lines follow a fixed trajectory. The equilateral triangle with the interval $[N_-, N_+]$ as its base contains a KTW-puzzle. The Young diagrams μ and ν can be read horizontally off the top where green lines correspond to North moves and red lines correspond to East moves. Our goal is to show that the number of such tilings of the upper half-plane where (λ, μ, ν) emerge as shown is the Littlewood-Richardson coefficient $c_{\mu,\nu}^{\lambda}$.

19
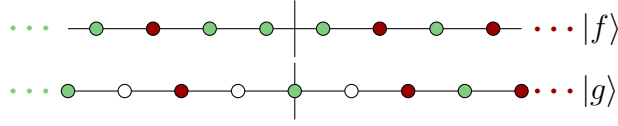
Figure 4.5: Basis elements $|f\rangle \in \mathcal{F}_0$ and $|g\rangle \in \mathcal{G}_1$. Here, $N_-(f) = -3$ and $N_+(f) = 3$.

4.3 The Fock spaces $\mathcal{F}$ and $\mathcal{G}$

We consider horizontal cross-sections below every row of the tiled upper-half plane and draw them as number lines with red and green dots on integers or half-odd integers. These are basis states of a Fock space similar to that of Section 3.2, though we have two types of particles (green and red). We assume discrete time and think of the cross-sections as different points in time. The decorations of the tiles trace out the paths particles take as they jump to different energy levels. This model is fermionic since no two particles can occupy the same site at the same time. To be precise, we describe these states with functions.

Let $\mathcal{G}_0$ and $\mathcal{G}_1$ be complex Hilbert spaces spanned by basis elements $|f\rangle$ indexed by functions $f : \mathbb{Z} + \frac{1}{2} \to \{-1, 0, 1\}$ in the case of $\mathcal{G}_0$ and $f : \mathbb{Z} \to \{-1, 0, 1\}$ in the case of $\mathcal{G}_1$. For these functions, we also require that there exist integers N_- and N_+ where $f(i) = -1$ for $i < N_-$ and $f(i) = 1$ for $i > N_+$. Define $N_-(f)$ and $N_+(f)$ as the largest and smallest such integers respectively. Define $\mathcal{G} = \mathcal{G}_0 \oplus \mathcal{G}_1$ and let $\mathcal{F} \subset \mathcal{G}$ be the Hilbert subspace with basis elements $|f\rangle$ where $f(i) \neq 0$ for all i in the domain of f. Similarly, let $\mathcal{F} = \mathcal{F}_0 \oplus \mathcal{F}_1$ where $\mathcal{F}_0 \subset \mathcal{G}_0$ and $\mathcal{F}_1 \subset \mathcal{G}_1$.

We define the inner product on $\mathcal{G}$ as the bra-ket $\langle f|g\rangle = \delta_{f,g}$ for basis elements $|f\rangle, |g\rangle \in \mathcal{G}$. We define a vacuum state in $\mathcal{G}_0$ as $|\varnothing\rangle$ where $\varnothing : \mathbb{Z} + \frac{1}{2} \to \{-1, 1\}$ is simply the sign function restricted to $\mathbb{Z} + \frac{1}{2}$. We do not define a vacuum state for $\mathcal{G}_1$.

As mentioned, we will represent basis elements in $\mathcal{G}$ with number lines as in Figure 4.5 where a green, white or red dot at position i indicates that $f(i)$ is -1, 0 or 1 respectively. For a basis element $|f\rangle \in \mathcal{G}$, diagrams are assumed to contain the interval $[N_-(f), N_+(f)]$, so that there is no ambiguity. We will refer to these pictures rather than their underlying functions when possible.

We assign a **charge** $c(f)$ to a basis element $|f\rangle \in \mathcal{G}$ where green dots contribute -1, red dots contribute 1 and white dots contribute 0. Precisely, the formula for charge is
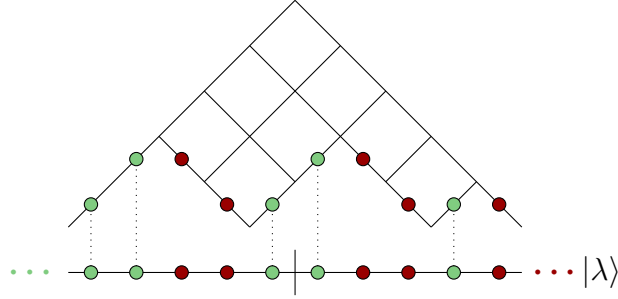
$$c(f) = \sum_i (f(i) - \operatorname{sign}(i))$$

20

Figure 4.6: A zero charge basis element $|\lambda\rangle \in \mathcal{F}_0$ and its associated Young diagram λ.

where the sum is over all i in the domain of f. The **emptiness number** $e(f)$ is the number of white dots in $|f\rangle$ or $e(f) = \#\{i \mid f(i) = 0\}$.

Note that there is a bijection between zero charge basis elements in $\mathcal{F}_0$ and Young diagrams. If $|f\rangle \in \mathcal{F}_0$ is a zero charge basis element, simply take the $\pm$-string given by $|f\rangle$ within the interval $[N_-(f), N_+(f)]$. That is, if $a = N_-(f) + \frac{1}{2}$, $b = N_+(f) - \frac{1}{2}$ and S is the $\pm$-string $\text{sign}(f(a))\,\text{sign}(f(a+1))\cdots\text{sign}(f(b))$, then the associated Young diagram is $\lambda(S)$. The main diagonal of the Young diagram will always be centred above 0 in the number line when drawn as in Figure 4.6. If a zero charge basis element $|f\rangle \in \mathcal{F}$ is associated to a Young diagram λ, we denote $|f\rangle$ by $|\lambda\rangle$ and also define $N_-(\lambda) = N_-(f)$ and $N_+(\lambda) = N_+(f)$.

We define the **shift operator**, denoted **S**, which shifts all the dots in $|f\rangle \in \mathcal{F}$ one step to the right. That is, if $f'(i+1) = f(i)$, then $|f'\rangle = \mathbf{S}|f\rangle$. We also define the **combine** operator $\sqcup : \mathcal{F}_0 \otimes \mathcal{F}_0 \to \mathcal{G}_0$; if $|f\rangle = |g\rangle \sqcup |h\rangle$ for basis elements $|g\rangle, |h\rangle \in \mathcal{F}_0$, then

$$f(i) = \begin{cases} \frac{1}{2}(g(i) - 1) & \text{if } i < 0 \\ \frac{1}{2}(h(i) + 1) & \text{if } i > 0. \end{cases}$$

In words, $|g\rangle \sqcup |h\rangle$ turns all red dots in $|g\rangle$ and green dots in $|h\rangle$ into white dots and then concatenates everything to the left of 0 in $|g\rangle$ with everything to the right of 0 in $|h\rangle$. Extend $\sqcup$ linearly and, for $k \geq 0$, define $\sqcup_k$ so that $|g\rangle \sqcup_k |h\rangle = (\mathbf{S}^{-k}|g\rangle) \sqcup (\mathbf{S}^k|h\rangle)$. For notational convenience, we allow the combine operator to act on the indexing functions within kets so that

$$|g\rangle \sqcup_k |h\rangle = |g \sqcup_k h\rangle.$$

Lastly, note that $\sqcup_k$ is injective if we restrict its domain to the subset of $\mathcal{F}_0 \times \mathcal{F}_0$ consisting of pairs of basis elements $|g\rangle, |h\rangle$ where $k \geq \max(N_+(g), -N_-(h))$; this is only circumstance where we use $\sqcup_k$.
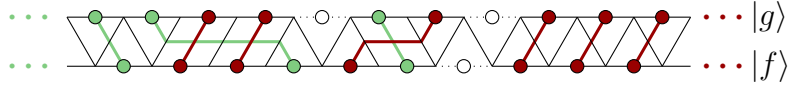
Figure 4.7: The element $|g\rangle$ is obtainable from $|f\rangle$ following $\mathbf{T}$.

Example 4.3.1. We apply $\sqcup_3$ to two basis elements in $\mathcal{F}_0$.
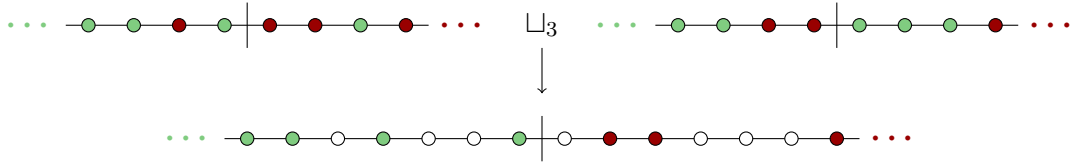
4.4 The transfer matrix T

The elements in $\mathcal{G}$ capture the state space of some toy physical system. We model the evolution of this system over time with transfer matrices that represent a discrete time step. Throughout, we define transfer matrices for basis elements in $\mathcal{G}$ and extend linearly.

We define $\mathbf{T} : \mathcal{G} \to \mathcal{G}$, which evolves elements using the α and β tiles in Figure 4.3. For a basis element $|f\rangle \in \mathcal{G}$, place a row of α and β tiles on the diagram of $|f\rangle$ so that coloured lines in the tiles match the dots on the diagram. The edges between tiles must also match, of course. Treating the top of this row as a new diagram yields an element $|g\rangle \in \mathcal{G}$ and we say $|g\rangle$ is **obtainable** from $|f\rangle$ **following** $\mathbf{T}$. An example is given in Figure 4.7. It is only possible to construct elements obtainable from $|f\rangle$ in one way; thus, $\mathbf{T}|f\rangle$ is the sum of all basis elements obtainable from $|f\rangle$ and we have

$$\langle g|\mathbf{T}|f\rangle = \begin{cases} 1 & \text{if } |g\rangle \text{ is obtainable from } |f\rangle \text{ following } \mathbf{T} \\ 0 & \text{otherwise.} \end{cases}$$

Note that a row of tiles must move coloured lines a half-odd integer to the left or right, so $\mathbf{T}$ sends elements in $\mathcal{G}_0$ to $\mathcal{G}_1$ and elements in $\mathcal{G}_1$ to $\mathcal{G}_0$. Following the paths that lines take after repeated applications of $\mathbf{T}$ produces images as in Figure 4.4, reading from the bottom to top. Note that the rules force green lines to move at least a half-step left and red lines to move at least a half-step right after each row of tiles and so eventually all the green lines will be to the left of the red lines; we formalize this with a lemma.

Lemma 4.4.1. *For a basis element* $|f\rangle \in \mathcal{F}_0$, *let* $p = \max(N_+(f), -N_-(f))$. *Then there exist unique nonnegative integer coefficients* $\tilde{c}_{g,h}^f$ *such that for all* $k \geq p$

$$\mathbf{T}^{2k}|f\rangle = \sum_{g,h} \tilde{c}_{g,h}^f |g \sqcup_k h\rangle$$

where the summation is over all basis elements $|g\rangle, |h\rangle \in \mathcal{F}_0$ *with* $p \geq \max(N_+(g), -N_-(h))$.

Proof. If $\langle d|\mathbf{T}^{2p}|f\rangle \neq 0$ for a basis element $|d\rangle$, then $|d\rangle \in \mathcal{G}_0$ and all its green dots are to the left of 0, all red dots are to the right of 0 and thus $|d\rangle = |g \sqcup_q h\rangle$ for some basis elements $|g\rangle, |h\rangle \in \mathcal{F}_0$. If we restrict to elements where $p \geq \max(N_+(g), -N_-(h))$, then there is a unique pair (g, h) where $|d\rangle = |g \sqcup_p h\rangle$. For $k \geq p$, we have

$$\mathbf{T}^{2k}|f\rangle = \mathbf{T}^{2(k-p)} \sum_{g,h} \tilde{c}_{g,h}^f |g \sqcup_p h\rangle$$

$$= \sum_{g,h} \tilde{c}_{g,h}^f |g \sqcup_k h\rangle$$

where the summation is restricted as in the lemma. $\qquad\square$

Next, in the above lemma, if $|f\rangle = |\lambda\rangle \in \mathcal{F}_0$ is a zero charge basis element, we can write $\mathbf{T}^{2k}|\lambda\rangle$ as a sum of other zero charge basis elements.

Lemma 4.4.2. *Let* $p = \max(N_+(\lambda), -N_-(\lambda))$ *for a zero charge basis element* $|\lambda\rangle \in \mathcal{F}_0$. *Then there exist unique nonnegative integer coefficients* $\tilde{c}_{\mu,\nu}^\lambda$ *such that for all* $k \geq p$

$$\mathbf{T}^{2k}|\lambda\rangle = \sum_{\mu,\nu} \tilde{c}_{\mu,\nu}^\lambda |\mu \sqcup_k \nu\rangle$$

where the summation is over all zero charge basis elements $|\mu\rangle, |\nu\rangle \in \mathcal{F}_0$ *with* $p \geq \max(N_+(\mu), -N_-(\nu))$.

Proof. Consider a basis element $|d\rangle \in \mathcal{G}_0$ where $\langle d|\mathbf{T}^{2k}|\lambda\rangle \neq 0$. From Lemma 4.4.1, we get that $|d\rangle = |g \sqcup_k h\rangle$ where $|g\rangle$ and $|h\rangle$ are basis elements in $\mathcal{F}_0$. We must prove $|g\rangle$ and $|h\rangle$ have zero charge. Note that $\mathbf{T}$ pushes the coloured dots outside the interval $[N_-(f), N_+(f)]$ outwards half a step, clearing space for one white dot without adding or destroying any coloured dots. Thus, $\mathbf{T}$ preserves charge, but increases the emptiness number by 1 and therefore, $c(d) = 0$ and $e(d) = 2k$.

Since $|d\rangle = |g \sqcup_k h\rangle$ has zero charge and $2k$ white dots, it must have k white dots on either side of 0. The white dots to the left of 0 correspond to k red dots in $|g\rangle$ to the left of k. Since $N_+(g) \leq k$, all dots to the right of k in $|g\rangle$ are red. Since $|g\rangle$ has no white dots, if it has ℓ green dots in the interval $[0, k]$, it must then have ℓ red dots pushed to the interval $(-\infty, 0]$ and thus $|g\rangle$ has zero charge (it has the same number of red dots to the left of 0 as green dots to the right of 0). Similarly for $|h\rangle$. $\qquad\square$

Theorem 4.4.3. *If each coefficient $\tilde{c}^\lambda_{\mu,\nu}$ in Lemma 4.4.2 is equal to the Littlewood-Richardson coefficient $c^\lambda_{\mu,\nu}$, the puzzle rule is true.*

Proof. For a zero charge basis element $|\lambda\rangle \in \mathcal{F}_0$, let $p = \max(N_+(\lambda), N_-(\lambda))$. Then Lemma 4.4.2 says there are $\tilde{c}^\lambda_{\mu,\nu}$ different ways of tiling the plane with the α and β tiles starting from $|\lambda\rangle$ and ending with $|\mu \sqcup_p \nu\rangle$ for zero charge basis elements $|\mu\rangle, |\nu\rangle$. As illustrated in Figure 4.4, the lines encoding the Young diagrams μ and ν also exit the left and right sides of the triangle, following a fixed trajectory. These Young diagrams are read left to right in the same way as in the puzzle rule. Thus, $\tilde{c}^\lambda_{\mu,\nu}$ count puzzles with boundary (λ, μ, ν) and the puzzle rule is true if $\tilde{c}^\lambda_{\mu,\nu} = c^\lambda_{\mu,\nu}$. $\qquad\square$

4.5 The transfer matrices $\mathbf{T}_\pm(x)$ and $\tilde{\mathbf{T}}_\pm(x)$

Since we want Littlewood-Richardson coefficients to come out of this picture, it is natural to expect Schur polynomials emerge as well. To this end, we define two new transfer matrices, $\mathbf{T}_\pm(x) : \mathcal{G} \to \mathcal{G}$. Both matrices have their own rules for which elements are obtainable from a basis element $|f\rangle \in \mathcal{G}$:

<table>
<tr><td align="center">$\mathbf{T}_-(x)$ rules</td><td align="center">$\mathbf{T}_+(x)$ rules</td></tr>
<tr><td>

- Place a row of α, β and γ_- tiles on the diagram of $|f\rangle$ so that edges match.
- At the boundaries, all tiles must move all lines one half-step to the **right**. That is, there are only γ_- tiles sufficiently to the left and α_+ tiles sufficiently to the right.

</td><td>

- Place a row of α, β and γ_+ tiles on the diagram of $|f\rangle$ so that edges match.
- At the boundaries, all tiles must move all lines one half-step to the **left**. That is, there are only α_- tiles sufficiently to the left and γ_+ tiles sufficiently to the right.

</td></tr>
</table>

Figure 4.8: The element $|g\rangle$ is obtainable from $|f\rangle$ following $\mathbf{T}_+(x)$.

In Figure 4.8, we show a possible evolution for $\mathbf{T}_+(x)$. As before, the output of these matrices applied to a basis element $|f\rangle \in \mathcal{G}$ is a sum over elements obtainable from $|f\rangle$ following the respective rules, but the terms are weighted by a spectral parameter x. Note that α_+ and β_+ tiles come in pairs when following $\mathbf{T}_+(x)$. For α_+ tiles, there will possibly be a sequence of β_- tiles between each pair. Similarly, α_- and β_- tiles come in pairs when following $\mathbf{T}_-(x)$. Each element obtainable from $|f\rangle$ is weighted by the number of these pairs:

<table>
<tr><td align="center">T_−(x) weights</td><td align="center">T_+(x) weights</td></tr>
</table>

- Let $w_-(f,g)$ be the number of pairs of $\boldsymbol{\alpha_-}$ and $\boldsymbol{\beta_-}$ tiles used in the evolution from $|f\rangle$ to $|g\rangle$ following $\mathbf{T}_-(x)$.
- If $|g\rangle$ is obtainable from $|f\rangle$ following $\mathbf{T}_-(x)$, then $\langle g|\mathbf{T}_-(x)|f\rangle = x^{w_-(f,g)}$.

- Let $w_+(f,g)$ be the number of pairs of $\boldsymbol{\alpha_+}$ and $\boldsymbol{\beta_+}$ tiles used in the evolution from $|f\rangle$ to $|g\rangle$ following $\mathbf{T}_+(x)$.
- If $|g\rangle$ is obtainable from $|f\rangle$ following $\mathbf{T}_+(x)$, then $\langle g|\mathbf{T}_+(x)|f\rangle = x^{w_+(f,g)}$.

Example 4.5.1. We apply $\mathbf{T}_+(x)$ to an element in $\mathcal{G}_0$.

We slightly modify these matrices to get Schur polynomials. Abusing notation, define

$$\tilde{\mathbf{T}}_+(x) = \mathbf{S}^{1/2}\mathbf{T}_+(x)$$
$$\tilde{\mathbf{T}}_-(x) = \mathbf{S}^{-1/2}\mathbf{T}_-(x).$$

That is, $\tilde{\mathbf{T}}_+(x)$ applies $\mathbf{T}_+(x)$ and shifts all the dots one half-step to the right and $\tilde{\mathbf{T}}_-(x)$ applies $\mathbf{T}_-(x)$ and shifts all the dots one half-step to the left. The reason we make this

25

Figure 4.9: An evolution of a zero charge basis element $|\lambda\rangle \in \mathcal{F}_0$ following $\mathbf{T}_+(x)$ and its corresponding evolution following $\tilde{\mathbf{T}}_+(x)$.

change is to view $\tilde{\mathbf{T}}_+(x)$ as a map from $\mathcal{F}_0$ to $\mathcal{F}_0$. We compare $\mathbf{T}_+(x)$ and $\tilde{\mathbf{T}}_+(x)$ in Figure 4.9.

Consider placing a row of tiles following $\tilde{\mathbf{T}}_+(x)$. Since the boundaries must only consist of α_- and γ_+ tiles, the half step to the right makes the boundary lines go straight up. The remaining dots within the boundaries are just permuted and thus $\tilde{\mathbf{T}}_+(x)$ preserves charge and emptiness number. If we restrict the domain of $\tilde{\mathbf{T}}_+(x)$ to $\mathcal{F}$, then α_+ and β_- tiles cannot be used in the evolution. As mentioned, α_+ tiles come in pairs and the leftmost α_+ of upper type must be to the right of a β_0 of lower type. This is not possible since $|f\rangle$ has no white dots and $\tilde{\mathbf{T}}_+(x)$ preserves emptiness number. There is no β_- since a sequence of β_- tiles must end with an α_+ tile. Therefore, the evolution can only use α_-, β_+ and γ_+ tiles. Similarly, $\mathbf{T}_-(x)$ only uses α_+, β_- and γ_- tiles when restricted to $\mathcal{F}$. Using these facts, we describe $\tilde{\mathbf{T}}_\pm(x)|_\mathcal{F}$ in words:

<div style="display:flex">

$\tilde{\mathbf{T}}_-(x)|_\mathcal{F}$ rules

- **Green** dots may stay in their current position or cross any number of consecutive **red** dots immediately to their **left**.
- **Red** dots may stay in their current position or move one step to the **right**.
- No two dots may occupy the same position at the same time.

$\tilde{\mathbf{T}}_+(x)|_\mathcal{F}$ rules

- **Red** dots may stay in their current position or cross any number of consecutive **green** dots immediately to their **right**.
- **Green** dots may stay in their current position or move one step to the **left**.
- No two dots may occupy the same position at the same time.

</div>

The number of **crossings** in the evolution from $|f\rangle$ to $|g\rangle$ is the number of times two

26

Figure 4.10: An element $|\lambda'\rangle$ obtainable from λ following $\tilde{\mathbf{T}}_+(x)$. The evolution removes the horizontal strip given by the grey boxes, leaving the Young diagram λ'.

lines cross in our visualizations; in Figure 4.9, there are three crossings. When applying $\tilde{\mathbf{T}}_\pm(x)$ to a basis element $|f\rangle \in \mathcal{F}_0$, the number of crossings is counted by $w_\pm(f, g)$ since the weighted $\alpha_\pm$ tiles are not used in $\mathcal{F}_0$.

4.6 Deriving Schur polynomials

For a zero charge basis element $|\lambda\rangle \in \mathcal{F}_0$, note what following $\tilde{\mathbf{T}}_+(x)$ is doing to the underlying Young diagram λ: it removes a horizontal strip. Green dots cross some number of consecutive red dots to the left, which corresponds to removing a strip as in Figure 4.10. If there are c crossings in the evolution of $|\lambda\rangle$ to $|\lambda'\rangle$ following $\tilde{\mathbf{T}}_+(x)$ then there are λ' has c fewer boxes and $\langle\lambda'|\tilde{\mathbf{T}}_+(x)|\lambda\rangle = x^c$. If we apply $\tilde{\mathbf{T}}_+(x_i)$ for $1 \leq i \leq n$ in descending order, this is analogous to running the algorithm given in Section 2.1, only keeping track of entries used in tableaux with monomials multiplied by Young diagrams that represent the unfilled portion of the diagram. We give an example which essentially runs the same computation as in Example 2.1.1.

Example 4.6.1. Computing $s_{21}(x_1, x_2, x_3)$:

$$\langle\varnothing|\prod_{i=1}^{3}\tilde{\mathbf{T}}_{+}(x_i)|⊞\rangle = \langle\varnothing|\prod_{i=1}^{2}\tilde{\mathbf{T}}_{+}(x_i)(|⊞\rangle + x_3|⊞\rangle + x_3|⊟\rangle + x_3^2|\square\rangle)$$

$$= \langle\varnothing|\tilde{\mathbf{T}}_{+}(x_1)(|⊞\rangle + x_2|⊞\rangle + x_2|⊟\rangle + x_2^2|\square\rangle$$
$$+ x_3|⊞\rangle + x_2 x_3|\square\rangle + x_2^2 x_3|\varnothing\rangle$$
$$+ x_3|⊟\rangle + x_2 x_3|\square\rangle$$
$$+ x_3^2|\square\rangle + x_2 x_3^2|\varnothing\rangle)$$

$$= \langle\varnothing|(|⊞\rangle + x_1|⊞\rangle + x_1|⊟\rangle + x_1^2|\square\rangle$$
$$+ x_2|⊞\rangle + x_1 x_2|\square\rangle + x_1^2 x_2|\varnothing\rangle$$
$$+ x_2|⊟\rangle + x_1 x_2|\square\rangle$$
$$+ x_2^2|\square\rangle + x_1 x_2^2|\varnothing\rangle$$
$$+ x_3|⊞\rangle + x_1 x_3|\square\rangle + x_1^2 x_3|\varnothing\rangle$$
$$+ x_2 x_3|\square\rangle + x_1 x_2 x_3|\varnothing\rangle$$
$$+ x_2^2 x_3|\varnothing\rangle$$
$$+ x_3|⊟\rangle + x_1 x_3|\square\rangle$$
$$+ x_2 x_3|\square\rangle + x_1 x_2 x_3|\varnothing\rangle$$
$$+ x_3^2|\square\rangle + x_1 x_3^2|\varnothing\rangle$$
$$+ x_2 x_3^2|\varnothing\rangle)$$

$$= x_2^2 x_3 + x_1^2 x_2 + x_1 x_2^2 + 2 x_1 x_2 x_3 + x_1 x_3^2 + x_2 x_3^2$$
$$= s_{21}(x_1, x_2, x_3).$$

Similarly, $\tilde{\mathbf{T}}_{-}(x)$ removes vertical strips so we get the Schur polynomial for the transpose of the underlying Young diagram. The supersymmetric Schur polynomials are obtained by removing vertical strips and then horizontal. We have the following lemma.

Lemma 4.6.2. *Let* $\mathbf{x} = (x_1, \ldots, x_n)$ *and* $\mathbf{y} = (y_1, \ldots, y_n)$. *Then*

$$s_\lambda(\mathbf{x}) = \langle\varnothing| \prod_{i=1}^{n} \tilde{\mathbf{T}}_+(x_i)|\lambda\rangle,$$

$$s_{\tilde\lambda}(\mathbf{x}) = \langle\varnothing| \prod_{i=1}^{n} \tilde{\mathbf{T}}_-(x_i)|\lambda\rangle,$$

$$s_\lambda(\mathbf{x}/\mathbf{y}) = \langle\varnothing| \prod_{i=1}^{n} \mathbf{T}_+(x_i) \prod_{i=1}^{n} \mathbf{T}_-(y_i)|\lambda\rangle.$$

Proof. The result follows from the previous remarks. We can use $\mathbf{T}_\pm(x)$ in place of $\tilde{\mathbf{T}}_\pm(x)$ in the third equation since the shift operators commute with the transfer matrices and cancel out. $\qquad\square$

We make a note before one more important lemma. If we restrict $\tilde{\mathbf{T}}_-(x)$ to elements with only green and white dots, it acts just as $\tilde{\mathbf{T}}_-(x)|_{\mathcal{F}}$ only where white dots behave as red. This is because $\tilde{\mathbf{T}}_-(x)$ also counts α_- tiles and, restricted to white and green dots, it only uses α_-, β_0 and γ_- tiles. As stated before, $\tilde{\mathbf{T}}_-(x)|_{\mathcal{F}}$ only uses α_+, β_- and γ_- tiles. Adding some dotted lines that represent the paths of white dots, we see the bijection:
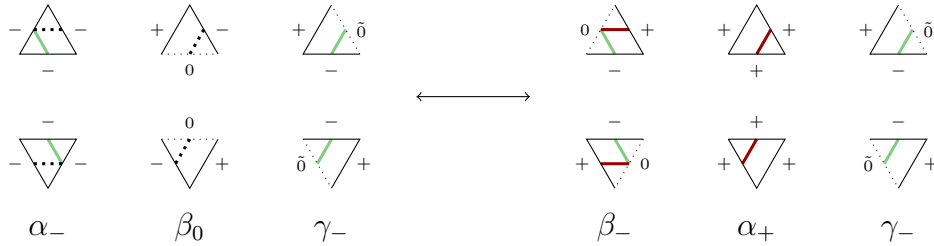
Similarly, restricting $\tilde{\mathbf{T}}_+(x)$ to red and white dots acts just as $\tilde{\mathbf{T}}_+(x)|_{\mathcal{F}}$ only where white dots behave as green.

Lemma 4.6.3. *For two zero charge basis elements* $|\mu\rangle, |\nu\rangle \in \mathcal{F}_0$, *let* $k \geq \max(N_+(\mu), -N_-(\nu))$. *Then*

$$\langle\varnothing \sqcup_k \varnothing| \prod_{i=1}^{n} \mathbf{T}_+(x_i) \prod_{i=1}^{n} \mathbf{T}_-(y_i)|\mu \sqcup_k \nu\rangle = s_\nu(\mathbf{x})s_{\tilde\mu}(\mathbf{y}).$$

Proof. From the commutativity of the shift operator, convert the matrices to their tilde versions. Applying $\prod_{i=1}^{n} \tilde{\mathbf{T}}_-(y_i)$ to $|\mu \sqcup_k \nu\rangle$ makes the red lines coming from $|\nu\rangle$ on the right all go straight up. From the above remarks, applying $\prod_{i=1}^{n} \tilde{\mathbf{T}}_+(x_i)$ should evolve the red lines so that they contribute $s_\nu(\mathbf{x})$. On the left, the green lines from $|\mu\rangle$ are evolved independently by $\prod_{i=1}^{n} \tilde{\mathbf{T}}_-(y_i)$ and then go straight up after $\prod_{i=1}^{n} \tilde{\mathbf{T}}_+(x_i)$, so this evolution contributes $s_{\tilde{\mu}}(\mathbf{y})$. $\qquad \square$

4.7   The Yang-Baxter equation

Consider the rhombi in Figure 4.11. The rhombi in each column may use the α and β tiles in any combination but we only allow one type of γ tile in each orientation. The rhombi that use the α and β tiles are assigned weight 1 and the γ tiles are assigned weight x. Rhombi of orientation $\diagup\!\!\!\!\diagdown$ may contain γ_- tiles, orientation $\diagdown\!\!\!\!\diagup$ may contain γ_+ tiles and orientation $\diamondsuit$ may contain γ_0 tiles. Rhombi in each row are rotations of each other (after redecorating with lines). Notice that rhombi in orientation $\diagdown\!\!\!\!\diagup$ use the same tiles as $\mathbf{T}_+(x)$ and $\mathbf{T}_-(x)$ uses tiles from orientation $\diagup\!\!\!\!\diagdown$. The transfer matrices we consider do not use the γ_0 tiles, but we include these to prove a more general Yang-Baxter equation for future use.

We consider all combinations of rhombi in Figure 4.11 arranged into hexagons where the labels of the exterior edges are fixed. It is convenient to depict the sum of all such rhombi with thick-edged rhombi arranged into hexagons. A spectral parameter is put inside each rhombus to indicate the weight given to the γ tile. When multiple shapes are put together, the internal edges are summed over while the outer edges are fixed. If it is not possible to combine the rhombi to get the exterior edges, the hexagon has weight 0. We illustrate with a few examples.

Example 4.7.1. There is only one possible hexagon in most cases. There is only one rhombus in orientation $\diagup\!\!\!\!\diagdown$ which has a $+$ on its bottom edges and so we have

In the next case, the weight is zero since we reach a point where no rhombus of orientation

30
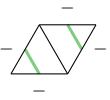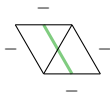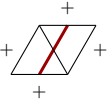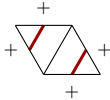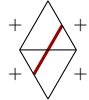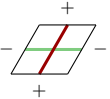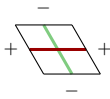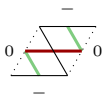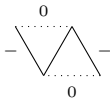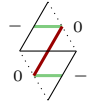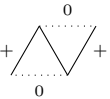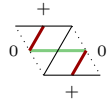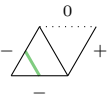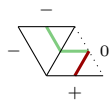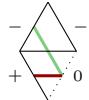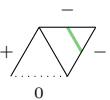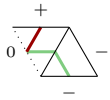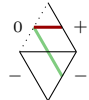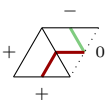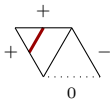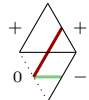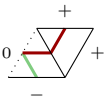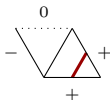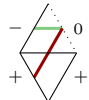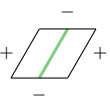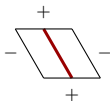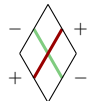
Figure 4.11: Weights assigned to all possible rhombi.

$\diagup\!\!\!\!\diagup$ will fit in the bottom right:

$$\diagup\!\!\!\boxed{\begin{smallmatrix} x \\ y \end{smallmatrix}} = \diagup\!\!\!\boxed{\begin{smallmatrix} x \\ y \end{smallmatrix}} = \diagup\!\!\!\boxed{\begin{smallmatrix} \\ y \end{smallmatrix}} = 0.$$

Lastly, one can compute:

$$\diagup\!\!\!\boxed{\begin{smallmatrix} x \\ y \end{smallmatrix}} = \diagup\!\!\!\boxed{} + \diagup\!\!\!\boxed{} + \diagup\!\!\!\boxed{} = x + y + z.$$

In general, we have the following theorem, which we later show is the Yang-Baxter equation.

Theorem 4.7.2. *If $x+y+z = 0$, then for every sequence of external edges $(a, b, c, d, e, f) \in \{-, +, 0\}^6$, we have*

$$\begin{smallmatrix} & e & \\ f & \boxed{\begin{smallmatrix} x \\ z \quad y \end{smallmatrix}} & d \\ a & & c \\ & b & \end{smallmatrix} = \begin{smallmatrix} & e & \\ f & \boxed{\begin{smallmatrix} y \\ z \\ x \end{smallmatrix}} & d \\ a & & c \\ & b & \end{smallmatrix}$$

Proof. If both sides of the equation only involve α and β tiles, then equality holds since each orientation allows α and β tiles in every combination; both sides of the equation have weight 1 in this case.

Also note that, if the equation holds for some set of external edges, it holds for any cyclic permutation of those edges. All rotations of each rhombus (ignoring decorations) are contained in Figure 4.11 and the three rhombi with nontrivial weight are rotations of each other. Thus, a 120 degree rotation of the edges amounts to permuting the weights. For instance, the filling in Example 4.7.1 has weight x; the filling with the same outer edges rotated 120 degrees counterclockwise has weight z. Also, rotating the edge labels 180 degrees amounts to swapping the left and right hand side of the equation since the tiles with nontrivial weight are invariant under 180 degree rotation. All rotations can then be obtained from 120 and 180 degree rotations.

The only hexagons left to check are those with external edges that allow γ tiles. Up to cyclic rotation, the external edge sequences can only be $(-, +, -, -, +, -)$, $(-, +, +, -, +, +)$ and $(+, -, +, -, +, -)$. The first two sequences are invariant under 180 degree rotations and thus equality holds from our previous remarks. For the last sequence, as shown in the third computation of Example 4.7.1, the left hand side has value $x + y + z$. However, there is no hexagon with the same external edges when computing the right hand side; therefore, if $x + y + z = 0$, the theorem holds. $\qquad\square$

Theorem 4.7.2 gives a Yang-Baxter equation if we think of rhombi as matrices that map pairs of edges to other pairs. Let V be the complex vector space spanned by labelled edges of the unit triangle, $\{\backslash\!\!+, \backslash\!-, \backslash\!0, \underline{+}, \underline{-}, \underline{0}, +\!\!/, -\!\!/, 0\!\!/\}$. We think of each rhombus in Figure 4.11 as a map from $V \otimes V$ to $V \otimes V$ that takes the bottom edges to the top edges, multiplying by the weight given in the table. For example, the last rhombus of orientation $\diagup\!\!\!\!\square$ maps $\underline{-} \otimes +\!\!/$ to $x(\,+\!\!/ \otimes \underline{-}\,)$ and other simple tensors to zero. If we denote the maps of each rhombus as $\mathbf{R}_1(x), \mathbf{R}_2(x), \ldots, \mathbf{R}_{30}(x)$, then $\mathbf{R}(x) = \mathbf{R}_1(x) + \mathbf{R}_2(x) + \cdots + \mathbf{R}_{30}(x)$. We show $\mathbf{R}(x)$ satisfies the classical Yang-Baxter equation.

Corollary 4.7.3. *If $x + y + z = 0$, then*

$$(\mathbf{I} \otimes \mathbf{R}(x))(\mathbf{R}(z) \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{R}(y)) = (\mathbf{R}(y) \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{R}(z))(\mathbf{R}(x) \otimes \mathbf{I}).$$

Proof. Consider applying $(\mathbf{I} \otimes \mathbf{R}(x))(\mathbf{R}(z) \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{R}(y))$ to the three bottom edges of a hexagon. We ignore the tensor symbol and glue the edges together to illustrate the hexagon. For $a, b, c \in \{-, +, 0\}$, we have

33

$$= \sum_{d,e,f} \quad \text{(hexagon diagram with } f, y, d, x, z, c, b, a, e\text{)} \quad f \text{(shaded hexagon with } e, d, a, b, c\text{)}$$

$$= (\mathbf{R}(y) \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{R}(z))(\mathbf{R}(x) \otimes \mathbf{I}) \quad {}_a\diagdown{}_b\diagup{}^c$$

where indices in the summations go over all values in $\{-, +, 0\}$ and the second last line follows from Theorem 4.7.2. $\square$

4.8 Commutativity of transfer matrices

To complete the proof of the puzzle rule, we need to show that the transfer matrices commute. Specifically, we have the following theorem.

Theorem 4.8.1. *The matrix* $\mathbf{T}$ *commutes with the matrices* $\mathbf{T}_\pm(x)$.

Proof. We decompose transfer matrices into the rhombi given in Figure 4.11 so that we can apply the Yang-Baxter equation. First consider $\langle g|\mathbf{T}|f\rangle$ for basis elements $|f\rangle, |g\rangle \in \mathcal{G}$. Since $\mathbf{T}$ only uses the α and β tiles, we can decompose it into an infinite row of rhombi of orientation $\diagdown$ or $\diagup$ as long as we set the weight of the γ tile to 0. We choose $\diagdown$ and decompose $\mathbf{T}$ into several copies of $\boxed{0}\diagdown$. Then we have

$$\langle g|\mathbf{T}|f\rangle = \text{(row of rhombi labeled } g_1, g_2, g_3, g_4 \text{ on top; } 0, 0, 0, 0 \text{ inside; } f_1, f_2, f_3, f_4 \text{ below)}$$

$$= \begin{cases} 1 & \text{if } |g\rangle \text{ is obtainable from } |f\rangle \text{ following } \mathbf{T} \\ 0 & \text{otherwise.} \end{cases}$$

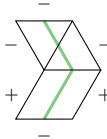The labels f_i and g_i are determined by the signs that f and g respectively assign ($-$, $+$ or 0). Here the row of rhombi simply acts an indicator function that returns 1 if $|g\rangle$ is obtainable from $|f\rangle$ and 0 otherwise. Now consider $\langle g|\mathbf{T}_-(x)|f\rangle$. Since $\mathbf{T}_-(x)$ includes γ_- tiles, we must decompose into rhombi of orientation $\diagup$. Since $\mathbf{T}_-(x)$ assigns γ_- tiles a weight of 1, we use the rhombus $\diagup\boxed{1}$ and account for the weight of the α_- and β_- tiles

34
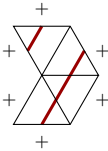
separately. That is

$$\langle g|\mathbf{T}_-(x)|f\rangle = x^{w_-(f,g)} \;\begin{matrix} g_1 & g_2 & g_3 & g_4 \\ \boxed{1}\;\boxed{1}\;\boxed{1}\;\boxed{1} \\ f_1 & f_2 & f_3 & f_4 \end{matrix}$$

$$= \begin{cases} x^{w_-(f,g)} & \text{if } |g\rangle \text{ is obtainable from } |f\rangle \text{ following } \mathbf{T}_-(x) \\ 0 & \text{otherwise.} \end{cases}$$

Again, the row of rhombi is just an indicator function. We consider $\langle g|\mathbf{T}\mathbf{T}_-(x)|f\rangle$ by stacking these diagrams together, however, accounting for the weight of $\mathbf{T}_+(x)$ requires a few tricks. Consider two rows of tiles placed by $\mathbf{T}\mathbf{T}_-(x)$ on top of the diagram of $|f\rangle$.

Sufficiently far to the left, there must only be [diagram] pieces and sufficiently far to the

right, there must only be [diagram] pieces. Both of these boundary regions have weight 1 and thus the nontrivial weight comes from the finite region between them. Now consider the portion of this region due to $\mathbf{T}_-(x)$. Since we can ignore the boundaries, for a basis element $|h\rangle \in \mathcal{G}$, we have

$$\langle h|\mathbf{T}_-(x)|f\rangle = x^{w_-(f,h)} \;^+\begin{matrix} h_1 & h_2 & & h_{k-1} & h_k \\ \boxed{1}\;\boxed{1} & \cdots & \boxed{1}\;\boxed{1} \\ f_1 & f_2 & & f_{k-1} & f_k \end{matrix}^+$$

where the labels f_i and h_i are determined by f and h respectively. Note that, due to the pieces on the boundary, any green lines passing through this finite region must stay within it. Let G be the number of such lines. Also note that all the green lines enter and exit the row using pairs of α_-, β_- or γ_- tiles. Since $x^{w_-(f,h)}$ counts all the α_- and β_- pairs, if we divide by x^G, we will cancel the weights of all those tiles and only count the γ_- tiles with x^{-1}. Setting $y = x^{-1}$, we have

$$y^G \langle h|\mathbf{T}_-(x)|f\rangle = \;^+\begin{matrix} h_1 & h_2 & & h_{k-1} & h_k \\ \boxed{y}\;\boxed{y} & \cdots & \boxed{y}\;\boxed{y} \\ f_1 & f_2 & & f_{k-1} & f_k \end{matrix}^+ .$$

Putting the matrices together, we have

$$y^G \langle g | \mathbf{T}\mathbf{T}_-(x) | f \rangle =$$

Now we multiply on the left by a rhombus z, setting $z = -y$. The only rhombus that can fit on the left is the rhombus which has weight 1; consequently, multiplication by this rhombus does not affect weight. We now apply the Yang-Baxter equation to "unzip" the matrices:

$$y^G \langle g | \mathbf{T}\mathbf{T}_-(x) | f \rangle =$$

$$=$$

$$= y^G \langle g | \mathbf{T}_-(x)\mathbf{T} | f \rangle.$$

Therefore, $\langle g | \mathbf{T}_-(x)\mathbf{T} | f \rangle = \langle g | \mathbf{T}\mathbf{T}_-(x) | f \rangle$ and $\mathbf{T}$ commutes with $\mathbf{T}_-(x)$. We can show $\mathbf{T}$ commutes with $\mathbf{T}_+(x)$ in a similar way. This time consider $\langle g | \mathbf{T}_+(x)\mathbf{T} | f \rangle$ and let R be the number of red lines within in the finite region enclosed by infinite sequences of

and pieces. Then we have

36

$$y^R \langle g | \mathbf{T}_+(x)\mathbf{T} | f \rangle =$$

$$= y^R \langle g | \mathbf{T}\mathbf{T}_+(x) | f \rangle$$

and $\mathbf{T}$ commutes with $\mathbf{T}_+(x)$. $\qquad\square$

One can also prove that that $\mathbf{T}_+(x)$ commutes with $\mathbf{T}_-(y)$, though we will not need this fact. We are now ready to prove the puzzle rule.

4.9 Completing the proof

Let $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$ and $|\lambda\rangle \in \mathcal{F}_0$. We begin with the equation established in Lemma 4.6.2:

$$s_\lambda(\mathbf{x}/\mathbf{y}) = \langle\varnothing| \prod_{i=1}^n \mathbf{T}_+(x_i) \prod_{i=1}^n \mathbf{T}_-(y_i) |\lambda\rangle.$$

Note that only $|\varnothing \sqcup_k \varnothing\rangle$ is obtainable from $|\varnothing\rangle$ after applying $\mathbf{T}^{2k}$. In other words, $\langle\varnothing \sqcup_k \varnothing| \mathbf{T}^{2k} = \langle\varnothing|$. We use this fact and set $k \geq \max(N_+(\lambda), -N_-(\lambda))$.

$$s_\lambda(\mathbf{x}/\mathbf{y}) = \langle\varnothing \sqcup_k \varnothing| \mathbf{T}^{2k} \prod_{i=1}^n \mathbf{T}_+(x_i) \prod_{i=1}^n \mathbf{T}_-(y_i) |\lambda\rangle.$$

Applying commutativity of transfer matrices from Theorem 4.8.1, we have

$$s_\lambda(\mathbf{x}/\mathbf{y}) = \langle\varnothing \sqcup_k \varnothing| \prod_{i=1}^n \mathbf{T}_+(x_i) \prod_{i=1}^n \mathbf{T}_-(y_i) \mathbf{T}^{2k} |\lambda\rangle.$$

Applying Lemma 4.4.2, we get the linear combination

$$s_\lambda(\mathbf{x}/\mathbf{y}) = \sum_{\mu,\nu} \tilde{c}_{\mu,\nu}^\lambda \langle \varnothing \sqcup_k \varnothing | \prod_{i=1}^n \mathbf{T}_+(x_i) \prod_{i=1}^n \mathbf{T}_-(y_i) | \mu \sqcup_k \nu \rangle$$

where $k \geq \max(N_+(\mu), -N_-(\nu))$. Applying Lemma 4.6.3, we have

$$s_\lambda(\mathbf{x}/\mathbf{y}) = \sum_{\mu,\nu} \tilde{c}_{\mu,\nu}^\lambda s_\nu(\mathbf{x}) s_{\tilde{\mu}}(\mathbf{y}).$$

Now applying Theorem 2.2.3, the Littlewood-Richardson coefficients uniquely satisfy this equation, so $\tilde{c}_{\mu,\nu}^\lambda = c_{\mu,\nu}^\lambda$ and hence from Theorem 4.4.3, the puzzle rule is true. □

Chapter 5

The Product of two Factorial Schur Polynomials

In this chapter, we consider the product of two factorial Schur polynomials with the same first set of variables. Molev and Sagan prove a product rule for these polynomials in [27] and Knutson and Tao reformulate the structure coefficients in the summation in terms of MS-puzzles in [17]. In this chapter, we show Zinn-Justin's proof which repurposes the constructions in Chapter 4 to build factorial Schur polynomials and MS-puzzles [37]. Gluing puzzles together and applying the Yang-Baxter equation of Theorem 4.7.2 yields an elegant combinatorial proof of Knutson and Tao's result.

5.1   Factorial puzzles

We assign spectral parameters to thick-edged rhombi using dashed arrows as follows:
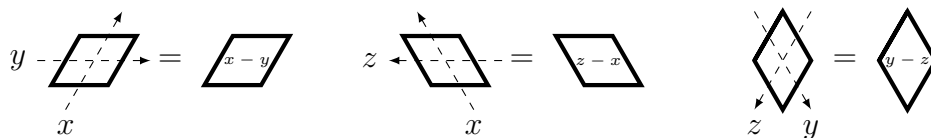
The dashed arrows must intersect two parallel edges and we only use arrows entering at 60 degree angles as above. Each dashed arrow is labelled with a variable and, when they cross within a rhombus, they assign a spectral parameter equal to the clockwise difference

Figure 5.1: Factorial puzzle representing $s_\lambda(x_1, \ldots, x_k | y_1, y_2, \ldots)$.

of their associated variables. That is, if the head of an arrow y is 60 degrees clockwise of the head of an arrow x, the rhombus is assigned the spectral parameter $x - y$. Rotating the rhombi and arrows does not change the spectral parameter assigned.

Let $\mathbf{x} = (x_1, \ldots, x_k)$, $\mathbf{y} = (y_1, y_2, \ldots)$ and $\square = (n - k)^k$ for a positive integers n and k. For $\lambda \subseteq \square$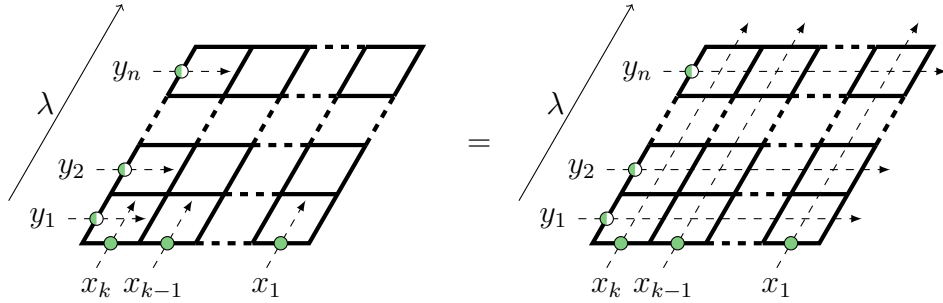, we define **factorial puzzles** as in Figure 5.1. Previously, we referred to puzzles as particular fillings of objects like these, but now we refer to the whole objects (the sum of all fillings) themselves as "puzzles." We only partially draw the dashed arrows and they are assumed to travel in straight lines through the rhombi like in the right hand side of Figure 5.1. If dashed arrows labelled by x_i and y_j cross within a rhombus, it is assigned the spectral parameter $x_i - y_j$. A filling of this shape must have a green line incident to an edge labelled solid green. Blank edges must have no lines incident to them. Edges labelled with half-coloured dots can be either green or blank depending on the Young diagram assigned to that side. The Young diagram λ is encoded using lattice paths within $\square$ as before where green corresponds to a North step and blank corresponds to an East step.

Note that this is equivalent to fixing a $-$ on each edge with a green dot, a $+$ on blank edges on left or right sides, and a 0 for each blank edge on the top side, so there is no ambiguity in how to use the old labels based this new system. Also note that it is important that λ is contained within $\square$ so that λ can fit on the left side and also use k green dots; the k green lines coming from the bottom side exist at these sites.

To see that these are indeed factorial Schur polynomials, first note that only the α_-, β_0 and γ_- tiles can occur within a factorial puzzle since they do not contain red lines. We add dotted lines to these tiles in such a way that crossings occur on γ_- tiles:

40

$$\alpha_- \qquad \beta_0 \qquad \gamma_-$$
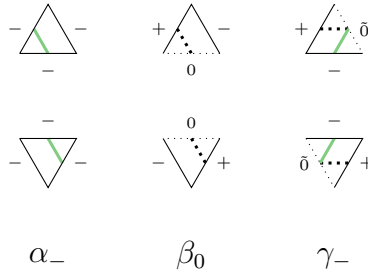
Each filling of a factorial puzzle corresponds to a unique semistandard Young tableau as shown in Figure 5.2. After reflecting, rotating and straightening the filling, we place boxes on each crossing, labelled with its row number. Green lines intersect boxes that correspond to one row of the tableau and dotted lines intersect boxes that correspond to a column. These objects are in bijection with semistandard Young tableaux with maximum entry k. Recall the definition of a factorial Schur polynomial:

$$s_\lambda(\mathbf{x}|\mathbf{y}) = \sum_{T \in \text{SSYT}(\lambda)} \prod_{\alpha \in \lambda} (x_{T(\alpha)} - y_{T(\alpha)+c(\alpha)})$$

where each tableau T has maximum entry k. From the above correspondence, we can sum over fillings of factorial puzzles instead of semistandard Young tableaux. We show each filling has the correct weight using the correspondence between the bottom two pictures in Figure 5.2.

Each box $\alpha = (i, j)$ in the tableau corresponds to a crossing in a filling which is assigned a weight of $x_a - y_b$ for some a, b. We have that $a = T(\alpha)$ as desired since boxes in T are assigned entry a if they come from the row marked by x_a in the centre image. We wish to show that $b = T(\alpha) + c(\alpha) = a + j - i$. Note that the dotted line beginning in column y_j in the centre image intersects the boxes corresponding to column j of the tableau. Then $b = j + d$ where d is the number of diagonal steps the dotted line takes before reaching the box in column y_b. Also, the dotted line must take i vertical steps before reaching the row x_a in the centre image since α is in row i of the tableau. Therefore, $a = i + d$ and hence $b = j + a - i = T(\alpha) + c(\alpha)$ as desired.

5.2 MS-puzzles

Zinn-Justin defines the **MS-puzzle** from Knutson and Tao [17] as in Figure 5.3 where the top right edge has a sequence of k white dots followed by $n - k$ red dots read left to right.
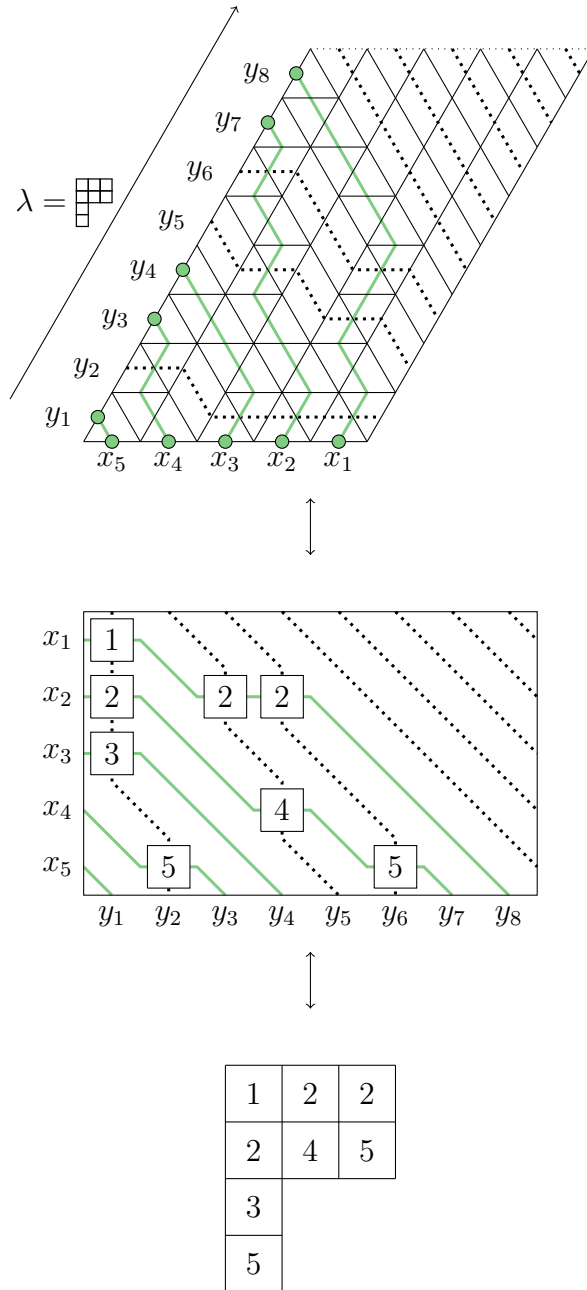
Figure 5.2: A filling of a factorial puzzle contributing to $s_\lambda(x_1, \ldots, x_5 | y_1, y_2, \ldots)$ where $\lambda = $ ▦. This filling has weight $(x_1-y_1)(x_2-y_1)(x_3-y_1)(x_5-y_2)(x_2-y_3)(x_4-y_4)(x_2-y_4)(x_5-y_6)$.
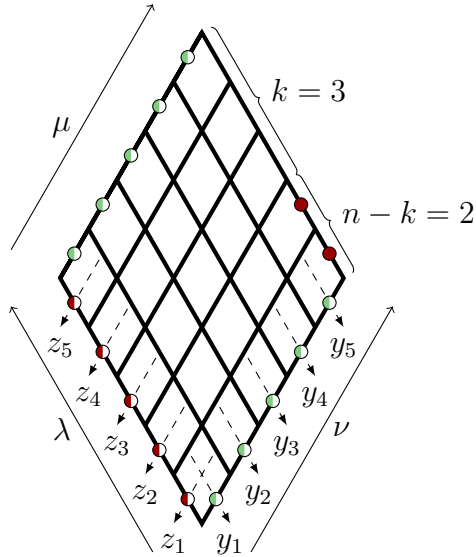
Figure 5.3: An MS-puzzle of weight $c^{\nu}_{\lambda,\mu}(y_1,\ldots,y_5;z_1,\ldots,z_5)$ and a filling of weight $(y_5 - z_1)$.

The other three edges encode Young diagrams. For $\mathbf{y} = (y_1,\ldots,y_n)$, $\mathbf{z} = (z_1,\ldots,z_n)$ and $\lambda,\mu,\nu \subseteq \square$, the weight assigned to these puzzles is denoted $c^{\nu}_{\lambda,\mu}(\mathbf{y};\mathbf{z})$. Also, ν and μ must use k green dots and λ must use $n-k$ red dots. Note that we can arbitrarily increase the size of an MS-puzzle without changing its weight. We illustrate this with a filling of weight $(y_5 - z_1)$ where we increase its size by one:
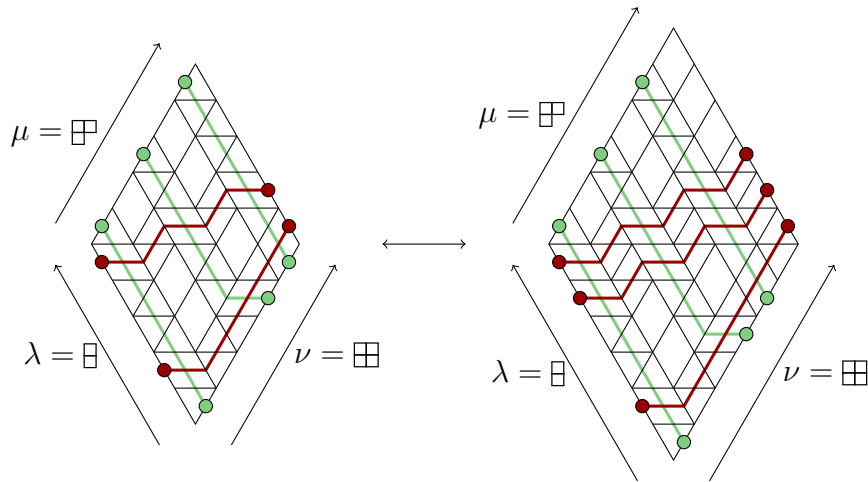
Increasing the size by one adds the leftmost red line to each filling, but this red line has no impact on the weight of the filling since, starting from the South West side, it must move horizontally k times to reach the leftmost red site on the North East side. This red line can then be deleted and the puzzle contracted by one unit. For this reason, we define $c_{\lambda,\mu}^{\nu}(y_1, y_2, \ldots; z_1, z_2, \ldots) = c_{\lambda,\mu}^{\nu}(y_1, \ldots, y_n; z_1, \ldots, z_n)$ where n is as large as it needs to be for the MS-puzzle to be defined. Next, we will need the following important lemma.

Lemma 5.2.1. *If $\mathbf{y} = (y_1, \ldots, y_n)$, $\overleftarrow{\mathbf{y}} = (y_n, \ldots, y_1)$ and $\lambda, \mu \subseteq \square$, then $c_{\lambda,\mu}^{\square}(\mathbf{y}; \overleftarrow{\mathbf{y}}) = \delta_{\lambda,\bar{\mu}}$ where $\bar{\mu}$ is the complement of μ within $\square$.*

Proof. First, note that the rhombi along the middle horizontal are assigned a spectral parameter of 0. If a red and green line cross in an "X" shape (if a γ_0 tile is used) along the middle horizontal, then the filling has weight zero. Thus, every site along the middle horizontal must be occupied by a single green or red line, not both.

For every red line exiting a site in the middle horizontal, there are r red lines and g green lines to its right. It must move diagonally up r times and move horizontal g times to reach the correct site on the North East wall. It can only make a horizontal move when it meets a green line, so it does so for every green line it meets. Therefore, all the green lines starting in the North East wall must move straight diagonally down to the middle horizontal.

Since $\square$ is on the South East wall, it must have k green dots followed by $n - k$ white dots reading from right to left. Reflecting the puzzle vertically and noting the symmetry of the tiles, it follows from the previous paragraph that the red lines in the bottom half of a filling must go diagonally straight up to the right from the South West wall to the middle horizontal.

The middle horizontal encodes μ when read from left to right since the green dots on the North West wall encode μ and follow a straight path to the middle horizontal. It must also encode λ when read from right to left. The complement diagram within $\square$ is what one gets when reading in the reverse order and thus $\mu = \bar{\lambda}$ if and only if there is a valid filling and hence $c_{\lambda,\mu}^{\nu}(\mathbf{y}; \overleftarrow{\mathbf{y}}) = \delta_{\lambda,\bar{\mu}}$. $\square$

5.3 The product

With the constructions above, we can prove the product rule for two factorial Schur polynomials with the same first set of variables.
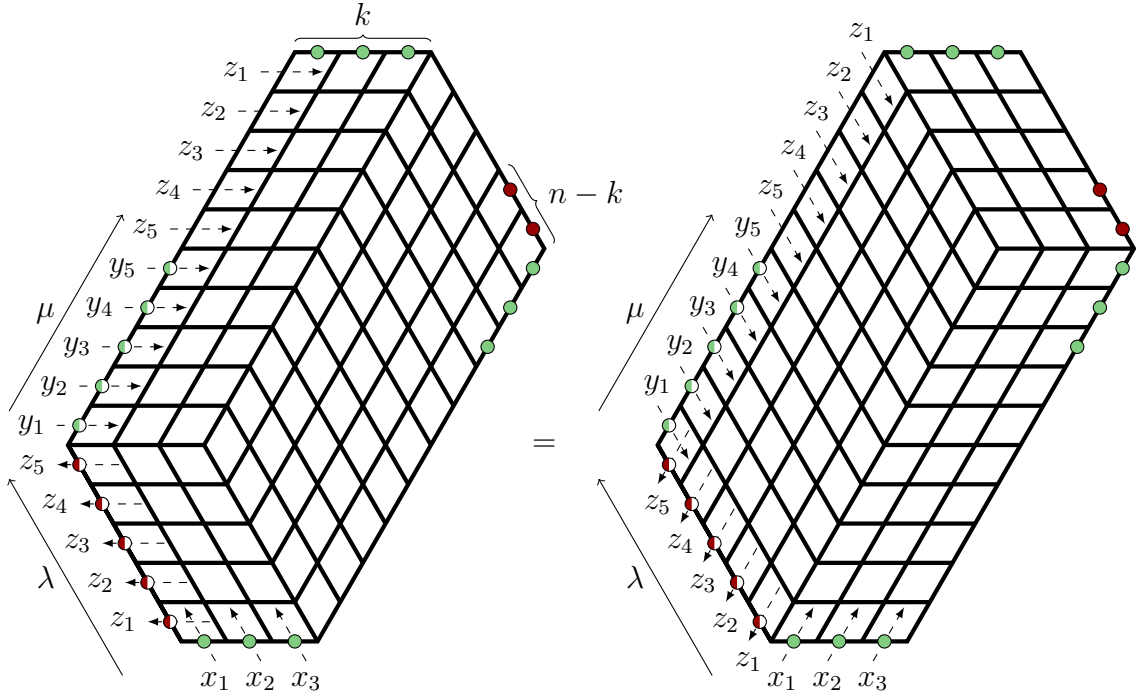
Figure 5.4: Product of two factorial Schur polynomials when $n = 5, k = 3$.

Theorem 5.3.1. *Let* $\mathbf{x} = (x_1, \ldots, x_k)$, $\mathbf{y} = (y_1, y_2, \ldots)$ *and* $\mathbf{z} = (z_1, z_2, \ldots)$. *Then we have*

$$s_\lambda(\mathbf{x}|\mathbf{z})s_\mu(\mathbf{x}|\mathbf{y}) = \sum_\nu c_{\lambda,\mu}^\nu(\mathbf{y}; \mathbf{z})s_\nu(\mathbf{x}|\mathbf{y}).$$

Proof. We use finitely many variables so that we can work with puzzles. Set $\mathbf{y} = (y_1, \ldots, y_n)$ and $\mathbf{z} = (z_1, \ldots, z_n)$ where $n \geq \lambda_1 + \mu_1 + k$. We later explain why this choice of n is sufficiently large. We assume $\lambda, \mu \subseteq \square$ since both sides are zero otherwise.

The proof is now just a matter of proving the equivalence of the puzzles shown in Figure 5.4 and reading off each side. In each puzzle, the top and bottom sides have length k and are entirely filled with green dots. The top right side is length n and has $n - k$ consecutive red dots read from right to left. The bottom right side is length $2n$ and has k consecutive green dots read from right to left. The Young diagrams $\lambda, \mu \subseteq \square$ are encoded with $n - k$ red dots and k green respectively. The dashed arrows entering the bottom are labelled with the variables $x_1, \ldots, x_k$ from left to right. Reading bottom to top, the arrows

45

exiting the bottom left side are labelled with $z_1, \ldots, z_n$ and the arrows entering the top left side are labelled $y_1, \ldots, y_n, z_n, \ldots, z_1$.

Note that the dashed arrows are assumed to bend at 120 degree angles when they meet a rhombus of a new orientation. In this way, the following hexagon is in the centre:
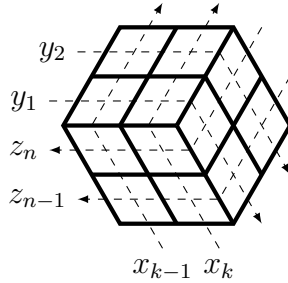
Observe that the unit hexagon in the centre has spectral parameters that sum to zero and thus we can apply the Yang-Baxter equation of Theorem 4.7.2:

Repeatedly applying the Yang-Baxter equation outwards from the centre, we see the left hand side in Figure 5.4 is equal to the right hand side. We now show that both sides represent the desired quantities. As visual aids, we give example fillings belonging to both sides of the equality in Figure 5.5 and Figure 5.6. We only draw the coloured lines and indicate weighted tiles, splitting each filling into five regions.

Consider Figure 5.5, a filling contributing to the left hand side of Figure 5.4. There is a unique way to fill regions A and B which contributes weight 1. The green lines in A are forced to travel diagonally up and to the left and then red lines must move straight horizontally right into region B where they can only travel diagonally, up and to the right. Next, we recognize region C as the MS-puzzle $c_{\lambda, \gamma}^{\square}(\mathbf{z}; \overleftarrow{\mathbf{z}})$ for some Young diagram γ. Then from Lemma 5.2.1, we have $\gamma = \bar{\lambda}$ and thus the North West side of C is λ read from right to left and C has one filling of weight 1. Region D is the factorial puzzle $s_\lambda(\mathbf{x}|\mathbf{z})$ only rotated 180 degrees. Lastly, E is the factorial puzzle $s_\mu(\mathbf{x}|\mathbf{y})$ only with the order of the x variables reversed. However, since factorial Schur polynomials are symmetric in the first set of variables, it is unchanged and the left hand side is the product $s_\lambda(\mathbf{x}|\mathbf{z})s_\mu(\mathbf{x}|\mathbf{y})$.

Consider Figure 5.6, a filling contributing to the right hand side of Figure 5.4. If green lines from D pass directly into E without entering regions A or C, there is a unique way

46
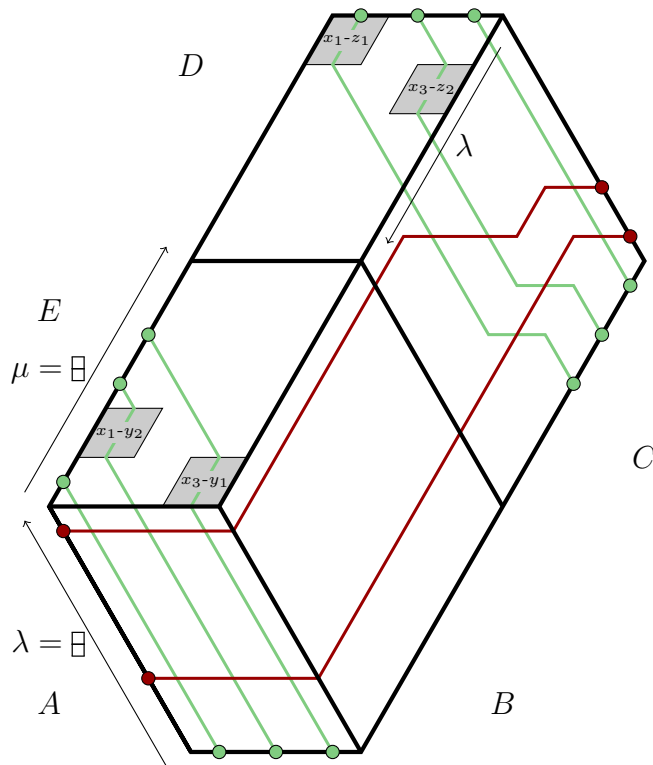
Figure 5.5: A filling of weight $(x_1 - y_2)(x_3 - y_1)(x_1 - z_1)(x_3 - z_2)$ contributing to the left hand side of Figure 5.4 when $\lambda = \mu = \boxminus$.
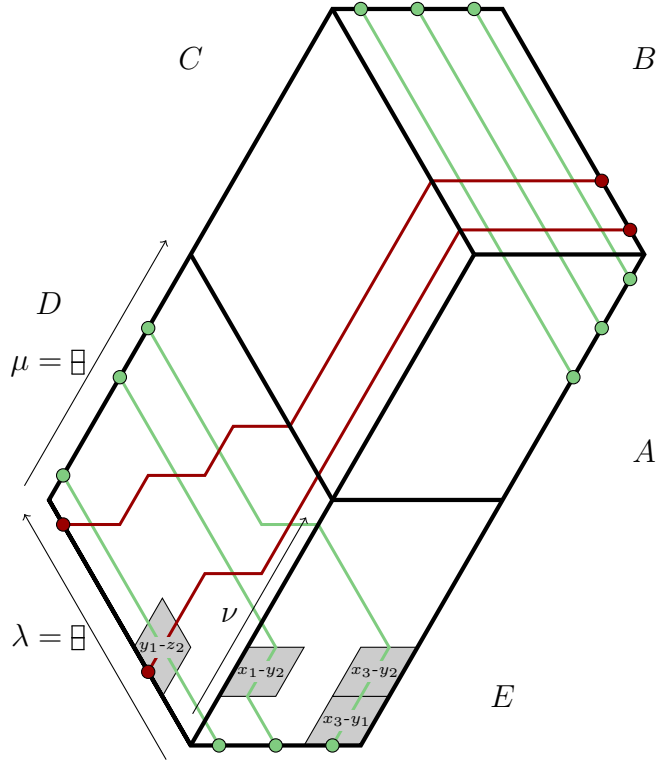
Figure 5.6: A filling of weight $(y_1 - z_2)(x_1 - y_2)(x_3 - y_1)(x_3 - y_2)$ contributing to the right hand side of Figure 5.4 when $\lambda = \mu = \boxminus$.

to fill regions A, B and C, contributing weight 1. We argue that setting $n \geq \lambda_1 + \mu_1 + k$ ensures the green lines in D behave this way. Number the sites along the North West side of D from left to right so that the rightmost green dot is at position $k + \mu_1$. From this position, a green line may take h horizontal steps to the right and end up at position $h + \mu_1 + k$ on the South East side of D. Therefore, $n \geq h + \mu + k$.

On the South West side of D, red lines beginning at sites with no blank sites to their left must take a horizontal step each time they cross a green line to ensure they reach the correct site on the North East side. The remaining red lines contribute to the width of λ and so there are λ_1 of them. Therefore, $h \leq \lambda_1$ and hence we must have $n \geq \lambda_1 + \mu_1 + k$ to ensure green lines in D enter directly into E. Then the region enclosed by D is the MS-puzzle $c_{\lambda,\mu}^{\nu}(\mathbf{y}; \mathbf{z})$ for a variable ν encoded on its South East side. Also, the region enclosed

Figure 5.7: A factorial puzzle representing $(-1)^{|\tilde{\lambda}|}s_{\tilde{\lambda}}(x_1, \ldots, x_k|y_1, y_2, \ldots)$.

by E is the factorial puzzle $s_\nu(\mathbf{x}|\mathbf{y})$. Since internal edges are summed over, we have

$$s_\lambda(\mathbf{x}|\mathbf{y})s_\mu(\mathbf{x}|\mathbf{z}) = \sum_\nu c^\nu_{\lambda,\mu}(\mathbf{y};\mathbf{z})s_\nu(\mathbf{x}|\mathbf{y}).$$

$\square$

5.4 An identity

We have a small original contribution after some experimentation with the factorial puzzles. We introduce the puzzles in Figure 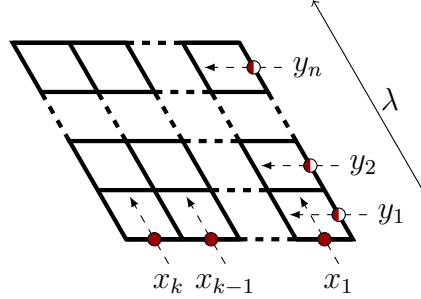5.7 which are simply the factorial puzzles of Figure 5.1 reflected horizontally, changing green lines to red. Due to the symmetry of the red and green decorations under horizontal reflection, this puzzle gives a factorial Schur polynomial with a few differences. The dashed arrows are now on opposite sides of the rhombi in this puzzle and thus contribute the opposite sign as they would in a standard factorial puzzle. Also, green dots correspond to a $-$ and white dots correspond to a $+$ in factorial puzzles whereas red dots correspond to a $+$ and white dots correspond to a $-$ in these puzzles. If all the signs in a $\pm$-string are flipped to the opposite sign, the new string encodes the conjugate of the complement of the original Young diagram. Thus, if λ is on the side of the puzzle in Figure 5.7, it gives a factorial puzzle on $\tilde{\lambda}$. Putting this all together, these puzzles represent $(-1)^{|\tilde{\lambda}|}s_{\tilde{\lambda}}(x_1, \ldots, x_k|y_1, y_2, \ldots)$. With this fact and the other constructions of this chapter, we show the following identity.

Proposition 5.4.1. *Let* $\mathbf{x}_\ell = (x_1, \ldots, x_\ell)$, $\mathbf{y} = (y_1, y_2, \ldots)$ *and* $\square = (n-k)^k$ *for positive integers* n, k *and* ℓ. *Then we have*

$$\sum_{\lambda \subseteq \square}(-1)^{|\lambda|}s_\lambda(\mathbf{x}_k|\mathbf{y})s_{\tilde{\lambda}}(\mathbf{x}_{n-k}|\mathbf{y}) = 0.$$

49
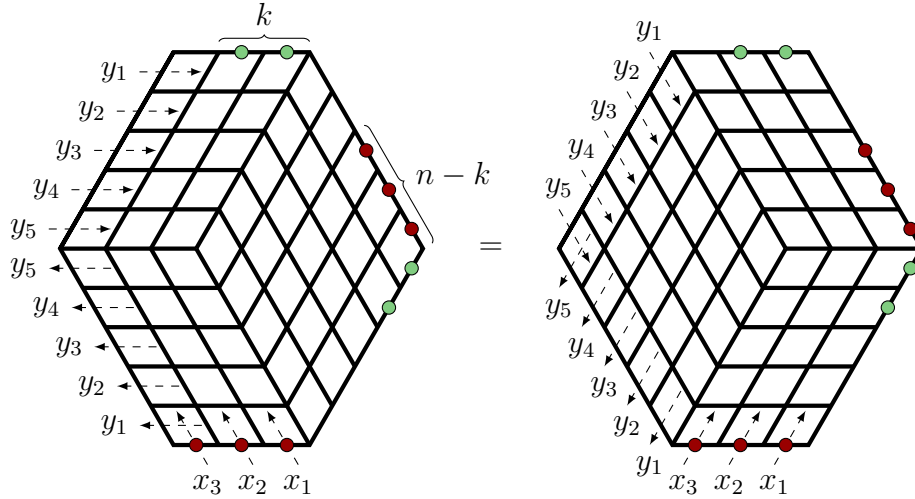
Figure 5.8: Illustration of Proposition 5.4.1 when $n = 5, k = 2$.

Proof. As in the previous section, we glue puzzles together and apply the Yang-Baxter equation of Theorem 4.7.2 to show the equality in Figure 5.8. The top and bottom of the hexagon have length $\max(n, n-k)$ and the sides have length n. Understanding this picture completes the proof. We show a filling contributing to the left hand side and attempt of filling the right hand side (which cannot be done).
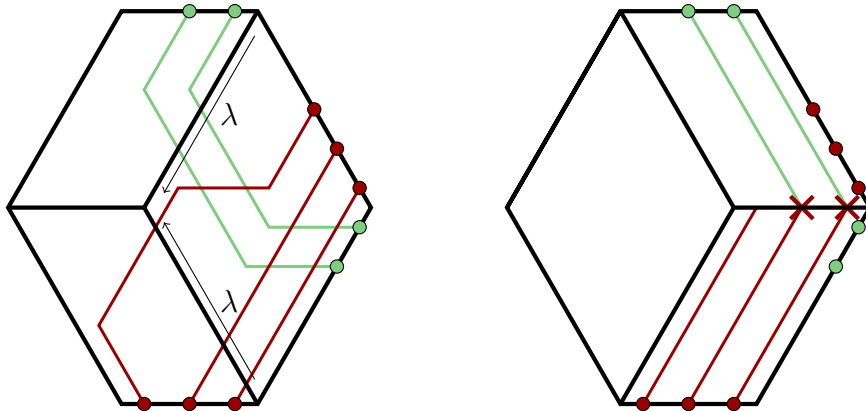
We first consider the puzzle on the left hand side. In any filling, the red lines entering the bottom left region from below will exit that region on the right, encoding some Young diagram $\lambda \subseteq \square$ indicated. Then, by Lemma 5.2.1, the green lines entering the top left region from above must form the Young diagram λ as indicated. For each $\lambda \subseteq \square$, the top

left region contributes $s_\lambda(\mathbf{x}_k|\mathbf{y})$ and the bottom left contributes $(-1)^{|\tilde{\bar{\lambda}}|}s_{\tilde{\bar{\lambda}}}(\mathbf{x}_{n-k}|\mathbf{y})$. Thus, the left hand side is

$$\sum_{\lambda \subseteq \square} (-1)^{|\tilde{\bar{\lambda}}|}s_\lambda(\mathbf{x}_k|\mathbf{y})s_{\tilde{\bar{\lambda}}}(\mathbf{x}_{n-k}|\mathbf{y}),$$

which is equivalent to the left hand side of Proposition 5.4.1.

Moving onto the right hand side, we see that green lines can only move diagonally down and to the right from the top. Similarly, the red lines can only move diagonally up and to the right from the bottom. Red and green lines must clash in the middle, but γ_0 tiles are only allowed in the regions of orientation $\diamondsuit$, so there is no such filling and the right hand side has weight 0. $\square$

Chapter 6

Further Applications

In this chapter, we give an overview of more current research related to the results and techniques of the previous chapters. We first detail product rules for Grothendieck polynomials found by Wheeler and Zinn-Justin [35] where they used the same approach as in Chapter 5. Next, we show how all the previous results can be thought of integrable vertex models and give a brief exposition on results in this area.

6.1 The product of two Grothendieck polynomials

As before, let $\square = (n-k)^k$ for positive integers n and k. Grothendieck polynomials were first introduced by Lascoux and Schützenberger in [20]; they are indexed by permutations $\sigma \in S_n$, but we only consider a certain subset which can be indexed by Young diagrams $\lambda \subseteq \square$. These are polynomial representatives in the K-theory of the Grassmannian. Wheeler and Zinn-Justin find Littlewood-Richardson coefficients for Grothendieck polynomials and *double* Grothendieck polynomials, along with their respective duals in [35]. They modify the approach of Chapter 5 to determine Littlewood-Richardson coefficients in terms of puzzles, showing the technique has general applicability. The results concerning the dual double Grothendieck polynomials are new; others were previously shown in [5, 29, 32].

The authors calculate double Grothendieck polynomials using puzzles with the same structure as factorial puzzles. The same tiles from Figure 4.3 are used in addition to a new tile which only appears in the upright orientation:

Assign the rhombi the following weights:

1	1	1	1	1	1	x	x	1	$1-x$	-1
1	1	1	1	1	x	1	x	1	$1-x$	-1
1	1	1	1	1	1	x	1	x	$1-x$	$-x$

Thick-edged rhombi are defined as before with these new weights and are assigned spectral parameters with dashed arrows, only with a division sign instead of subtraction:

Let $\mathbf{x} = (x_1, \ldots, x_k)$, $\mathbf{y} = (y_1, \ldots, y_n)$ and $\lambda \subseteq \square$. The **double Grothendieck polynomials** (left) and their duals (right) are given by the following puzzles:

$$G^\lambda(\mathbf{x}; \mathbf{y}) = $$

$$G_\lambda(\mathbf{x}; \mathbf{y}) = $$

The standard **Grothendieck Polynomials** and their duals are then given by $G^\lambda(\mathbf{x}) = G^\lambda(\mathbf{x}; 1, \ldots, 1)$ and $G_\lambda(\mathbf{x}) = G_\lambda(\mathbf{x}; 1, \ldots, 1)$. Next, the following Yang-Baxter equation holds for edge labels $(a, b, c, d, e, f) \in \{-, +, 0\}^6$:

Puzzles are glued together again to construct the following equality when, for instance $n = 5, k = 3$:
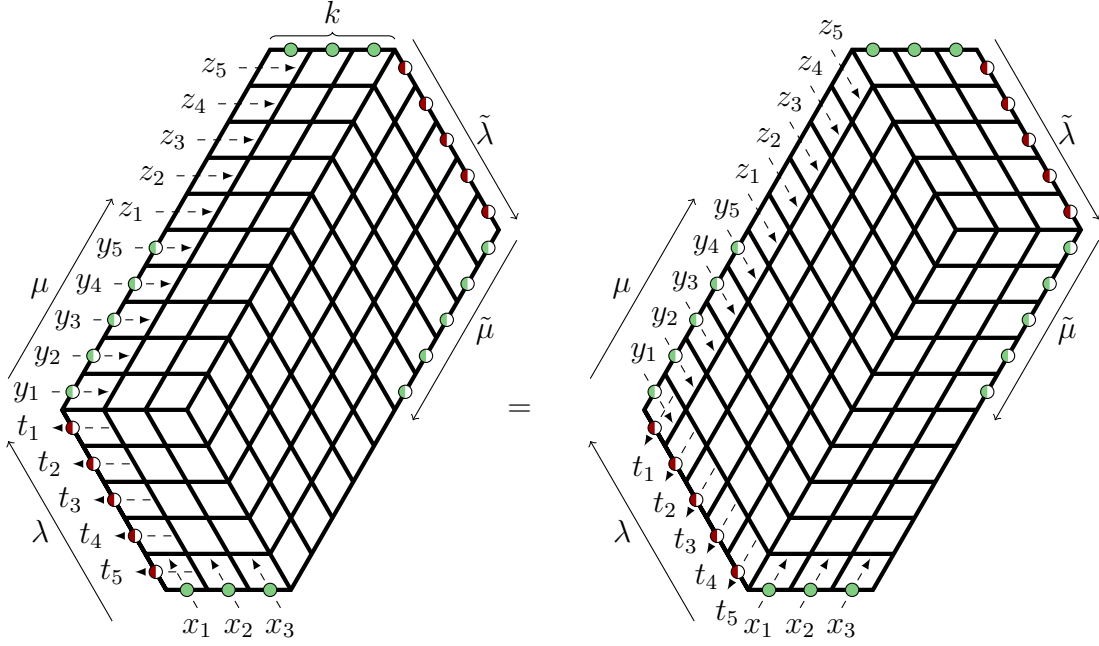
Reading this picture gives the following identity:

$$\sum_\nu c_{\lambda,\tilde\mu}^{\nu,\tilde\lambda}(\mathbf{t};\mathbf{z})G_\nu(\mathbf{x};\mathbf{z})G^\mu(\mathbf{x};\mathbf{y}) = \sum_\nu c_{\lambda,\tilde\nu}^{\mu,\tilde\lambda}(\mathbf{t};\mathbf{y})G_{\tilde\mu}(\mathbf{x};\mathbf{z})G^{\tilde\nu}(\mathbf{x};\mathbf{y})$$

where the coefficients $c_{\lambda,\nu}^{\mu,\gamma}(\mathbf{t};\mathbf{y})$ are determined by diamond puzzles which are generalizations of MS-puzzles with some slight variations.

Through various specializations, this identity allows proof of several product rules for Grothendieck polynomials. For example, taking the specialization $\tilde\lambda = \mu = \varnothing, t = \overleftarrow{\tilde{y}}$, one can show

$$G_\lambda(\mathbf{x};\mathbf{z})G_\mu(\mathbf{x};\mathbf{y}) = \sum_\nu c_{\lambda,\mu}^{\nu,\varnothing}(\mathbf{z};\mathbf{y})G_\nu(\mathbf{x};\mathbf{y}).$$

In total, the authors find Littlewood-Richardson coefficients determined by the puzzles $c_{\lambda,\nu}^{\mu,\gamma}(\mathbf{t};\mathbf{y})$ for several products, which we list for completeness: $G^\lambda(\mathbf{x})G^\mu(\mathbf{x})$, $G_\lambda(\mathbf{x})G_\mu(\mathbf{x})$, $G^\lambda(\mathbf{x};\mathbf{y})G^\mu(\mathbf{x};\mathbf{y})$, $G_\lambda(\mathbf{x};\mathbf{y})G_\mu(\mathbf{x};\mathbf{y})$, $G_\lambda(\mathbf{x};\mathbf{z})G_\mu(\mathbf{x};\mathbf{y})$, $G^\lambda(\mathbf{x};\mathbf{y})G^\mu(\mathbf{x};\overleftarrow{\mathbf{y}})$ and $G_\lambda(\mathbf{x};\mathbf{y})G_\mu(\mathbf{x};\overleftarrow{\mathbf{y}})$.
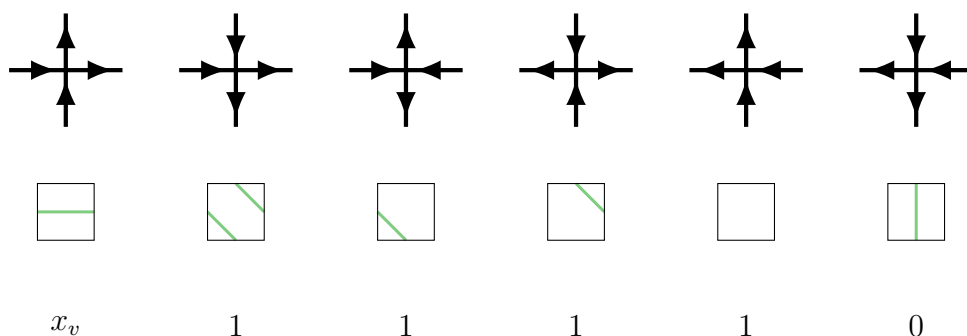
6.2 Vertex models

A **vertex model** is a tool from statistical mechanics that can model, for instance, ice or crystal structures on a lattice where edges represent bonds and vertices represent particles. Many results in combinatorics and integrability are shown through vertex models rather than puzzles. In fact, all the previous puzzles can be viewed as summations over states in vertex models called partition functions.

6.2.1 Puzzles as vertex models

We consider the **six-vertex model** where each vertex has four edges labelled by two incoming and outgoing arrows. There are thus $\binom{4}{2} = 6$ possible vertices, hence the name. Vertex models may be within a fixed boundary or be given a different topology (for instance, torus or cylinder) by identifying boundaries. Each vertex v is assigned a weight called a **Boltzmann weight** $B_v(x_v)$ with a spectral parameter x_v which may depend on v. A filling of a lattice with these vertices is called a state. The **partition function** of a vertex model is given by

$$Z = \sum_V \prod_{v \in V} B_v(x_v)$$

where the summation is over all states V of the vertex model. Below, we have each vertex in a six-vertex model where we associate a familiar looking tile and specify a Boltzmann weight.

| x_v | 1 | 1 | 1 | 1 | 0 |

We let the spectral parameter x_v depend on the coordinate of v in the vertex model. A model with the weights above may be called a "five-vertex model" since one of the vertices has weight zero. Consider filling a $k \times n$ lattice with the above vertices (of nonzero weight)

with incoming arrows along the left and right boundaries, outgoing arrows along the top boundary and a Young diagram $\lambda \subseteq \square$ along the bottom boundary. We show a possible filling using the vertices and tiles:

$$\lambda = \text{⊞}\qquad\qquad\qquad\lambda = \text{⊞}$$
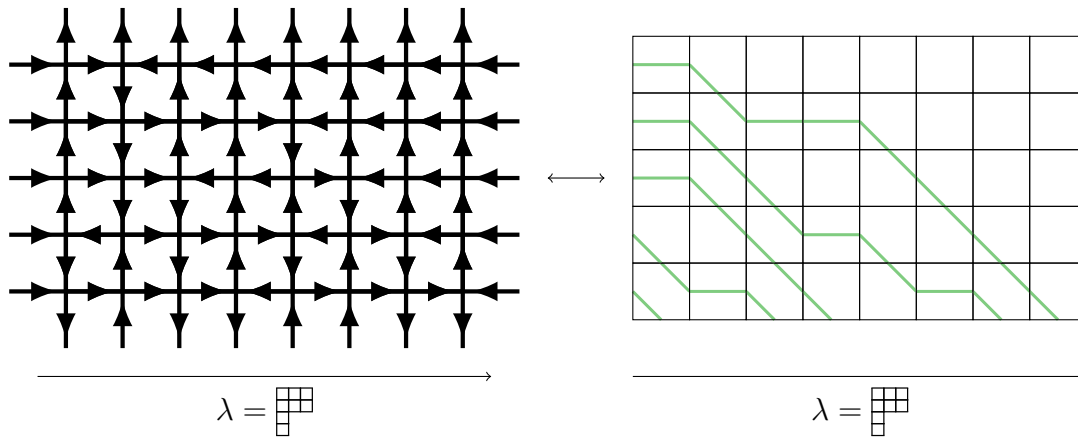
The right hand side is simply the same image as in the middle of the bijection in Figure 5.2. From here, it is straightforward assign spectral parameters to each coordinate in the lattice to get Schur functions, factorial Schur functions or Grothendieck polynomials as we have shown previously. These models are therefore integrable and have an underlying Yang-Baxter equation.

6.2.2   Alternating sign matrices

With the appropriate boundary conditions, six-vertex models count a number of combinatorial objects such as fully packed loops, domino tilings and alternating sign matrices [11, 36]. Alternating sign matrices (ASMs) are square matrices with entries in $\{-1, 0, 1\}$ where each row and column sum to 1 and the nonzero entries alternate along columns and rows. For example, a size 4 ASM is given by

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

ASMs are in bijection with states of $n \times n$ six-vertex models where the top and bottom boundaries point outward and the side boundaries point inward. To get an ASM from a state in the vertex model, replace each vertex with the following integer in place:

| 0 | 0 | 1 | -1 | 0 | 0 |
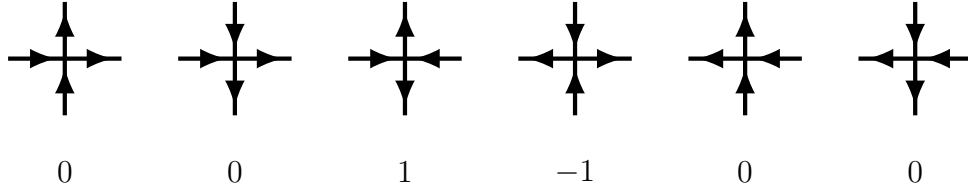
From this bijection, we can set the Boltzmann weight of every vertex to 1 and the partition function counts the number of ASMs of size n. Kuperberg computes this partition function with the help of a Yang-Baxter equation, showing the number of ASMs of size n is $\prod_{k=0}^{n-1} \frac{(3k+1)!}{(n+k)!}$ [19].

6.2.3 Cauchy identities

Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$. The **Cauchy identity** is important in symmetric function theory and states

$$\sum_\lambda s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}) = \prod_{1 \le i,j \le n} \frac{1}{1 - x_i y_j}.$$

Some generalizations of Cauchy identities are found using integrable vertex models. Wheeler, Betea and Zinn-Justin have contributions in a series of papers on Cauchy identities for Hall-Littlewood polynomials derived from a six-vertex model in [2, 3, 34]. We display a result of Motegi and Sakai who use a five-vertex model to find a Cauchy identity for β-Grothendieck polynomials [28].

The **β-Grothendieck polynomials** $G^\lambda(\mathbf{x}; \beta)$ and $G_\lambda(\mathbf{x}; \beta)$ are deformations of the Grothendieck polynomials and their duals given in the previous section. They are given by the following determinantal formulas:

$$G^\lambda(\mathbf{x}; \beta) = \frac{\det[x_j^{\lambda_i + n - i}(1 + \beta z_j)^{i-1}]_{1 \le i,j \le n}}{\prod_{i<j}(x_i - y_j)}$$

$$G_\lambda(\mathbf{x}; \beta) = \frac{\det[x_j^{\lambda_i + n - i}(1 + \beta z_j^{-1})^{1-i}]_{1 \le i,j \le n}}{\prod_{i<j}(x_i - y_j)}.$$

Setting $\beta = -1$ yields $G^\lambda(\mathbf{x}) = G^\lambda(\mathbf{x}; -1)$ and $G_\lambda(\mathbf{x}) = G_\lambda(\mathbf{x}; -1)$; further, setting $\beta = 0$, one recovers the Schur functions $G^\lambda(\mathbf{x}; 0) = G_\lambda(\mathbf{x}; 0) = s_\lambda(\mathbf{x})$. They show

$$\sum_\lambda G^\lambda(\mathbf{x}; \beta) G_\lambda(\mathbf{y}; \beta) = \prod_{i=1}^n \left(\frac{1 + \beta x_i}{1 + \beta y_i^{-1}} \right)^{n-1} \prod_{1 \le i,j \le n} \frac{1}{1 - x_i y_j}.$$

Setting $\beta = 0$ neatly reveals the Cauchy identity for Schur polynomials.

6.2.4 Other results

Buciumas and Scrimshaw derive double Grothendieck polynomials indexed by vexillary (2143-avoiding) permutations as partition functions of a six-vertex model [6]. They show these Grothendieck polynomials are equal to flagged factorial Grothendieck polynomials and also derive a determinantal formulas for double Schubert polynomials corresponding to vexillary permutations.

Wheeler and Zinn-Justin use an **R**-matrix coming from a six-vertex model to show new combinatorial formulas for Hall polynomials and generalized inverse Kostka polynomials [33]. Bump, McNamara and Nakasuji use integrability to study the product of factorial Schur functions and t-deformed Weyl denominators (some function of t) [7]. Specializing t to 0, -1 or ∞ recovers three known definitions of factorial Schur functions. They are also able to show that the factorial Schur functions $s_\lambda(\mathbf{x}|\mathbf{y})$ are asymptotically symmetric in the second set of variables $\mathbf{y}$ using their Yang-Baxter equation.

Di Francesco and Guitter use a twenty-vertex model to address various enumeration problems such as "quarter-turn symmetric domino tilings." They show partition functions are equivalent to the number of tilings in some known problems and conjecture that one of their partition functions counts the number of domino tilings of a certain triangle [12].

Bibliography

[1] A. Berele and A. Regev. Hook Young diagrams with applications to combinatorics and to representations of Lie superalgebras. *Advances in Mathematics*, 64:118–175, 1987.

[2] D. Betea and M. Wheeler. Refined Cauchy and Littlewood identities, plane partitions and symmetry classes of alternating sign matrices. *Journal of Combinatorial Theory, Series A*, 137:126–165, 2016.

[3] D. Betea, M. Wheeler, and P. Zinn-Justin. Refined Cauchy/Littlewood identities and six-vertex model partition functions: II. Proofs and new conjectures. *Journal of Algebraic Combinatorics*, 42:555–603, 2015.

[4] L. C. Biedenharn and J. D. Louck. A new class of symmetric polynomials defined in terms of tableaux. *Advances in Applied Mathematics*, 10:396–438, 1989.

[5] A. S. Buch. A Littlewood-Richardson rule for the K-theory of Grassmannians. *Acta Mathematica*, 189(1):37–78, 2002.

[6] V. Buciumas and T. Scrimshaw. Double Grothendieck polynomials and colored lattice models. Submitted, 2020.

[7] D. Bump, P. J. McNamara, and M. Nakasuji. Factorial Schur functions and the Yang-Baxter equation. 2014.

[8] J.-S. Caux and J. Mossel. Remarks on the notion of quantum integrability. In *Proceedings, 24th IUPAP International Conference on Statistical Physics*, pages 2599–2614, Cairns, Australia, 2011.

[9] J. Clemente-Gallardo and G. Marmo. Towards a definition of quantum integrability. *International Journal of Geometric Methods in Modern Physics*, 6(1):129–172, 2009.

[10] L. Corwin, Y. Ne'eman, and S. Sternberg. Graded Lie algebras in mathematics and physics (Bose-Fermi symmetry). *Reviews of Modern Physics*, 47(3):573–603, 1975.

[11] P. Di Francesco. Integrable combinatorics. In *Proceedings of International Congress of Mathematicians*, volume 3, pages 2599–2614, Rio de Janeiro, 2018.

[12] P. Di Francesco and E. Guitter. Twenty-vertex model with domain wall boundaries and domino tilings. *Electronic Journal of Combinatorics*, 27(2):P2.13, 2020.

[13] J. V. Jeugt, J. W. B. Hughes, R. C. King, and J. Thierry-Mieg. Character formulas for irreducible modules of the Lie superalgebras sl(m/n). *Journal of Mathematical Physics*, 31(9):2278–2304, 1990.

[14] M. Jimbo and T. Miwa. Solitons and infinite dimensional Lie algebras. *Research Institute for Mathematical Sciences, Kyoto University*, 19:943–1001, 1983.

[15] V. G. Kac. Lie superalgebras. *Advances in Mathematics*, 26:8–96, 1977.

[16] D. Karanth, D. Richmond, and J. R. Schmidt. A first course in the Yang-Baxter equation. *Canadian Journal of Physics*, 86(10):1153–1176, 2011.

[17] A. Knutson and T. Tao. Puzzles and (equivariant) cohomology of Grassmannians. *Duke Mathematical Journal*, 119(2):221–260, 2003.

[18] A. Knutson, T. Tao, and C. Woodward. The honeycomb model of $GL_n(\mathbb{C})$ tensor products II: puzzles determine facets of the Littlewood-Richardson cone. *Journal of the American Mathematical Society*, 17(1):19–48, 2003.

[19] G. Kuperberg. Another proof of the alternative-sign matrix conjecture. *International Mathematics Research Notices*, 1996(3):139–150, 1996.

[20] A. Lascoux and M.-P. Schützenberger. Structure de Hopf de l'anneau de cohomologie et de l'anneau de Grothendieck d'une variété de drapeaux. *C. R. Acad. Sci. Paris Sér. I Math.*, 295(11):629–633, 1982.

[21] I. G. Macdonald. Schur functions: theme and variations. *Séminaire Lotharingien de Combinatoire*, 28(498/S–27):5–39, 1992.

[22] N. Metropolis, G. Nicoletti, and G.-C. Rota. A new class of symmetric functions. *Advances in Mathematics Supplementary Studies*, 7B:563–575, 1981.

[23] L. C. Mihalcea. Giambelli formulae for the equivariant quantum cohomology of the Grassmannian. *Transactions of the American Mathematical Society*, 360(5):2285–2301, 2008.

[24] E. Moens. *Supersymmetric Schur functions and Lie superalgebra representations*. PhD thesis, Ghent University, 2007.

[25] A. Molev. Factorial supersymmetric Schur functions and super Capelli identities. *Canberra, ACT: Australian National University, Centre for Mathematics and its Applications, School of Mathematical Sciences*, 1996.

[26] A. Molev. Comultiplication rules for the double Schur functions and Cauchy identities. *Electronic Journal of Combinatorics*, 16(R13), 2009.

[27] A. Molev and B. Sagan. A Littlewood-Richardson rule for factorial Schur functions. *Transactions of the American Mathematical Society*, 351(11):4429–4443, 1999.

[28] K. Motegi and K. Sakai. Vertex models, TASEP and Grothendieck polynomials. *Journal of Physics A: Mathematical and Theoretical*, 46(35):355201, 2013.

[29] O. Pechenik and A. Yong. Equivariant K-theory of Grassmannians II: the Knutson-Vakil conjecture. *Compositio Mathematica*, 153(4):667–677, 2017.

[30] J. H. Perk and H. Au-Yang. Yang-Baxter equations. *Encyclopedia of Mathematical Physics*, 5:465–473, 2006.

[31] R. P. Stanley. Unimodality and Lie superalgebras. *Studies in Applied Mathematics*, 32:263–281, 1985.

[32] R. Vakil. A geometric Littlewood-Richardson rule. *Annals of Mathematics, Second Series*, 164(2):371–422, 2006.

[33] M. Wheeler and P. Zinn-Justin. Hall polynomials, inverse Kostka polynomials and puzzles. *Journal of Combinatorial Theory*, 159:107–163, 2016.

[34] M. Wheeler and P. Zinn-Justin. Refined Cauchy/Littlewood identities and six-vertex model partition functions: III. Deformed bosons. *Advances in Mathematics*, 299:543–600, 2016.

[35] M. Wheeler and P. Zinn-Justin. Littlewood-Richardson coefficients for Grothendieck polynomials from integrability. *Journal für die Reine und Angewandte Mathematik*, 2019(757):159–195, 2019.

[36] P. Zinn-Justin. Integrability and combinatorics: selected topics. http://www.lpthe.jussieu.fr/~pzinn/semi/intcomb.pdf, 2008. Les Houches lecture notes.

[37] P. Zinn-Justin. Littlewood-Richardson coefficients and integrable tilings. *The Electronic Journal of Combinatorics*, 16(1), 2009.