Representation-theoretic approaches to several problems in probability

Kieran Ryan

Supervisor: Prof. Sasha Sodin

Submitted in partial fulfillment of the requirements
of the Degree of Doctor of Philosophy

School of Mathematical Sciences
Queen Mary University of London

Abstract

In this thesis we study certain random walks on the two-dimensional lattice, known as the Manhattan and Lorentz Mirror models, and certain quantum spin systems which are generalisations of the quantum Heisenberg model. The topics are united by the fact that we use the Brauer and walled Brauer algebras, and the representation theory of these algebras, to study both.

We give an overview of Brauer and walled Brauer algebras, as well as that of the symmetric group and the classical groups, and the representation theory of general finite-dimensional algebras. A key feature of the representation theory of the groups and algebras studied in this thesis is called Schur-Weyl duality. We give an account of this theory, as well as applying it to our work on quantum spin systems.

We study the Manhattan and Lorentz Mirror models on a cylinder of finite width. We give an estimate on the vertical distance travelled by the walk along the cylinder, as the cylinder width grows large. We use the Brauer algebra to depict paths of these walks through the cylinder.

Our work on quantum spin systems is split into two parts, studying two classes of models. The first is a class on the complete graph, and the second is an inhomogeneous class, which includes models on the complete bipartite graph. In each case we give the free energy, and formulae for certain magnetisation and total spin observables. We then use these results to give formulae for points of phase transitions, as well as to describe the phases of the models. For the complete graph models, we are able to draw phase diagrams.

Acknowledgements

First and foremost, I am immensely grateful to my supervisor, Sasha Sodin, for his guidance, both mathematical and personal, his warmth, his enthusiasm for maths, his rigour as a mentor, his keen eye as an editor, his patience and understanding, and his humour. I am ten times the mathematician and researcher I was when I started this doctoral programme, and that is in large part down to him.

Many thanks to Jakob Björnberg and Hjalmar Rosengren, my collaborators for the third paper presented in this thesis, with whom it was a joy to work. Special thanks to Davide Macera, who introduced me to interchange processes and the quantum Heisenberg model, and who was a wonderful collaborator. Many thanks to Daniel Ueltschi, Inna Entova-Aizenbud, Dave Platt, Arun Ram and Thomas Prellberg, for careful listening and technical advice.

I would like to thank my assessors, Daniel Ueltschi and Boris Khoruzhenko, and my assessors in my first, second and third year reviews, Ilya Goldsheid and Sasha Gnedin; after a long time of mostly solitary study, it is always a joy to talk about maths with dedicated listeners.

Thanks to all the members (past and present) of the probability group at Queen Mary, especially Stephen Muirhead, Matthias Täufer, Anna Maltsev and Mira Shamis. Thanks to Shabnam Beheshti, Tomasz Popiel and Alex Fink for teaching and holistic support, and a massive thanks to all the admin staff at Queen Mary.

Thanks to the admin and teaching staff at LSE, who made me feel immensely welcome when I taught there.

A massive shout out to the maths PhD cohort at Queen Mary. Thanks for the pints, the common room lunches, the France House memories, the cups of tea, the democratic organising, the union organising, the spoons breakfasts, some great house parties, and the solidarity and friendship through lockdown.

Thanks to my maths friends from Imperial and beyond, especially Costanza, Chloe, Colin, and the common room crew - you know who you are.

It takes a village to raise a kid, and it takes many teachers to raise a mathematician. I've been immensely fortunate throughout my education to have had teachers who saw and nurtured my love for maths. So, an enormous and many-years-in-the-making thanks to Mrs. Barry, Mrs. Cowen, Mrs. Summers, Mrs. Shiel, Mrs. Lester, Mrs. Wallace and Mrs. Ward, Mrs. Jones and Mrs. Stead, Mr. Kendal, Mr. Carney and Mrs. Sweeney, all of St. Joseph's R.C. Primary School, to Sarah Martin, Mark Lewis, Neil Platt, and especially

Tony Haimes, all of St. Anselm's College.

And thanks to the many incredible lecturers I had at Imperial College, for showing me the wild and wonderful extent of what maths can do: Jonathan Mestel, Emma McCoy, Richard Thomas, Phil Ramsden, Eric Keaveny, Mauricio Barahona, Sebastian van Strien, David van Dyck, Ari Laptev, Andrew Walton, Robert Nürnberg, Dan Crisan, Kevin Buzzard, Igor Krasovsky, Alex Mijatović, Gustav Holzegel, Alessio Corti, Matthew Towers, Jeroen Lamb, Sasha Ivanov, Madeleine Whybrow, Ed Segal, Travis Schedler, Eleonora Di Nezza, Daniel Pomerleano, John Britnell, Jonathan Marshall, and Martin Liebeck.

Thanks to Goronwy, Mayya, Francesca, Alex, Róisín and Tom, for chats, parkruns, bakeoff, pub quizzes, zooms, patience, love and support.

And all my love to Mum, Dad, Richard, Jennifer, Monty, Helen, Roisin, Alex and Peter, you guys are the best.

This research was supported by the EPSRC Studentship 1936327, and supported in part by the European Research Council starting grant 639305 (SPECTRUM).

Statement of Originality

I, Kieran Ryan, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature: Kieran Ryan

Date: November 7, 2021

Details of collaboration and publications: The results of this thesis are contained in three works. Firstly, "The Manhattan and Lorentz Mirror Models - A result on the Cylinder with low density of mirrors" [90] (Chapter 4). Secondly, "The free energy of a class of $O_{2S+1}(\mathbb{C})$-invariant spin $\frac{1}{2}$ and 1 quantum spin systems on the complete graph" [89] (Chapter 5). Thirdly, "Heisenberg models and Schur–Weyl duality" [12] (Chapter 6), which is joint work with Jakob Björnberg and Hjalmar Rosengren, both of Chalmers University of Technology and University of Gothenburg, Sweden.

Contents

1 Introduction **8**

2 Representation Theory **17**

2.1 Representation theory 17

2.1.1 Representation theory of finite-dimensional algebras 17

2.1.2 The symmetric group 21

2.1.3 The Brauer algebra 23

2.1.4 The walled Brauer algebra 27

2.1.5 Classical groups 29

2.1.6 Branching rules 33

3 Schur-Weyl duality **38**

3.0.1 Proof of the double centraliser theorem, and statements of Schur Weyl Duality 39

3.0.2 Invariant theory proofs of Schur-Weyl duality 43

3.0.3 Constructing irreducibles in Schur-Weyl duality 47

4 The Manhattan and Lorentz Mirror Models **57**

4.1 Introduction 57

4.2 Definitions, and the Brauer algebra 60

4.3 Results 65

5 Quantum Spin Systems on the complete graph **70**

5.1 Classical and quantum spin systems 70

5.2 Introduction to quantum spin systems on the complete graph 77

5.2.1 Models and results 79

5.2.2 Phase diagram for spin $S = \frac{1}{2}$ 82

5.2.3 Phase diagram for spin $S = 1$ 84

5.2.4 Higher spins 86

5.3 The Brauer algebra 89

5.4 Proof of Theorem 5.2.1 91

5.5 Proofs of Theorems 5.2.2, 5.2.3 and 5.2.4 95

5.6 Phase diagrams 98

5.6.1 Higher spins; Proof of Theorem 5.2.10 98

5.6.2 Spin $\frac{1}{2}$, $\theta = 2$ 101

5.6.3 Spin 1; $\theta = 3$ 103

5.7 Branching Coefficients 109
5.7.1 Useful lemmas for all θ 109
5.7.2 Spin $\frac{1}{2}$; $\theta = 2$ 113
5.7.3 Spin 1; $\theta = 3$ 113
5.8 Numerical proof of Remark 5.6.5 114
5.9 Equivalence of $Q_{i,j}$ and $P_{i,j}$ 115

6 Quantum Spin Systems on the complete bipartite graph 117
6.1 Introduction and results 117
6.1.1 Two-block models: Free energy 119
6.1.2 Two-block models: Phase transition and critical temperature 121
6.1.3 Two-block models: Correlations and magnetisation 122
6.1.4 Multi-block models 125
6.1.5 Heuristics for phase diagrams and extremal Gibbs states 127
6.2 Free energy and correlations 128
6.2.1 Interchange model: proof of Theorem 6.1.1 128
6.2.2 Walled Brauer algebra: proof of Theorem 6.1.2 132
6.2.3 Correlation functions: proof of Theorem 6.1.7 137
6.2.4 Magnetisation term: proof of Theorem 6.1.8 139
6.3 The phase-transition 141
6.3.1 Existence of a phase transition: proof of Proposition 6.1.3 142
6.3.2 Formulas for β_c: proofs of Propositions 6.1.4 and 6.1.5 144
6.4 Multi-block models 149
6.5 Form of the maximiser of F for $c > 0$ 150
6.6 The trace-inequality (6.79) 155
6.7 Equivalence of $Q_{i,j}$ and $P_{i,j}$ in the WB-model 155

Glossary 158

Chapter 1

Introduction

Representation theory, as well as being a rich and diverse field in is own right, has found countless interesting applications to probability and physics. In this thesis we present an analysis of certain quantum spin systems, and certain random walks in random environments. The topics are unified by our applying the Brauer algebra, its subalgebra called the walled Brauer algebra, and the representation theory of these algebras, to both of these subjects.

While similar to the well-studied symmetric group algebra, the Brauer and walled Brauer algebras are not group algebras, and their representation theory is often more nuanced. The representation theory of all three algebras is summarised in Chapter 2. A key part of their representation theory, called Schur-Weyl duality, proves crucial in our applications to quantum spin systems. The main objectives of this work are firstly to convey the results on the random walks and spins systems in Chapters 4, 5 and 6, and secondly to deliver the unified account of Schur-Weyl duality found in Chapter 3.

*

Observe that a permutation σ in the symmetric group S_n can be depicted as a diagram of the form in Figure 1.1, which represents the permutation $(24)(56) \in S_6$. Multiplication $\sigma\tau$ of two permutations σ and τ is then given by concatenation of the diagrams - placing the diagram of σ above that of τ and joining the lines together. Recall that the symmetric group algebra $\mathbb{C}S_n$ is the vector space with basis given by permutations S_n.

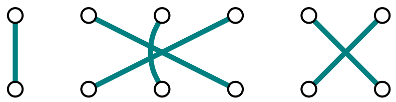

$= (24)(56) \in S_6$

Figure 1.1: The permutation $(24)(56) \in S_6$.

The Brauer algebra $\mathbb{B}_{n,\theta}$ is the vector space with basis given by a larger set of such diagrams (so, $\mathbb{B}_{n,\theta}$ contains $\mathbb{C}S_n$ as a subalgebra); specifically, all diagrams which are pairings of the $2n$ vertices. In particular, an upper vertex can be connected to another upper vertex, and the same with lower vertices. Multiplication is still given by concatenation, and one multiplies the result by the parameter θ to the power of the number of internal loops removed in the concatenation. See Figure 1.2 for an example.

Let $0 \le m \le n$. The walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ is similar, but defined as the span of the diagrams described above which have a certain property. Draw a line (a "wall")

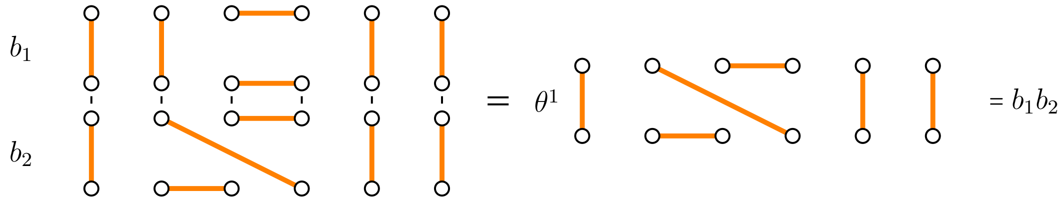


Figure 1.2: Two diagrams b_1 and b_2 (left), and their product (right). The concatenation contains one loop, so we multiply the concatenation (with middle vertices removed) by θ^1.

separating the leftmost $2m$ vertices and the rightmost $2(n-m)$. See Figure 1.3. Then the certain property is that an edge connecting two upper vertices (or two lower vertices) must cross the wall, and an edge connecting an upper with a lower vertex must not cross the wall. Multiplication is the same as in the Brauer algebra.

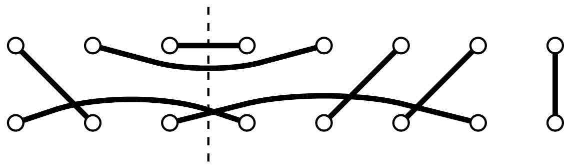


Figure 1.3: A diagram in the basis of the walled Brauer algebra $\mathbb{B}_{8,3}(\theta)$. Notice that all edges connecting two upper vertices (or two lower) cross the wall, and all edges connecting an upper vertex to a lower vertex do not.

The Brauer algebra $\mathbb{B}_{n,\theta}$ was introduced by Brauer [19] as having Schur-Weyl duality with the orthogonal group $O(\theta)$. The original Schur-Weyl duality [104] intimately links the representation theory of the symmetric group S_n with that of the general linear group $GL(\theta)$. Specifically, it studies the action of the two groups on the vector space $V^{\otimes n}$, where $GL(\theta)$ acts diagonally and S_n acts by permuting the tensor factors. Among other things, the theory gives the decomposition of tensor space $V^{\otimes n}$ as a representation of either group, or of the direct product of the two groups. Much more recently, the walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ was introduced as having a Schur-Weyl duality with the general linear group in the work of Turaev [99], Koike [60] and Benkart et al. [8]. The action of the general linear group in this case is different: $g \in GL(\theta)$ acts diagonally, as itself on the first m tensor factors, and as its inverse-transpose on the remaining $n-m$ factors.

One of the aims of this thesis to give an account of Schur-Weyl duality in these three instances ($GL(\theta) - S_n$, $O(\theta) - \mathbb{B}_{n,\theta}$ and $GL(\theta) - \mathbb{B}_{n,m,\theta}$) which is streamlined, as self-contained as possible, and unified (the latter in the sense that the account covers the representation theory of both the classical groups and the Brauer and walled Brauer algebras). This account is given in Chapter 3. Its need arises as a result of the major work on the algebras being relatively recent (in particular the walled Brauer algebra).

The Brauer algebra and walled Brauer algebra have been studied extensively in their own right. One key difference from the symmetric group algebra $\mathbb{C}S_n$ is that while $\mathbb{C}S_n$ is semisimple, $\mathbb{B}_{n,\theta}$ and $\mathbb{B}_{n,m,\theta}$ are non-semisimple for some values of θ, essentially meaning their representation theory is less straightforward. We encounter these non-semisimple cases in the applications in Chapters 5 and 6. For a more comprehensive overview of the two algebras, as well as an overview of their representations, see Sections 2.1.3 and 2.1.4.

*

Our first set of applications pertains to certain random walks on the two-dimensional lattice, the Manhattan and Lorentz Mirror models. Chapter 4 presents the results of the paper "The Manhattan and Lorentz Mirror Models - A result on the Cylinder with low density of mirrors" [90].

For the Manhattan model, imagine the two-dimensional lattice with directions like the streets and avenues of Manhattan (see the left diagram in Figure 1.4). With a fixed number p between 0 and 1, at each intersection of the lattice, place, independently with probability p, a mirror at 45° to the lattice, which reflects a walker left or right. The orientation (i.e. whether it is pointing northwest or northeast) is always so that the walker follows the directions of the lattice. If there is no mirror, the walker continues straight on. For the Mirror model, the lattice has no directions (every road is a two-way street). We still place mirrors at 45° to the lattice at each intersection, independently with probability p, but each time we then toss a $50-50$ coin to determine its orientation. The main question of interest for these models is whether the walks are bounded or not, and beyond that, the nature of the walks.

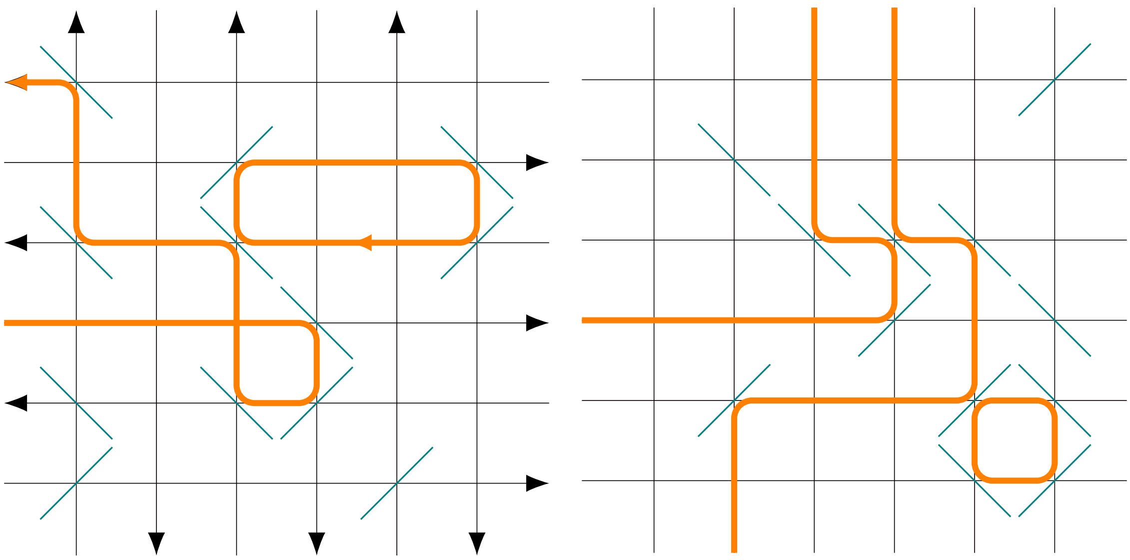


Figure 1.4: Examples of the Manhattan model (left) and Mirror model (right), with mirrors in blue, and a few paths of the walker highlighted in orange. Note that the orientation of a mirror in the Manhattan case is determined by the Manhattan directions of the lattice.

It is straightforward to see that when $p = 1$ the paths are bounded with probability 1 in the Manhattan model. Grimmett [46] gave a simple argument for the same result on the Mirror model. For both models, the same result is expected to hold for all $0 < p \leq 1$. The two models do not have fully identical behaviour though. Kozma and Sidoravicius [62] showed that for all $p > 0$, the probability that two points n steps apart are connected by a path in the Mirror model decays slower than $(2n + 1)^{-1}$. In contrast, Cardy et al. [6] showed that for $p \geq \frac{1}{2}$ this decay in the Manhattan model is exponential in n (and so paths are bounded with probability 1), and the same is expected for all $p > 0$.

One approach that simplifies the models is by considering them on a lattice cylinder (say, of width n). If the cylinder has height n too, one can transfer results back to the planar lattice (see [62]). We consider in Chapter 4 the random variable V given by the highest row of the cylinder the walker reaches above its starting point. A crude argument gives that $\mathbb{P}[V \le k] \le \mathbb{P}[G \le k]$, G a geometric random variable with parameter $(\frac{p}{2})^n$, since a path cannot pass through a street fully occupied by mirrors arranged appropriately. Li [66] showed that for fixed p, $V = O(n^{10})$, with probability exponentially close to 1. The approach we take in Chapter 4 is based on observing that the models on the cylinder can be thought of as Markov chains on the basis of diagrams of the Brauer algebra (or the walled Brauer algebra). Our result is for a low density of mirrors, specifically when $p \le n^{-1}$. In this case, we find that V behaves like p^{-2}: that is, for any $\alpha > 0$,

$$\mathbb{P}[V \le \alpha p^{-2}] \le C_1 e^{-C_2 \alpha}, \tag{1.1}$$

for some constants C_1, C_2 depending on the model chosen. Note we also have the lower bound, valid for all $p < \frac{1}{2}$:

$$\mathbb{P}[V \ge \alpha p^{-2}] \le 2\alpha. \tag{1.2}$$

See Theorem 4.1.1 for full details. Moving forward, one would like to extend this to a theorem for fixed, small values of p on the cylinder, and approach a result on the square cylinder to transfer to $\mathbb{Z}^2$.

*

Our second set of applications is to quantum spin systems. Chapters 5 and 6 present, respectively, the results of the papers "The free energy of a class of $O_{2S+1}(\mathbb{C})$-invariant spin $\frac{1}{2}$ and 1 quantum spin systems on the complete graph" [89], and the joint work with Jakob Björnberg and Hjalmar Rosengren "Heisenberg models and Schur–Weyl duality" [12].

Quantum spin systems are models which aim to derive the macroscopic properties of matter from microscopic interactions of particles via their (quantum) spins, with the particles arranged in a lattice. In particular, often the main goal is to find (and describe) abrupt changes in the model when the parameters involved (for example, temperature) are varied - these are called phase transitions. See the beginning of Chapter 5 for a more detailed introduction to spin systems. In Chapters 5 and 6 we investigate phase transitions for certain explicit classes of models. The results are of two types. Firstly, in each chapter, we obtain for the models considered an explicit formula for a function known as the free energy. The free energy is a function of the parameters of the model (i.e. temperature), and points at which it is non-analytic indicate points of phase transition. The second type of result uses the free energy results to give explicit formulae for points of phase transition, and through the free energy and other working, investigates the properties of the models in different regions of the parameter space. The models studied are generalisations of the well-studied quantum Heisenberg model.

In the classical Heisenberg model, a particle at a site i in a lattice of dimension d is given a spin $\sigma_i = (\sigma_i^{(1)}, \sigma_i^{(2)}, \sigma_i^{(3)}) \in \mathbb{S}^2$, the two-sphere. This models particles of a magnetic material being magnetised in the direction σ_i. Allow a parameter β to represent inverse temperature. Then, for a given β, the probability that a configuration $\sigma = (\sigma_i)_{1 \le i \le n}$ occurs is

$$\phi_{H,\beta}(\sigma) = \phi_\beta(\sigma) = \frac{1}{Z(\beta)} e^{-\beta H(\sigma)}. \tag{1.3}$$

Here the function $Z(\beta) = \int d\sigma \; e^{-\beta H(\sigma)}$ is called the partition function and is the normalisation constant which makes the measure a probability measure, and

$$H(\sigma) = -\sum_{i,j} \sigma_i \cdot \sigma_j \tag{1.4}$$

is the Hamiltonian describing the energy of the configuration, and the sum is over nearest neighbours in the lattice. Spins want to be aligned: the more aligned the spins are, the lower the energy from H. The configurations with lowest energy are those with the spins at all vertices pointing in the same direction. These are relatively few in number, compared with the vast number of configurations which would give a large energy, with neighbouring spins being much less aligned. Tuning β determines which of these types of configuration dominates the measure - for $d \geq 3$ there is a phase transition between these two behaviours. See the beginning of Chapter 5 for more detail.

Two natural generalisations of the Heisenberg model are the XXZ model and the bilinear-biquadratic model. The XXZ model has Hamiltonian

$$H(\sigma) = -\sum_{i,j} K_1(\sigma_i)^{(1)}(\sigma_j)^{(1)} + K_2(\sigma_i)^{(2)}(\sigma_j)^{(2)} + K_1(\sigma_i)^{(3)}(\sigma_j)^{(3)}, \tag{1.5}$$

that is, we give a certain weight K_2 to the interaction in the 2 direction, and a second weight K_1 to the other two directions. (In the literature the K_2 term is often in the 3 direction - the model is equivalent either way. For technical reasons we prefer the 2 direction, and we will maintain this convention throughout this thesis.) For example, if $K_1 > 0$ and $K_2 < 0$, the system wants adjacent spins which point in the $1-3$ plane to be aligned, but those pointing in the 2-axis to be anti-aligned. The bilinear-biquadratic Heisenberg model has Hamiltonian

$$H = -\sum_{i,j} \left(J_1(\sigma_i \cdot \sigma_j) + J_2(\sigma_i \cdot \sigma_j)^2 \right). \tag{1.6}$$

From the first term, for $J_1 > 0$ adjacent spins want to align and for $J_1 < 0$ they want to anti-align. The second term, for $J_2 > 0$, prefers adjacent spins to be either aligned or anti-aligned, but not orthogonal to one another, and vice-versa for $J_2 < 0$.

A quantum spin system is analogous to a classical one, but described using an operator of the form $e^{-\beta H}$ (instead of a probability measure), acting on the phase space of the model, which is a vector space $V^{\otimes n}$, a tensor product of a local phase space V for each of the n particles in the lattice. Here H the Hamiltonian is a Hermitian operator. Each of the Heisenberg, XXZ, and bilinear-biquadratic models have their quantum analogues. The analogue of the Heisenberg model (1.4) has Hamiltonian

$$H = -\sum_{i,j} (\boldsymbol{S}_i \cdot \boldsymbol{S}_j), \tag{1.7}$$

where $\boldsymbol{S}_i = (S_i^{(1)}, S_i^{(2)}, S_i^{(3)})$, $(\boldsymbol{S}_i \cdot \boldsymbol{S}_j) = (S_i^{(1)} S_j^{(1)} + S_i^{(2)} S_j^{(2)} + S_i^{(3)} S_j^{(3)})$ and $S_i^{(k)}$ are explicit Hermitian operators acting on the i^{th} tensor factor of $V^{\otimes n}$. The quantum analogue of the

XXZ model (1.5) has Hamiltonian

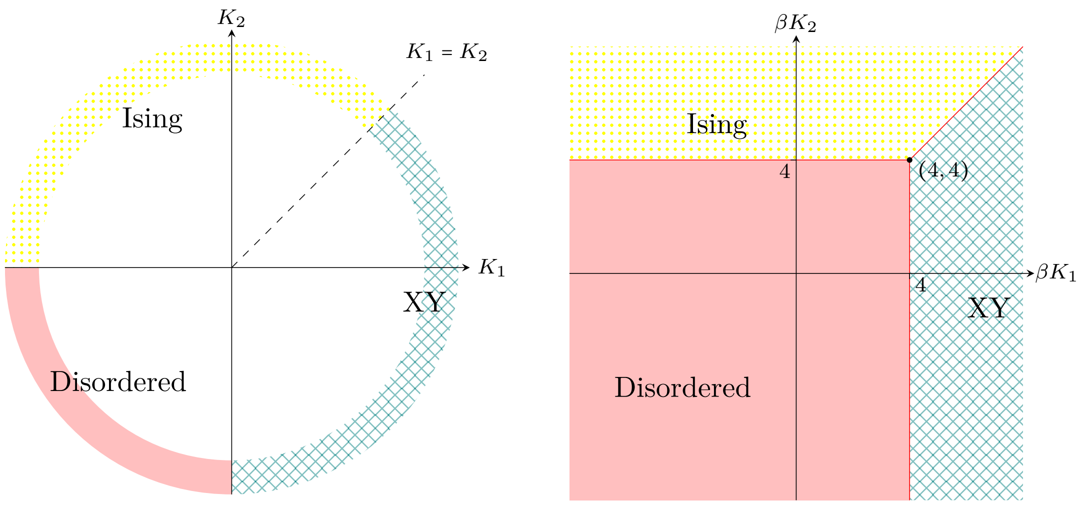
$$H = -\sum_{i,j}\left(K_1 S_i^{(1)}S_j^{(1)} + K_2 S_i^{(2)}S_j^{(2)} + K_1 S_i^{(3)}S_j^{(3)}\right). \tag{1.8}$$

The model can be shown to be equivalent to that with K_2 on the 1 or 3 axis instead of the 2 axis, and, if the underlying graph is bipartite, to the model with K_1 replaced with $-K_1$. On $\mathbb{Z}^d$, $d \geq$, a phase transition is expected for all parameters, and has been shown for several cases, see for example Dyson, Lieb and Simon [38] and Kennedy [58]. Not all cases have been shown to have a transition though, most notably the Heisenberg model itself. The quantum analogue of the bilinear-biquadratic model (1.6) has Hamiltonian

$$H = -\sum_{i,j}\left(J_1(\boldsymbol{S}_i\cdot\boldsymbol{S}_j) + J_2(\boldsymbol{S}_i\cdot\boldsymbol{S}_j)^2\right). \tag{1.9}$$

The expected phase diagram of this model is given in Ueltschi's paper [101]. Again transitions have been proved for some cases of the parameters, for example in [101] or Lees [64], but not all.

A common simplification of a spin system is achieved by replacing the lattice with the complete graph - known as the mean field approximation. Often results in the mean field carry over to or approximate the behaviour of models on the lattice, and often the approximation makes computations easier. Chapter 5 studies a model on the complete graph which, for $\theta = 2$ is the XXZ model (1.8), and for $\theta = 3$ is the bilinear-biquadratic model (1.9). Figure 1.5 and Figure 1.6 give the phase diagrams obtained from our free energy results (Theorem 5.2.1) for these two models, respectively.

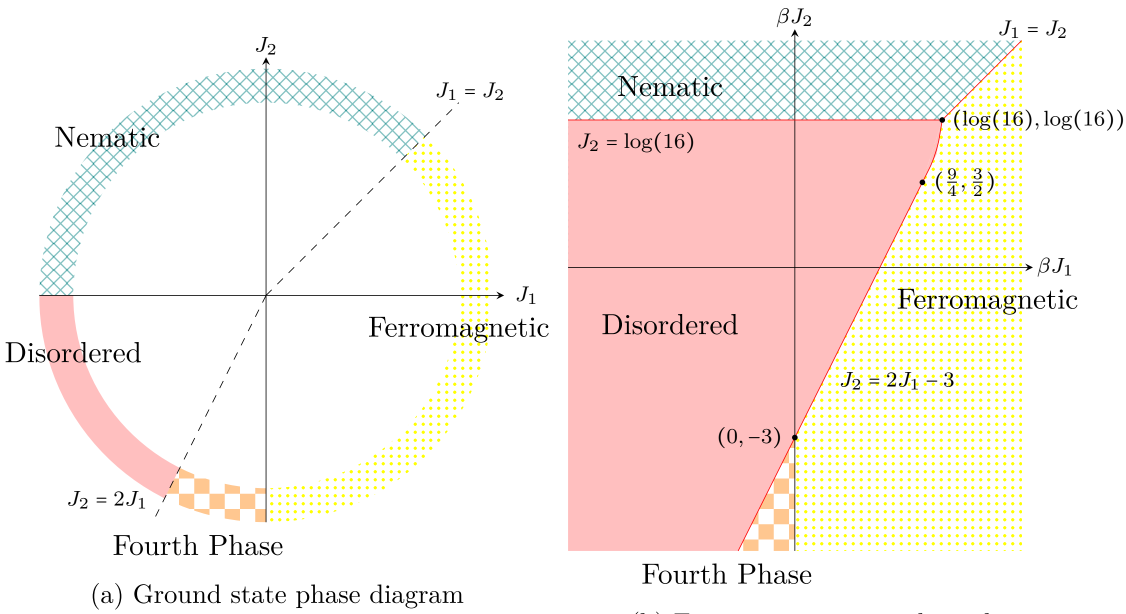


(a) Ground state phase diagram

(b) Finite temperature phase diagram

Figure 1.5: On the left, the ground state phase diagram for the quantum Heisenberg XXZ model with $\theta = 2$ and with Hamiltonian (1.8). The line $K_1 = K_2 \geq 0$ gives the Heisenberg ferromagnet. On the right, the phases at finite temperature, where varying temperature is given by varying the modulus $\|\beta(K_1, K_2)\|$. Transitions between phases (points of non-analyticity of the free energy) shown in red lines.

The right hand diagram in each case shows that the points of non-analyticity of the free energy (and points of phase transition) are given by the red lines. The left hand diagram describes the “ground state” behaviour - how the model behaves as temperature approaches zero (β approaches ∞). For the XXZ model, quite intuitively, in the "Ising" phase the K_2 term of (1.8) dominates, and spins want to be aligned along the 2-axis. In the "XY" phase, the K_1 terms dominate, and the spins want to be aligned, but in the $1-3$ plane. For $K_1, K_2 \leq 0$, there is no phase transition, and the model remains "disordered" for all temperatures.



(a) Ground state phase diagram

(b) Finite temperature phase diagram

Figure 1.6: On the left, the ground state phase diagram for the quantum bilinear-biquadratic Heisenberg model with Hamiltonian (1.9), and $\theta = 3$. On the right, the phases at finite temperature, where varying temperature is given by varying the modulus $\|\beta(J_1, J_2)\|$. Transitions between phases (points of non-analyticity of the free energy) shown in red lines (proved in the region $J_2 \geq J_1$, expected as shown for the rest of the plane).

For the bilinear-biquadratic model, again fairly intuitively, in the "Ferromagnetic" region, the J_1 term dominates, and the model behaves like the usual Heisenberg model (1.7). In the "Nematic" region, the J_2 term dominates. In the classical setting one would expect spins to align or anti-align. See the introduction of Chapter 5 for more detail on how this is expected to manifest in the quantum case. In the "Disordered" region, there is no phase transition. In the "Fourth phase" the behaviour is unclear, although some properties of the Ferromagnetic region are exhibited. Further work could investigate this region further, in addition to attempting to transfer some of these results to the lattice $\mathbb{Z}^d$.

As noted above, the other results of Chapter 5 give further features of the model, and support the interpretations of the phase diagrams above. See Theorems 5.2.3 and 5.2.4. We also give the free energy and phase diagrams for versions of the model with $\theta > 3$; see Theorem 5.2.2 and Figure 5.3b.

In Chapter 6 we study models very similar to those in Chapter 5, except essentially we work on the complete bipartite graph rather than the complete graph. This is a step closer to reality, since the lattice $\mathbb{Z}^d$ is bipartite, and this step is indeed significant - observe that for $K_1 = K_2 = -1$ in the XXZ model (known as the antiferromagnet), Figure 1.5 shows there is no phase transition, whereas [33] shows that on $\mathbb{Z}^d$, $d \geq 3$, there is one. This makes intuitive sense, since for the antiferromagent, adjacent spins want to be anti-aligned, and it is only possible for every pair of adjacent spins in the graph to be anti-aligned if the graph is bipartite. The models studied in Chapter 6 are a general family, but for example when $\theta = 3$ they are the two special cases of the bilinear-biquadratic model (1.9) where $J_1 = J_2$, and $J_1 = 0$. Again, we give free energy results, Theorems 6.1.1, 6.1.2 and 6.1.9, and then (for certain values of the parameters involved) find formulae for the points of phase transition. See Propositions 6.1.3, 6.1.4 and 6.1.5. We also give further features of the models, which back up the presence of phase transitions and help describe the phases of the models. See Theorems 6.1.7 and 6.1.8.

The key technical application of the Brauer algebra $\mathbb{B}_{n,\theta}$ and walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ to the results of Chapters 5 and 6 is the following. Regardless of the underlying graph, the operator $e^{-\beta H}$ of the model in question (eg. (1.8) or (1.9)) can be written as the action of a certain element of one of the algebras $\mathbb{C}S_n$, $\mathbb{B}_{n,\theta}$ or $\mathbb{B}_{n,m,\theta}$. The free energy, the key quantity in both papers, is given by the formula $\lim_{|\mathcal{V}|\to\infty} \frac{1}{|\mathcal{V}|} \log \operatorname{Tr}\left[e^{-\beta H}\right]$. Here $\mathcal{V}$ is the set of vertices of the underlying graph. It is not hard to see that an eigendecomposition of the operator $e^{-\beta H}$ on tensor space $V^{\otimes n}$ is very useful for computing the free energy. Schur-Weyl duality, as noted above, gives a decomposition of the action of the algebras on $V^{\otimes n}$ into irreducibles, and so a block decomposition of $e^{-\beta H}$. When the underlying graph is the complete graph or complete bipartite graph, $e^{-\beta H}$ is the action of a central (or nearly central) element of the algebra, so it acts on irreducibles as scalars. This means the blocks of the decomposition of $e^{-\beta H}$ are scalars, so we have an eigendecomposition of $e^{-\beta H}$.

For generic parameters considered, the block decomposition is slightly more nuanced than described above, and requires the careful restriction of irreducibles of one of the algebras to another one: for example, in Chapter 5, from the Brauer algebra $\mathbb{B}_{n,\theta}$ to the symmetric group $\mathbb{C}S_n$. For some cases these restrictions are well-studied, for some we work out complete, explicit formulae, and for some they are ill-understood (in particular $\mathbb{B}_{n,\theta}$ to $\mathbb{C}S_n$), and we can give formulae in some cases. In fact the main reason this problem is difficult is the non-semisimplicity of $\mathbb{B}_{n,\theta}$ for certain values of θ, as discussed above. See Section 5.7 and Sections 6.2.1 and 6.2.2.

This approach of applying the representation theory of the algebra in question to diagonalise the action of a central element is well-trodden, although has mostly been done with group algebras in the past. The book of Diaconis [29] collects several examples of such applications to problems in probability. Perhaps the most famous therein is the application to shuffling cards via random transpositions [30], where Diaconis and Shahshahani obtain the mixing time (the time when the cards become "well-shuffled") by diagonalising the action of a central element of $\mathbb{C}S_n$. A key insight in their work was interpreting the shuffling as a random walk on the symmetric group.

There are strong links between random walks on the symmetric group and quantum spin systems. Powers [83] showed that (the $\theta = 2$ version of) the quantum Heisenberg model has a probabilistic representation as a continuous time version of the random walk studied by Diaconis and Shahshahani, known as the interchange process, with an added weight depending on cycle lengths. Tóth [97] used this probabilistic representation to give a bound on the free energy of the model. Since then several models, including the Heisenberg model for general θ and (for certain values of the parameters) those studied in Chapters 5 and 6 have been shown to have similar representations as random processes, and results about the models have given rise to results about the processes, and vice-versa. See Ueltschi [101] and Nachtergaele [77] for details.

The work of Chapters 5 and 6 follows a line of work approaching the interchange process and the Heisenberg model with the representation theory of $\mathbb{C}S_n$. Alon and Kozma [3] estimated the number of cycles of length k in the unweighted interchange process, on any graph. Berestycki and Kozma [9] gave an exact formula for the same on the complete graph, and studied the phase transitions present. In [4] Alon and Kozma gave a formula for the magnetisation of the weighted process (equivalent to the $\theta = 2$ Heisenberg model) on any graph, which simplifies greatly in the mean-field. Chapters 5 and 6 are most directly inspired by the paper of Björnberg [13], who showed a phase transition in the weighted interchange process.

Looking forward, one would like to build on this work in a few ways. Certainly transferring some of the results to the models on the lattice is desirable. For $\theta > 3$ we only have a partial phase diagram for the model of Chapter 5. Completing this picture requires obtaining more information about the restriction of irreducibles from $\mathbb{B}_{n,\theta}$ to $\mathbb{C}S_n$, when $\theta > 3$. One would also like to apply similar methods to obtain results similar to those of [3] on the interchange processes which are associated with the model of Chapter 5. Remark 5.2.11 details why there are nuances in obtaining certain such results. In Chapter 6 one of the pleasantly surprising results is that two noticeably different models (6.5) and (6.10) have exactly the same free energy. This is intriguing, and warrants further investigation - perhaps the associated random processes have analogous similarities?

Chapter 2

Representation Theory

2.1 Representation theory

2.1.1 Representation theory of finite-dimensional algebras

In this section we give an introduction to the representation theory of finite dimensional algebras over $\mathbb{C}$, giving results that we will use in later sections. We will not prove results in this section. We follow mainly Etingof et al. [34], with some results from Fulton and Harris [42], some from Sections 9 and 10 of Curtis and Reiner [26], and some from Lecture notes of Fayers [35].

We follow Chapter 2 of [34] for this section on general representation theory of algebras, unless stated differently. A (unital, associative) algebra over $\mathbb{C}$ is a vector space A with a multiplication which is associative, distributive, has an identity, and satisfies $(ca_1)(a_2) = (a_1)(ca_2) = c(a_1a_2)$ for all $a_1, a_2 \in A$, and $c \in \mathbb{C}$. The identity, which we denote by 1, is unique. The centre $Z(A)$ is the set of elements of A which commute with every element of A.

A left (resp. right) ideal of A is a subspace I of A such that ab (resp. ba) lie in I for all $b \in I$, $a \in A$. A two-sided ideal, which we will usually just call an ideal, is a left ideal which is also a right ideal. For A_1, A_2 two algebras, an algebra homomorphism from A_1 to A_2 is a linear map $\phi : A_1 \to A_2$ which sends identity to identity and satisfies $\phi(ab) = \phi(a)\phi(b)$ for all $a, b \in A_1$.

A (left) representation (or module) of A is a pair (ρ, M), M a complex vector space, $\rho : A \to \text{End}(M)$ an algebra homomorphism into the algebra of endomorphisms of M (linear maps from M to M). We will use the terms representation and module interchangeably, and often we will denote a module by only its vector space or its homomorphism. All the representations we will consider will be finite-dimensional, so often we will assume without stating that modules are finite-dimensional. For a module (ρ, M), and for $a \in A$, $v \in M$, we will often denote $\rho(a)v$ as $a \cdot v$ or av. A right module is the same as a left module, except that $\rho(ab) = \rho(b) \circ \rho(a)$, instead of $\rho(ab) = \rho(a) \circ \rho(b)$ in the case of a left module. We write its action as $\rho(a)(v) = va$. For two algebras A and B, an $A - B$ bimodule is a vector space M which is a left A-module and a right B-module, such that $(av)b = a(vb)$ for all $a \in A$, $b \in B$, $v \in M$.

A submodule (or subrepresentation) N of a module M is a subspace N of M which is sent to itself by any $a \in A$. A irreducible (or simple) module is a module with no non-zero proper submodules. Given two representations (ρ_1, M_1), (ρ_2, M_2) of A, the space $M_1 \oplus M_2$ is a representation of A, with action $\rho_1 \oplus \rho_2$. A non-zero module is indecomposable if it cannot be written as a direct sum of two non-zero submodules (and decomposable if it can be written as such). The regular module ${}_AA$ of A is the space A itself, with action given by left multiplication $a_1 \cdot a_2 := a_1 a_2$.

Lemma 2.1.1 (Schur's lemma, Proposition 1.16 of [34])**.** *Let S and T be two simple modules of an algebra A. If $\phi : S \to T$ is a module homomorphism, then $\phi = 0$ or ϕ is an isomorphism.*

Given two representations (ρ_1, M_1), (ρ_2, M_2) of two different algebras A_1 and A_2 respectively, we call $M_1 \boxtimes M_2$ the module of the algebra $A_1 \otimes A_2$ with vector space $M_1 \otimes M_2$, and action $\rho_1 \boxtimes \rho_2(a_1 \otimes a_2)(v_1 \otimes v_2) = (a_1 v_1) \otimes (a_2 v_2)$. The module $M_1 \boxtimes M_2$ is irreducible if M_1 and M_2 are irreducible, and all irreducibles of $A_1 \otimes A_2$ are of the form $M_1 \boxtimes M_2$, for unique M_1 and M_2. Note we use the box tensor product to differentiate this representation from the usual tensor products of representations of groups - see below.

We will occasionally use another tensor product. Let A, B be two algebras, and M an $A - B$ bimodule, N a left B module. Then we define the vector space $M \otimes_B N$ as the vector space $M \otimes N$, quotiented by the relations $mb \otimes n = m \otimes bn$, for all $m \in M$, $n \in N$, $b \in B$. Then $M \otimes_B N$ is an A module, with action $a \cdot (m \otimes n) = (am) \otimes n$.

A module M is a semisimple module if it is a direct sum of simple modules; this is equivalent to the statement that for any submodule N of M there is another submodule N' of M such that $M = N \oplus N'$, $N \cap N' = 0$. One finds that any submodule or quotient module of a semisimple module is also semisimple.

If a module M is semisimple and $M = \sum_{T \in \mathcal{T}} T$ is a sum of a collection of simple submodules, then every simple submodule of M is isomorphic to one of the $T \in \mathcal{T}$.

Theorem 2.1.2 (Artin-Wedderburn, Proposition 2.16 of [34])**.** *The following are equivalent:*

1. *An algebra A is semisimple if it is isomorphic to a direct sum of matrix algebras: $A \cong \bigoplus_{i=1}^{n} Mat_{r_i}(\mathbb{C})$;*

2. *The regular representation ${}_AA$ of A is semisimple.*

If these properties hold we call the algebra A a semisimple algebra.

Theorem 2.1.3 (Density Theorem, Theorem 2.5 of [34])**.** *Let A be an algebra, and $(\rho_1, V_1), \dots, (\rho_k, V_k)$ be pairwise non-isomorphic finite-dimensional simple modules of A. Then the map $\rho_1 \oplus \dots \oplus \rho_k : A \to \bigoplus_{i=1}^{k} End_{\mathbb{C}} V_i$ is surjective.*

Similarly, an algebra is called a simple algebra if it has no non-zero proper ideals.

Lemma 2.1.4. *Let A be an algebra. The following are equivalent:*

1. *The algebra A is simple;*

2. *The algebra A is isomorphic to a matrix algebra: $A \cong Mat_r(\mathbb{C})$;*

3. *There is some simple module S of A such that the regular representation ${}_AA$ is isomorphic to $S^{\oplus n}$ for some $n \in \mathbb{N}$.*

We will encounter some representations (and algebras) which are not semisimple. Let us define the radical of an A module M, $\operatorname{rad} M$ as the intersection of all maximal submodules of M. One can prove that $\operatorname{rad} M = 0$ if and only if M is semisimple. The quotient $M / \operatorname{rad} M$, called the head of M, is semisimple - it is essentially the "semisimple part" of M.

We follow Section 25 of [26] for the following paragraphs on idempotents. An element $e \in A$ is an idempotent if $e^2 = e$. Two idempotents e, f are orthogonal if $ef = fe = 0$. An idempotent is primitive if it is non-zero and cannot be written as the sum of two non-zero orthogonal idempotents. A primitive decomposition of an idempotent e is a finite set of pairwise orthogonal idempotents which sum to e. Let us note: 0 and 1 are idempotents. If e is an idempotent, $1 - e$ is an idempotent orthogonal to e. If e, f are orthogonal idempotents, then $e + f$ is an idempotent. If e, f are idempotents which commute, then ef is an idempotent. The only invertible idempotent is 1. A central idempotent is an idempotent that lies in the centre $Z(A)$ of A.

If A is finite-dimensional, every idempotent e has a primitive decomposition, and eAe is a subalgebra of A under the same operations as A, with identity element e. Then the following are equivalent:

1. e is a primitive idempotent in A;
2. e is a primitive idempotent in eAe;
3. e is the only non-zero idempotent in eAe.

One can use idempotents to decompose an algebra. An algebra A is indecomposable if 1 is a primitive central idempotent (that is, 1 is a primitive idempotent in $Z(A)$). If A is an algebra and E is a primitive decomposition of 1 in $Z(A)$, then

1. E is the unique primitive decomposition of 1 in $Z(A)$;
2. E consists of all the primitive central idempotents in A;
3. A has only finitely many central idempotents.

A primitive decomposition of 1 in $Z(A)$ gives a unique decomposition

$$A = Ae_1 \oplus \cdots \oplus Ae_k, \tag{2.1}$$

into indecomposable algebras. Similarly, it gives a decomposition of a module M into $M = e_1M \oplus \cdots \oplus e_kM$, where e_iM is annihilated by each Ae_j, $i \neq j$. Let us note that sometimes we do not need a primitive decomposition of 1 in $Z(A)$ to give certain useful decompositions. If $\{e_1, \ldots, e_k\}$ satisfy the weaker condition that they are primitive idempotents in A summing to 1 which are "half-orthogonal", then

$${}_AA = Ae_1 \oplus \cdots \oplus Ae_k, \tag{2.2}$$

decomposes ${}_AA$ into indecomposable modules. By half-orthogonal we mean the following: for all $1 \le i < j \le k$, we have $e_i e_j = 0$ (and not necessarily $e_j e_i = 0$). We will find such a decomposition of the symmetric group algebra in Lemma 2.1.7, where the idempotents are the Young symmetrisers, and we will use it in Chapter 3.

We follow Chapters 3 and 4 of [34] for the following results on representations of groups. A representation of a group G is a pair (ρ, M), M a complex vector space, $\rho : G \to GL(M)$ a group homomorphism into the group of invertible linear maps on M. The definitions of subrepresentations, simple/irreducible, semisimple and indecomposable representations, direct sums and box tensor products of representations are all analogous to the equivalent definitions for modules of algebras. In addition, we define the usual tensor product of representations of a group, $M_1 \otimes M_2$: For M_1, M_2 two representations of a group G, let $g \cdot (m_1 \otimes m_2) = gm_1 \otimes gm_2$ for all $g \in G$, $m_1 \in M_1$, $m_2 \in M_2$, and extend linearly. The trivial representation always exists: it is the module $(\rho, \mathbb{C})$ with $\rho(a) = 1$ for all $a \in A$.

The group algebra $\mathbb{C}G$ of a group G is the vector space of formal finite sums $\sum_{g \in G} a_g g$ (with all but finitely many of the scalars a_g equal to zero); $\mathbb{C}G$ is the vector space with basis G. The multiplication in $\mathbb{C}G$ is the multiplication in G linearly extended. Group algebras of finite groups are semisimple algebras.

A representation (ρ, M) of G is equivalent to a module (ρ', M) of $\mathbb{C}G$ by linear extension $\rho'(\sum_{g \in G} a_g g) \coloneqq \sum_{g \in G} a_g \rho(g)$, or in reverse, simply restricting ρ' to G; we will refer to either as simply representations or modules of G. The finite dimensional irreducible representations of G are therefore in bijection with the finite-dimensional irreducible modules of $\mathbb{C}G$.

Given any representation (ρ, M) of a group G, the dual representation is the pair (ρ^*, M^*), where M^* is the dual vector space of M, with $\rho^*(g) = (\rho(g^{-1}))^*$, the adjoint of the map $\rho(g^{-1})$. Explicitly, for $w \in M^*$, $v \in M$, $\rho^*(g)(w)(v) = (\rho(g^{-1}))^*(w)(v) = w(\rho(g^{-1})(v))$.

The character of a representation (ρ, M) of an algebra A is the function $\chi : A \to \mathbb{C}$ as $\chi(a) = \mathrm{Tr}(\rho(a))$, the trace of the map $\rho(a)$. We say χ is irreducible (resp. indecomposable, semisimple) if its associated representation is. Calculations yield that if χ_M, χ_N are the characters of two modules M and N of an algebra A, then $\chi_{M \oplus N} = \chi_M + \chi_N$, and for A a group algebra, $\chi_{M \otimes N} = \chi_M \chi_N$ and $\chi_{M^*} = \overline{\chi_M}$.

The conjugacy classes of a group are the equivalence classes of the conjugacy relation: $g, h \in G$ are conjugate if there is some $\pi \in G$ with $\pi^{-1} g \pi = h$. For G finite, the centre $Z(\mathbb{C}G)$ is the span of the conjugacy class sums $\sum_{g \in C} g$, for conjugacy classes C. The central elements of any algebra act as scalars on the irreducible representations. For G finite, the irreducible characters of the group algebra form a basis for the space of class functions (functions constant on conjugacy classes); hence the number of irreducible representations of the group algebra of a finite group is the number of conjugacy classes. Two representations of a semisimple algebra are isomorphic if and only if they have the same character.

For G finite, there is a natural inner product on the space of class functions, $\langle \alpha, \beta \rangle = \frac{1}{|G|} \sum_{g \in G} \alpha(g) \overline{\beta(g)}$. The irreducible representations of G are orthonormal with respect to this inner product. Let us note that when G is a classical group (defined later in this section), there is an analogous inner product (see Proposition 1.12 of [72]):

$$\langle \alpha, \beta \rangle = \int_G \alpha(g)\overline{\beta(g)} d\mu(g), \tag{2.3}$$

where μ is the Haar measure on the group. The irreducible (finite dimensional) representations of G are orthonormal with respect to this inner product.

We follow Chapter 2 of [34] once again. A Lie algebra $\mathfrak{g}$ is a vector space with a multiplication $[\cdot,\cdot]$ which is bilinear, satisfying $[x, x] = 0$ for all $x \in \mathfrak{g}$, and the Jacobi identity, $[x,[y,z]] + [z,[x,y]] + [y,[z,x]] = 0$ for all $x, y, z \in \mathfrak{g}$. Similarly to the algebra case, we can define Lie sub-algebras, direct sums and tensor products of Lie algebras, and Lie algebra homomorphisms. For a vector space V, the general linear Lie algebra on V, $\mathfrak{gl}(V)$, is the Lie algebra of linear maps from V to V, with multiplication given by the Lie bracket $[x, y] = xy - yx$. A representation of a Lie algebra is a pair (ϕ, V) with ϕ a Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{gl}(V)$. Similarly to the algebra case, we can define subrepresentations, direct sums of representations, irreducible/simple and semisimple representations. Let us note that the tensor product of two representations of a Lie algebra $\mathfrak{g}$, V_1 and V_2, is the vector space $V_1 \otimes V_2$, with the action $x \cdot (v_1 \otimes v_2) = (x \cdot v_1) \otimes v_2 + v_1 \otimes (x \cdot v_2)$.

The irreducible representations of all of the major groups and algebras in this thesis are indexed by some set of partitions or tuples. For a given group or algebra G, we will denote by ψ_ρ^G the irreducible of G corresponding to the partition or tuple ρ. We will also denote its character by χ_ρ^G and its dimension by d_ρ^G.

2.1.2 The symmetric group

We follow James [57] for this section. For $n \in \mathbb{N}$, the symmetric group S_n is the group of bijections of the set $\mathcal{N} = \{1, \ldots, n\}$, under composition. The symmetric group is a fundamental feature of representation theory. One of its many interesting features is its relationship with the general linear group, called Schur-Weyl duality, which is the focus of Chapter 3. Let us recall a number of results on the symmetric group and its representations.

We often call elements of S_n permutations. We often write $\sigma \in S_n$ in disjoint cycle notation: for example $\sigma = (1,3,4)(2,5)$ is the permutation exchanging 2 and 5, and sending 1 to 3, 3 to 4, and 4 to 1, while fixing all other numbers. The cycle type of σ is the tuple given by the lengths of its cycles; the cycle type of $\sigma = (1,3,4)(2,5) \in S_6$ is $(3,2,1)$, the 1 from fixing the 6. We often drop any ones when describing cycle types, so we would write our example as $(3,2)$. The elements (i,j), the elements of cycle type (2), or $(2, 1^{n-2})$, called the transpositions, generate the symmetric group. Cycle type determines conjugacy classes, ie. $\sigma, \tau \in S_n$ are conjugate if and only if σ and τ have the same cycle type.

A partition $\lambda = (\lambda_1, \ldots, \lambda_s)$ of a natural number n is an ordered list of non-negative integers which are non-increasing and which sum to n. We often denote partitions by the Greek letters $\lambda, \rho, \mu, \pi, \xi$, etc. By the working above, the conjugacy classes, and therefore also the irreducible representations of S_n, are in bijection with partitions of n, that is, tuples of non-negative, non-increasing integers $\lambda = (\lambda_1, \ldots, \lambda_r)$ with $\sum_{i=1}^r \lambda_i = n$. We write

$\lambda \vdash n$ to denote λ being a partition of n.

We can construct the irreducibles of S_n explicitly; they are called the Specht modules, following [57]. We need some notation. For a partition $\lambda \vdash n$, the Young diagram of λ is the array of boxes with λ_i boxes in the i^{th} row, arranged with the first box of each row above the first box of the next. When it is unambiguous, we will denote the Young diagram of λ simply by λ. We will sometimes write $|\lambda|$ for the number of boxes in λ, that is, $|\lambda| = n$ is just another way of writing $\lambda \vdash n$. See Figure 2.1 for an illustration of the Young diagrams of the partitions $(5,5,3,1)$, $(4,1,1)$ respectively. For a partition λ, the conjugate partition λ^T is the partition with Young diagram obtained by transposing the diagram of λ (so λ_i^T is the length of the i^{th} column of λ). A tableau of size n is

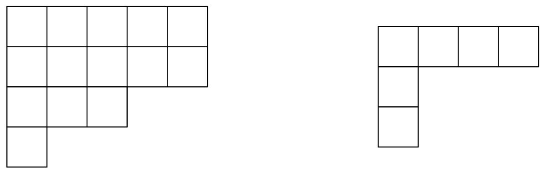


Figure 2.1: The Young diagrams of the partitions $(5,5,3,1)$ and $(4,1,1)$.

a Young diagram of some shape $\lambda \vdash n$ with each box filled with a unique number from $\mathcal{N} = \{1,\dots,n\}$. We say such a tableau has shape λ. We call the set of tableaux of size n $\mathcal{T}(\mathcal{N})$, and we sometimes write $|\tau| = |\lambda|$, the number of boxes in τ and λ. Sometimes we will require the numbers in the tableaux to be a subset U of $\mathcal{N}$; in this case we call the set of such tableaux $\mathcal{T}(U)$. Note $|\tau| = |U|$ for all $\tau \in \mathcal{T}(U)$. We label the set of tableaux with shape λ by $\mathcal{T}_\lambda(U)$. We say a standard tableau is a tableau with its entries strictly increasing along rows and down columns, and we denote by $\mathcal{ST}(U)$ the set of standard tableaux with entries in U, and by $\mathcal{ST}_\lambda(U)$ the set of standard tableaux of shape $\lambda \vdash |U|$, with entries from U.

Let us now give a first definition of the Specht modules. The symmetric group acts on tableaux by permuting the entries. Say $\tau, \tau' \in \mathcal{T}(\mathcal{N})$ are related if one can permute the entries of each row of τ to get τ'. The symmetric group acts on the equivalence classes of this relation, by $\sigma \cdot \{\tau\} = \{\sigma\tau\}$. The vector space spanned by these classes is therefore a module of S_n (often denoted M^τ). The subrepresentation of M^τ spanned by elements $(\sum_{h\in\mathcal{C}(\tau)} \operatorname{sgn}(h)h)\cdot\{\tau\}$, where τ has shape λ, is called the Specht module, which we denote $\psi_\lambda^{S_n}$.

Theorem 2.1.5 (Theorem 4.12 of [57])**.** *For each $\lambda \vdash n$, the space $\psi_\lambda^{S_n}$ is an irreducible representation of S_n, and $\psi_\lambda^{S_n} \cong \psi_\mu^{S_n}$ if and only if $\lambda = \mu$. Hence the Specht modules $\psi_\lambda^{S_n}$, for $\lambda \vdash n$ are a complete set of pairwise non-isomorphic irreducible representations of S_n.*

We write $\chi_\lambda^{S_n}$ for the character of the Specht module $\psi_\lambda^{S_n}$, and $d_\lambda^{S_n}$ for its dimension. For a subset $U \subset \mathcal{N} = \{1,\dots,n\}$, and $\tau \in \mathcal{T}(U)$, define the element

$$z'_\tau = \sum_{g\in\mathcal{R}(\tau)} \sum_{h\in\mathcal{C}(\tau)} \operatorname{sgn}(h)hg \in \operatorname{Sym}(U) \le S_n, \tag{2.4}$$

where $\mathcal{R}(\tau)$ is the set of permutations in $\operatorname{Sym}(U)$ which preserve the rows of τ (as sets), and similar for $\mathcal{C}(\tau)$ and columns.

Lemma 2.1.6. *The element z'_τ satisfies $(z'_\tau)^2 = c_\tau z'_\tau$, for some $c_\tau \in \mathbb{C}$.*

Hence the element $z_\tau := \frac{1}{c_\tau} z'_\tau$ is an idempotent, called the Young symmetriser with tableau τ.

Lemma 2.1.7 (Theorem 4.3 of [42]). *If $\lambda \vdash n$, $\tau \in \mathcal{ST}_\lambda(\mathcal{N})$, one can rewrite the Specht module $\psi_\lambda^{S_n}$ as the span of the elements $z_\tau \tau$, τ standard, in the space spanned by all tableaux of shape λ. Equivalently, $\psi_\lambda^{S_n}$ can be realised as the subspace given by $\mathbb{C}S_n z_\tau$ of the regular representation ${}_{\mathbb{C}S_n}\mathbb{C}S_n$.*

Notice that the dimension $d_\lambda^{S_n}$ of the Specht module is therefore the number of standard tableaux of shape λ. The Young symmetrisers are a set of primitive idempotents, but are not in general orthogonal. They are, however, "half-orthogonal", in the following sense. For $U \subset \mathcal{N}$, introduce a total order $<$ on the set $\mathcal{ST}_\lambda(U)$ with $\tau < \tau'$ if when reading the entries left to right along consecutive rows, the first entries m of τ and m' of τ' in the same box with $m \neq m'$ have $m < m'$.

Lemma 2.1.8 (Proposition 2.3 of [43]). *If $\tau, \tau' \in \mathcal{ST}_\lambda(U)$ with $\tau < \tau'$ with respect to the order described above, then $z_\tau z_{\tau'} = 0$. (And it is not true in general that $z_{\tau'} z_\tau = 0$).*

Lemma 2.1.9 (Theorem 4.3 of [42]). *We have a decomposition $\mathbb{C}S_n = \bigoplus_{\tau \in \mathcal{ST}(\mathcal{N})} \mathbb{C}S_n z_\tau$ of the group algebra $\mathbb{C}S_n$ into minimal left ideals, and therefore a decomposition of the regular representation ${}_{\mathbb{C}S_n}\mathbb{C}S_n = \bigoplus_{\tau \in \mathcal{ST}(\mathcal{N})} \mathbb{C}S_n z_\tau$, where each $\mathbb{C}S_n z_\tau$ is a copy of the irreducible representation $\psi_\lambda^{S_n}$ of S_n, for τ shape λ.*

We finish this section with a result which describes how a particular central element of $\mathbb{C}S_n$ acts on irreducibles. For $\lambda \vdash n$, we label by $\mathrm{ct}(\lambda)$ the sum of contents of the boxes of the Young diagram of λ, where the content of the box in row i and column j is given by $j - i$.

Lemma 2.1.10. *The sum of all transpositions, $\sum_{1 \le i < j \le n} (i, j) \in \mathbb{C}S_n$ acts on the irreducible $\psi_\lambda^{S_n}$ as the scalar $\mathrm{ct}(\lambda)$.*

2.1.3 The Brauer algebra

We follow several references for this section, but mainly Cox et al. [24]. The Brauer algebra was introduced by Brauer [19] in 1937 as having the same relationship (Schur-Weyl duality) with the orthogonal group as the symmetric group does with the general linear group. See Chapter 3 for a full description of Schur-Weyl duality. The Brauer algebra has been studied widely in its own right, particularly since the late 1980s. Broadly, its representation theory (over the complex numbers) has been shown to be related to, but more nuanced than, that of the symmetric group. Although the Brauer algebra has a distinguished basis (see definitions below), it is not a group algebra, so there are several results about group algebras which do not hold in general for the Brauer algebra. A key such result is that over the complex numbers the algebra is not always semisimple.

In the semisimple case the irreducible representations and characters were found by Brown [20] and Ram [84], respectively. The non-semisimple case was studied in a series of papers by Hanlon and Wales [50], [51], [52] and [53]. The criterion for semisimplicity was worked on in these papers and by Wenzl [103], Doran, Hanlon and Wales [31] and was settled over an arbitrary field by Rui [87]. The blocks (essential information on the irreducible representations) of the algebras in the non-semisimple cases were determined by Cox, de Visscher and Martin [24].

Let $\theta \in \mathbb{C}$. The Brauer algebra $\mathbb{B}_{n,\theta}$ is the (formal) complex span of the set of pairings of $2n$ vertices. We think of pairings as graphs, which we will call diagrams, with each vertex having degree exactly 1. We arrange the vertices in two horizontal rows, labelling the upper row (the northern vertices) $1^{\mathsf{N}}, 2^{\mathsf{N}}, \ldots, n^{\mathsf{N}}$, and the lower (southern) $1^{\mathsf{S}}, \ldots, n^{\mathsf{S}}$. We call an edge connecting two northern vertices (or two southern) a bar, and an edge connecting a northern vertex and a southern vertex a NS-path. The number of northern bars in a diagram is the same as the number of southern bars, and we refer to either as simply the number of bars of the diagram.

Multiplication of two diagrams is given by concatenation. If b_1, b_2 are two diagrams, we align the northern vertices of b_1 with the southern of b_2, and the result is obtained by removing these middle vertices (which produces a new diagram), and then multiplying the result by $\theta^{l(b_1,b_2)}$, where $l(b_1,b_2)$ is the number of loops in the concatenation. See Figure 2.2. This defines $\mathbb{B}_{n,\theta}$ as an algebra. One can readily check that the dimension of $\mathbb{B}_{n,\theta}$ is $2n!! = (2n-1)(2n-3)\cdots 5 \cdot 3 \cdot 1$.

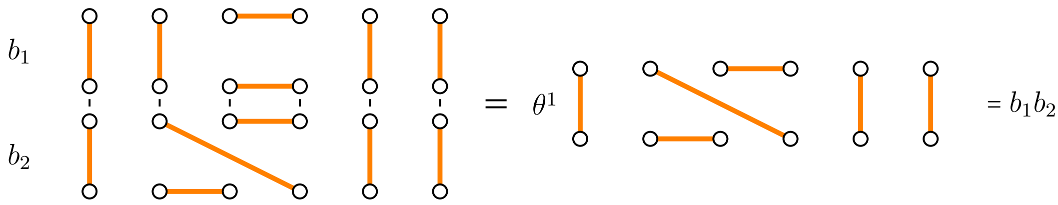


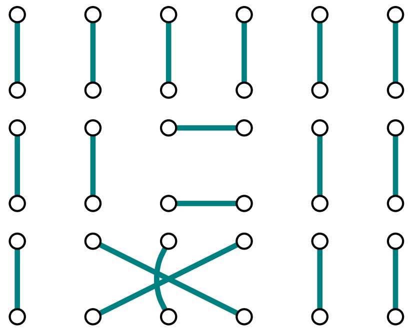
Figure 2.2: Two diagrams b_1 and b_2 (left), and their product (right). The concatenation contains one loop, so we multiply the concatenation (with middle vertices removed) by θ^1.

We call the set of diagrams (the basis of $\mathbb{B}_{n,\theta}$) B_n. Note that diagrams with no bars are exactly permutations, where $\sigma \in S_n$ is represented by the diagram where i^{S} is connected to $\sigma(i)^{\mathsf{N}}$, so $S_n \subset B_n$. Moreover the multiplication defined above reduces to multiplication in S_n, so $\mathbb{C}S_n$ is a subalgebra of $\mathbb{B}_{n,\theta}$. We write id for the identity - its diagram has all its edges vertical. We denote the transposition S_n swapping i and j by (i,j), and we write $(\overline{i,j})$ for the diagram with i^{N} connected to j^{N}, and i^{S} connected to j^{S}, and all other edges vertical. See Figure 2.3. Note that just as the transpositions (i,j) generate the symmetric group, the Brauer algebra is generated by the transpositions and the elements $(\overline{i,j})$.

Occasionally we will write a diagram $b \in B_n$ in what we call edge notation, that is, as a list of the pairs of vertices which are connected in the graph. In general, this looks like

$$b = \prod_{j=1}^{n} (b_j, b'_j), \tag{2.5}$$

where $b_j, b'_j \in \{t^\xi \ : \ \xi = \mathsf{N}, \mathsf{S},\ 1 \le t \le n\}$; this denotes the diagram with the pairs of vertices (b_j, b'_j) connected to each other, for all $j = 1, \ldots, n$. Note that for a diagram $b \in B_n$ there are many such products which can represent it, in fact, any permutation of the n pairs (b_j, b'_j) and swapping the positions of any of the b_j with b'_j gives the same diagram b. For example, the identity can be written as $\prod_{j=1}^{n}(j^{\mathsf{N}}, j^{\mathsf{S}})$, the transposition can be written as $(i,j) = (i^{\mathsf{N}}, j^{\mathsf{S}})(j^{\mathsf{N}}, i^{\mathsf{S}})\prod_{k\neq i,j}(k^{\mathsf{N}}, k^{\mathsf{S}})$, and $(\overline{i,j}) = (i^{\mathsf{N}}, j^{\mathsf{N}})(i^{\mathsf{S}}, j^{\mathsf{S}})\prod_{k\neq i,j}(k^{\mathsf{N}}, k^{\mathsf{S}})$.



$= \mathrm{id} \in S_6$

$= (\overline{3,4})$

$= (2,4) \in S_n$

Figure 2.3: The identity element, the element $(\overline{3,4})$, and the transposition $(2,4) \in S_6$, all lying in B_6.

Let us turn to representations. As noted at the beginning of this section, the Brauer algebra is not always semisimple. To be precise, Wenzl [103] showed it is semisimple when the multiplicative parameter θ is not an integer, and for $\theta \in \mathbb{N}$, Brown [20] showed it is semisimple if and only if $\theta \geq n-1$. The negative integers are less straightforward, although it always holds that $\mathbb{B}_{n,\theta}$ is semisimple for $\theta < -2n+4$ - see Rui [87], who gave a criterion for the Brauer algebra to be semisimple over any field. Brown [20] gave a description of the irreducible modules in the semisimple case, and Ram [84] gave a description of their characters. In this thesis we will need results in all cases (both semisimple and non-semisimple). The Brauer algebra is what is known as a cellular algebra [45], and more specifically, a tower of recollement [22]. While we will not explore this theory, we will note that it gives an explicit description of a set of "standard" indecomposable modules, called the cell modules, and the irreducible modules of a such an algebra. In particular, generically (that is, when $\mathbb{B}_{n,\theta}$ is semisimple), the cell modules are the irreducible modules, although when $\mathbb{B}_{n,\theta}$ is not semisimple, they are no longer all irreducible, and the irreducible is then its head (the quotient by its radical). Let us describe the cell modules.

Notice that if $b \in B_n$ is a diagram with k bars, then for any $a \in B_n$, ab and ba are both (scalar multiples of) diagrams with at least k bars. Hence, if we define $\mathbb{B}_{n,\theta}^{k}$ to be the subspace of $\mathbb{B}_{n,\theta}$ spanned by diagrams in B_n with at least k bars, then $\mathbb{B}_{n,\theta}^{k}$ is an ideal of $\mathbb{B}_{n,\theta}$. In fact, we have a descending chain of ideals:

$$\mathbb{B}_{n,\theta} = \mathbb{B}_{n,\theta}^{0} \supset \mathbb{B}_{n,\theta}^{1} \supset \cdots \supset \mathbb{B}_{n,\theta}^{\lfloor \frac{n}{2} \rfloor}. \tag{2.6}$$

Let $\xi_k = \theta^{-k}\prod_{i=1}^{k}(\overline{i, n+1-i})$ (ξ_k is a idempotent). Notice that the ideal $\mathbb{B}_{n,\theta}^{k}$ can be written as $\mathbb{B}_{n,\theta}\xi_k\mathbb{B}_{n,\theta}$. Let $B_n\langle k\rangle$ be the set of diagrams in B_n with exactly k bars. The quotient $\mathbb{A}_{n,\theta}^{k} := \mathbb{B}_{n,\theta}^{k}/\mathbb{B}_{n,\theta}^{k+1}$ is an algebra (we define $\mathbb{A}_{n,\theta}^{\lfloor \frac{n}{2} \rfloor} := \mathbb{B}_{n,\theta}^{\lfloor \frac{n}{2} \rfloor}$). One can easily picture $\mathbb{A}_{n,\theta}^{k}$ as the vector space with basis $B_n\langle k\rangle$, with multiplication the same as $\mathbb{B}_{n,\theta}$, except that if two diagrams multiply to give a diagram with more than k bars, then the result is set to zero. As a vector space, we have the decomposition:

$$\mathbb{B}_{n,\theta} = \mathbb{A}_{n,\theta}^{0} \oplus \mathbb{A}_{n,\theta}^{1} \oplus \cdots \oplus \mathbb{A}_{n,\theta}^{\lfloor \frac{n}{2} \rfloor}. \tag{2.7}$$

Note $\mathbb{A}_{n,\theta}^{0}$ is just the symmetric group algebra $\mathbb{C}S_n$. Each $\mathbb{A}_{n,\theta}^{k}$ is also a module for $\mathbb{B}_{n,\theta}$, by left multiplication j it is a quotient of the submodule $\mathbb{B}^k_{n,\theta}$ of the regular representation). Notice that the subspace of $\mathbb{A}^k_{n,\theta}$ spanned by diagrams with a fixed set of k southern bars is a submodule of $\mathbb{A}^k_{n,\theta}$. To decompose the representation $\mathbb{A}^k_{n,\theta}$ into indecomposable modules, we use this observation, along with the Specht modules of the symmetric group.

Any diagram $b \in B_n\langle k\rangle$ (i.e. with k bars), is determined by its k northern bars, k southern bars, and $n-2k$ NS-paths. For fixed northern and southern bars, the possible $n-2k$ NS-paths can be bijected with S_{n-2k} by deleting the bars (along with their vertices), and shifting the remaining vertices together. Hence each diagram $b \in B_n\langle k\rangle$ can we written uniquely as $b = \sigma \otimes (a_{\mathsf{N}}, a_{\mathsf{S}})$, for $\sigma \in S_{n-2k}$, $(a_{\mathsf{N}}, a_{\mathsf{S}}) \in \mathfrak{a}^{k,\mathsf{N}}_n \times \mathfrak{a}^{k,\mathsf{S}}_n$, where $\mathfrak{a}^{k,\mathsf{N}}_n$ is the set of choices of k northern bars and $\mathfrak{a}^{k,\mathsf{S}}_n$ the set of choices of k southern bars on the $2n$ vertices. As a vector space, we have $\mathbb{A}^k_{n,\theta} = \mathbb{C}S_{n-2k} \otimes \mathbb{C}(\mathfrak{a}^{k,\mathsf{N}}_n \times \mathfrak{a}^{k,\mathsf{S}}_n)$. Our observation above can now be written as the fact that $\mathbb{C}S_{n-2k} \otimes \mathbb{C}(\mathfrak{a}^{k,\mathsf{N}}_n \otimes a_{\mathsf{S}})$ is a submodule of $\mathbb{A}^k_{n,\theta}$ for each $a_{\mathsf{S}} \in \mathfrak{a}^{k,\mathsf{S}}_n$ (and it is straightforward to prove that they are all isomorphic). These submodules are decomposed using the Specht modules.

The following lemma was proved by Brown [20] in the semisimple case, and follows from cellular theory in the general case. See [24]. Recall from the introduction of this section that the head of a representation M is defined as the quotient $M/\operatorname{rad} M$, where $\operatorname{rad} M$ is the radical of M, the intersection of all maximal submodules of M.

Lemma 2.1.11.

1. *Let $\lambda \vdash n-2k$, $\tau \in \mathcal{ST}_\lambda(\{1,\ldots,n-2k\})$, and fix $a_{\mathsf{S}} \in \mathfrak{a}^{k,\mathsf{S}}_n$. The space $\Delta^{\mathbb{B}_{n,\theta}}_\lambda := (\mathbb{C}S_{n-2k}z_\tau) \otimes \mathbb{C}(\mathfrak{a}^{k,\mathsf{N}}_n \otimes a_{\mathsf{S}})$ is a left ideal of $\mathbb{A}^k_{n,\theta}$, and an indecomposable representation of both $\mathbb{A}^k_{n,\theta}$ and the Brauer algebra $\mathbb{B}_{n,\theta}$.*

2. *As vector spaces, we have*

$$\mathbb{B}_{n,\theta} \cong \bigoplus_{\substack{0\le k\le \lfloor \frac{n}{2}\rfloor \\ \lambda\vdash n-2k}} (\Delta^{\mathbb{B}_{n,\theta}}_\lambda)^{\oplus \dim(\Delta^{\mathbb{B}_{n,\theta}}_\lambda)}. \tag{2.8}$$

3. *The space $\psi^{\mathbb{B}_{n,\theta}}_\lambda$ defined as the head of the representation $\Delta^{\mathbb{B}_{n,\theta}}_\lambda$ is simple, and the representations $\psi^{\mathbb{B}_{n,\theta}}_\lambda$, $\lambda \vdash n-2k$, $0 \le k \le \lfloor \frac{n}{2}\rfloor$ are a complete set of irreducible representations of $\mathbb{B}_{n,\theta}$.*

Remark 2.1.12. A more concise way to write the above representation is in the following way. Let $\xi_k = \theta^{-k}\prod_{i=1}^k(\overline{i, n+1-i})$ (note ξ_k is an idempotent). Notice that the space $\mathbb{B}_{n,\theta}\xi_k$ is a $\mathbb{B}_{n,\theta} - \xi_k\mathbb{B}_{n,\theta}\xi_k$ bimodule from the multiplication in $\mathbb{B}_{n,\theta}$, and $\mathbb{C}S_{n-2k}$ (and thereby its submodules) is a left module for $\xi_k\mathbb{B}_{n,\theta}\xi_k \cong \mathbb{B}_{n-2k,\theta}$ (i.e. any $b \in \mathbb{B}_{n-2k,\theta}$ with a bar kills $\mathbb{C}S_{n-2k}$, $\mathbb{C}S_{n-2k}$ acts by the regular representation). Then $\Delta^{\mathbb{B}_{n,\theta}}_\lambda$ can be written as the $\mathbb{B}_{n,\theta}$-module $\mathbb{B}_{n,\theta}\xi_k \otimes_{\xi_k\mathbb{B}_{n,\theta}\xi_k} (\mathbb{C}S_{n-2k}z_\tau)$, with the action of left multiplication on the left tensor factor.

So, the irreducible representations of the Brauer algebra $\mathbb{B}_{n,\theta}$ are indexed by partitions $\lambda \vdash n-2k$, $0 \le k \le \lfloor \frac{n}{2}\rfloor$; let us denote their characters by $\chi^{\mathbb{B}_{n,\theta}}_\lambda$, and dimensions by $d^{\mathbb{B}_{n,\theta}}_\lambda$.

The analogue in the Brauer algebra of the sum of transpositions, and its action on irreducibles (see Lemma 4.1 of [24]) is given by

$$\Delta_\lambda^{\mathbb{B}_{n,\theta}}\left(\sum_{x,y}((x,y)-(\overline{x,y}))\right)=(\mathrm{ct}(\lambda)+k(1-\theta))\mathrm{id}. \tag{2.9}$$

Since the irreducible $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ is a quotient of $\Delta_\lambda^{\mathbb{B}_{n,\theta}}$, it is clear that we also have

$$\psi_\lambda^{\mathbb{B}_{n,\theta}}\left(\sum_{x,y}((x,y)-(\overline{x,y}))\right)=(\mathrm{ct}(\lambda)+k(1-\theta))\mathrm{id}. \tag{2.10}$$

2.1.4 The walled Brauer algebra

We follow [23] and [80] for this section. The walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ is a subalgebra of the Brauer algebra $\mathbb{B}_{n,\theta}$, introduced by Turaev [99], Koike [60] and [8] as having a Schur-Weyl duality with the general linear group $GL(\theta)$, when $GL(\theta)$ acts on tensor space $V^{\otimes n}$ as m tensor multiples of its natural representation, and $n-m$ tensor multiples of the dual of its natural representation. See Chapter 3 for a dull account of the Schur-Weyl duality. As with the Brauer algebra, the walled Brauer algebra has a distinguished basis, but is not a group algebra, and is not always semisimple over the complex numbers. In the semisimple case, its irreducible representations (the cell modules) and their characters were given by Halverson [47] - these are also studied by Nikitin [80]. Cox, de Visscher, Doty and Martin [23] gave an account of the blocks of the walled Brauer algebra in a similar manner to the paper [24] for the Brauer algebra, and also gave a semisimplicity criterion over an arbitrary field.

Let $m \le n$. Returning to the $2n$ labelled vertices we used to define the Brauer algebra, draw a line (a "wall") separating the leftmost $2m$ vertices and the rightmost $2(n-m)$. Let $B_{n,m}$ be the set of diagrams in B_n with the condition that any bar must cross the wall, and any NS-path must not cross the wall. See Figure 2.4. The walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ is the span of $B_{n,m}$, with multiplication as in the Brauer algebra. It is a straightforward exercise to show that the property defining $\mathbb{B}_{n,m,\theta}$ is preserved under concatenation of diagrams, so $\mathbb{B}_{n,m,\theta}$ is indeed a subalgebra of $\mathbb{B}_{n,\theta}$.

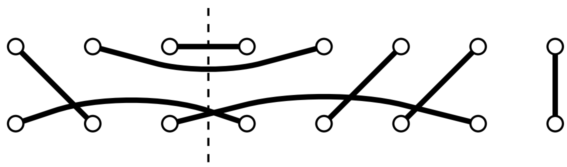


Figure 2.4: A diagram in the basis $B_{8,3}$ of the walled Brauer algebra $\mathbb{B}_{8,3}(\theta)$. Notice that all edges connecting two upper vertices (or two lower) cross the wall, and all edges connecting an upper vertex to a lower vertex do not.

The group algebra $\mathbb{C}[S_m \times S_{n-m}]$ is a subalgebra of $\mathbb{B}_{n,m,\theta}$ whose basis $S_m \times S_{n-m}$ consists of those diagrams with no edges crossing the wall. The transposition (i,j) lies in the walled Brauer algebra if and only if $1 \le i,j \le m$ or $m+1 \le i,j \le n$. The element $(\overline{i,j})$ lies in the walled Brauer algebra if and only if $1 \le i \le m < j \le n$. The elements (i,j) and $(\overline{i,j})$ generate the walled Brauer algebra.

The representation theory of the walled Brauer algebra is immensely similar to that of the Brauer algebra. The walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ is semisimple when $\theta \notin \mathbb{Z}$, and when $\theta \in \mathbb{Z}$, it is semisimple if and only if $|\theta| \geq n-1$ (Theorem 6.3 of [23]). The walled Brauer algebra is also cellular and a tower of recollement, and its cell modules can be defined analogously to those of the cell modules of the Brauer algebra. Again for each cell module, the corresponding irreducible is defined to be the semisimple head of the cell module. Let us be precise.

Analogous to the chain of ideals (2.12), in the walled Brauer algebra we have the descending chain of ideals

$$\mathbb{B}_{n,m,\theta} = \mathbb{B}^0_{n,m,\theta} \supset \mathbb{B}^1_{n,m,\theta} \supset \cdots \supset \mathbb{B}^{\min\{m,n-m\}}_{n,m,\theta}, \tag{2.11}$$

where $\mathbb{B}^k_{n,m,\theta}$ is the span of diagrams in $B_{n,m}$ with at least k bars. The ideal $\mathbb{B}^k_{n,m,\theta}$ can be written as $\mathbb{B}_{n,m,\theta}\xi_k\mathbb{B}_{n,\theta}$, where $\xi_k = \theta^{-k}\prod_{i=1}^k(\overline{i,n+1-i})$. Let $B_{n,m}\langle k\rangle$ be the set of diagrams in $B_{n,m}$ with exactly k bars. The quotient $\mathbb{A}^k_{n,m,\theta} := \mathbb{B}^k_{n,m,\theta}/\mathbb{B}^{k+1}_{n,m,\theta}$ is an algebra (we define $\mathbb{A}^{\min\{m,n-m\}}_{n,m,\theta} := \mathbb{B}^{\min\{m,n-m\}}_{n,m,\theta}$). One can picture $\mathbb{A}^k_{n,m,\theta}$ as the vector space with basis $B_{n,m}\langle k\rangle$, with multiplication the same as $\mathbb{B}_{n,m,\theta}$, except that if two diagrams multiply to give a diagram with more than k bars, then the result is set to zero. It is a module for $\mathbb{B}_{n,m,\theta}$ with action by this (left) multiplication. As a vector space, we have the decomposition:

$$\mathbb{B}_{n,m,\theta} = \mathbb{A}^0_{n,m,\theta} \oplus \mathbb{A}^1_{n,m,\theta} \oplus \cdots \oplus \mathbb{A}^{\min\{m,n-m\}}_{n,m,\theta}. \tag{2.12}$$

Analogously to the Brauer algebra case, can write $\mathbb{A}^k_{n,m,\theta} = \mathbb{C}(S_{m-k}\times S_{n-m-k})\otimes\mathbb{C}(\mathfrak{a}^{k,\mathsf{N}}_{n,m}\times\mathfrak{a}^{k,\mathsf{S}}_{n,m})$, where $\mathfrak{a}^{k,\mathsf{N}}_{n,m}$ is the set of choices of k northern bars (that all cross the wall) and $\mathfrak{a}^{k,\mathsf{S}}_{n,m}$ the same for southern bars. Once again, the space $\mathbb{C}(S_{m-k}\times S_{n-m-k})\otimes\mathbb{C}(\mathfrak{a}^{k,\mathsf{N}}_{n,m}\otimes a_{\mathsf{S}})$ is a submodule of $\mathbb{A}^k_{n,m,\theta}$ for each fixed $a_{\mathsf{S}}\in\mathfrak{a}^{k,\mathsf{S}}_{n,m}$, and we further decompose these submodules using the Specht modules. For $\lambda\vdash m-k$, $\mu\vdash n-m-k$ partitions, we write $(\lambda,\mu)\vdash(m-k,n-m-k)$ for short. For two disjoint sets $P,Q\subset\mathcal{N}=\{1,\ldots,n\}$, we write $\mathcal{ST}_{\lambda,\mu}(P\cup Q)$ for the set of pairs of standard tableaux (τ,π) of shape λ,μ and entries from P and Q respectively, where λ and μ are partitions with size $|P|$ and $|Q|$ respectively.

Lemma 2.1.13.

1. *Let $0\leq k\leq\min\{m,n-m\}$, $(\lambda,\mu)\vdash(m-k,n-m-k)$, $(\tau,\pi)\in\mathcal{ST}_{\lambda,\mu}(\{1,\ldots,m-k\}\uplus\{m-k+1,\ldots,n-2k\})$, and fix $a_{\mathsf{S}}\in\mathfrak{a}^{k,S}_n$. The space $\Delta-lamu := (\mathbb{C}S_{m-k}z_\tau\otimes\mathbb{C}S_{n-m-k}z_\pi)\otimes\mathbb{C}(\mathfrak{a}^{k,\mathsf{N}}_{n,m}\otimes a_S)$ is a left ideal of $\mathbb{A}^k_{n,m,\theta}$, and an indecomposable representation of both $\mathbb{A}^k_{n,m,\theta}$ and the walled Brauer algebra $\mathbb{B}_{n,m,\theta}$.*

2. *As vector spaces, we have*

$$\mathbb{B}_{n,m,\theta} \cong \bigoplus_{\substack{0\leq k\leq\min\{m,n-m\}\\(\lambda,\mu)\vdash(m-k,n-m-k)}} (\Delta^{\mathbb{B}_{n,m,\theta}}_{\lambda,\mu})^{\oplus\dim(\Delta^{\mathbb{B}_{n,m,\theta}}_{\lambda,\mu})}. \tag{2.13}$$

3. *The space $\psi^{\mathbb{B}_{n,m,\theta}}_{(\lambda,\mu)}$ defined as the head of the representation $\Delta^{\mathbb{B}_{n,m,\theta}}_{\lambda,\mu}$ is simple, and the representations $\psi^{\mathbb{B}_{n,m,\theta}}_{(\lambda,\mu)}$, $(\lambda,\mu)\vdash(m-k,n-m-k)$, $0\leq k\leq\min\{m,n-m\}$ are a complete set of irreducible representations of $\mathbb{B}_{n,m,\theta}$.*

Remark 2.1.14. Similarly to the Brauer algebra case, we can write the above representation $\Delta_{\lambda,\mu}^{\mathbb{B}_{n,m,\theta}}$ in a more concise way. Let $\xi_k = \theta^{-k}\prod_{i=1}^{k}(\overline{i,n+1-i})$ (note ξ_k is an idempotent). Notice that the space $\mathbb{B}_{n,m,\theta}\xi_k$ is a $\mathbb{B}_{n,m,\theta} - \xi_k\mathbb{B}_{n,m,\theta}\xi_k$ bimodule from the multiplication in $\mathbb{B}_{n,m,\theta}$, and $\mathbb{C}(S_{m-k}\times S_{n-m-k})$ (and thereby its submodules) is a left module for $\xi_k\mathbb{B}_{n,m,\theta}\xi_k \cong \mathbb{B}_{n-2k,m-k,\theta}$. Then $\Delta_{\lambda,\mu}^{\mathbb{B}_{n,m,\theta}}$ can be written as $\mathbb{B}_{n,m,\theta}\xi_k \otimes_{\xi_k\mathbb{B}_{n,m,\theta}\xi_k} (\mathbb{C}S_{m-k}z_\tau \otimes \mathbb{C}S_{n-m-k}z_\pi)$, with the action of left multiplication on the left tensor factor.

Hence, the irreducible representations of $\mathbb{B}_{n,m,\theta}$ are indexed by

$$\{(\lambda,\mu) \mid \lambda \vdash m-k,\ \mu \vdash n-m-k,\ k = 0,\dots,\min\{m,n-m\}\ \}. \tag{2.14}$$

Analogous to the sum of transpositions (2.1.10) for the symmetric group algebra and (2.9), (2.9) for the Brauer algebra, we have (see, for example, Lemma 4.1 of [23]) in the walled Brauer algebra

$$\Delta_{\lambda,\mu}^{\mathbb{B}_{n,m,\theta}}\left(\sum_{\substack{1\le i<j\le m\\ m<i<j\le n}}(i,j) - \sum_{1\le i\le m<j\le n}(\overline{i,j})\right) = (\mathrm{ct}(\lambda)+\mathrm{ct}(\mu)-k\theta)\mathrm{id}. \tag{2.15}$$

Since the irreducible $\psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}$ is a quotient of $\Delta_{\lambda,\mu}^{\mathbb{B}_{n,m,\theta}}$, it is clear that we also have

$$\psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}\left(\sum_{\substack{1\le i<j\le m\\ m<i<j\le n}}(i,j) - \sum_{1\le i\le m<j\le n}(\overline{i,j})\right) = (\mathrm{ct}(\lambda)+\mathrm{ct}(\mu)-k\theta)\mathrm{id}. \tag{2.16}$$

2.1.5 Classical groups

We follow Goodman and Wallach [44] for this section. We will study the representation theory of the general linear and orthogonal groups (and in order to study the latter, the special orthogonal group).

Let V be a complex vector space of dimension θ. We define a classical group to be a subgroup of $GL(\theta)$ preserving some (symmetric or skew-symmetric bilinear, or Hermitian or skew-Hermitian sesquilinear) form, as well as those groups intersected with the special linear group of invertible maps on V with determinant 1. As noted above, we will concentrate on the general linear, orthogonal and special orthogonal groups. Let $GL(\theta)$ be the group of invertible linear maps from V to V; each basis of V gives a realisation of $GL(\theta)$ as matrices. Fix $(\cdot,\cdot)$, a non-degenerate, symmetric bilinear form on V. We will use this form to define our orthogonal group, that is, $O(\theta)$ is the set of $g \in GL(\theta)$ leaving the form invariant: $(gv,gu) = (v,u)$ for all $v,u \in V$. The determinant of an element of the orthogonal group is ± 1; the special orthogonal group $SO(\theta)$ is the subgroup of $O(\theta)$ of maps whose determinant is 1.

Let V^* be the dual space of V (the vector space of linear functions from V to $\mathbb{C}$). There is always a canonical isomorphism between V and $(V^*)^*$ (the dual of V^*), given by $\mathcal{L}: V \to (V^*)^*$ as $\mathcal{L}(v)(\phi) = \phi(v)$. For any basis $\{f_i\}_{i=1}^{\theta}$ of V, we write $\{f_i^*\}_{i=1}^{\theta}$ for the dual basis of V^*, that is, $f_i^*(f_j) = \delta_{i,j}$. With the form $(\cdot,\cdot)$, there is a further canonical isomorphism, this time between V and V^*, given by

$$L : V \to V^* \text{ with } u \mapsto L(v)(u) = (v, u). \tag{2.17}$$

We can therefore regard the dual basis $\{f_i^*\}_{i=1}^{\theta}$ as elements of V. One can find a basis $\{e_i\}_{i=1}^{\theta}$ satisfying $e_i^* = e_i$ for all i; we call this the standard basis with respect to the bilinear form, or simply the standard basis. In this basis, $g \in O(\theta)$ if and only if its matrix satisfies $g = g^{-\mathsf{T}}$, the inverse transpose of g. Sometimes we will use the following basis (which given the form $(\cdot,\cdot)$, always exists): $\{f_i\}_{i=1}^{\theta}$, where $f_i^* = f_{\theta+1-i}$ for all $1 \le i \le \theta$.

A rational (resp. polynomial) representation of a classical group G is a finite-dimensional representation (ρ, M) of G such that, for all $g \in G$, $\rho(g)$ written as a matrix (in some basis of M) is such that all its entries are rational functions (resp. polynomial functions) in the functions $g_{i,j}$, $1 \le i, j \le \theta$ (the matrix entries of g). We note that this definition is independent of the basis of M chosen.

Lemma 2.1.15. *A finite-dimensional representation (ρ, M) of a classical group G is rational if and only if for all $g \in G$, the matrix entries of $\rho(g)$ are polynomials in the functions $g_{i,j}$, $1 \le i, j \le \theta$, and $\det^{-1}$, the function taking $g \in G$ to the inverse of its determinant.*

Note for $G = SO(\theta)$, rational and polynomial representations are equivalent, since all its matrices have determinant 1. The only representations of G we will consider are (finite-dimensional) rational ones, so we will drop the word rational unless we need to be specifically clear.

Each of the classical groups described has an associated Lie algebra. Recall that a Lie algebra $\mathfrak{g}$ is a vector space with a multiplication $[\cdot,\cdot]$ which is bilinear, satisfying $[x, x] = 0$ for all $x \in \mathfrak{g}$, and the Jacobi identity, $[x,[y,z]] + [z,[x,y]] + [y,[z,x]] = 0$ for all $x, y, z \in \mathfrak{g}$. The general linear Lie algebra $\mathfrak{gl}(\theta)$ is the Lie algebra associated with $GL(\theta)$; it is the space of $\theta \times \theta$ matrices with multiplication given by the Lie bracket $[x, y] = xy - yx$. The general bilinear form $(\cdot,\cdot)$ defined above, given a basis $\{f_i\}_{i=1}^{\theta}$ of V, defines a $\theta \times \theta$ matrix S such that $(v, w) = v^{\mathsf{T}} S w$ for all $v, w \in V$, where on the right hand side the vectors v and w are written as column vectors in the chosen basis. The Lie algebra associated with the orthogonal and special orthogonal groups is the same, $\mathfrak{so}(\theta)$, the space of $\theta \times \theta$ matrices satisfying $x^{\mathsf{T}} S + Sx = 0$, with multiplication also given by the Lie bracket. Note that in the standard basis, the matrix S is the identity, so $\mathfrak{so}(\theta)$ is the space of skew-symmetric matrices. In the basis $\{f_i\}_{i=1}^{\theta}$, where $f_i^* = f_{\theta+1-i}$ for all i, S is the matrix with entries $S_{i,j} = \delta_{\theta+1-i,j}$, for all $1 \le i, j \le \theta$.

Let us note that each representation of $GL(\theta)$ or $SO(\theta)$ is equivalently a representation of its associated Lie algebra, via the differential map. In particular, the natural representation of each classical group G is defined to be the vector space V itself, with $g \in G$ acting as itself. In this case, the corresponding representation of the Lie algebra $\mathfrak{g}$ is also the vector space V, with $x \in \mathfrak{g}$ acting as itself.

We will not give a full account of the representation theory of our chosen classical groups; instead collecting important results that will be used in later parts of this work. We follow Sections 3.1.4, 5.5.4 and 5.5.5 of [44]. Fix the basis of V, $\{f_i\}_{i=1}^{\theta}$, where $f_i^* = f_{\theta+1-i}$ for all i; this fixes each of $GL(\theta)$, $O(\theta)$, $SO(\theta)$ and $\mathfrak{gl}(\theta)$ and $\mathfrak{so}(\theta)$ as matrices. For each of the Lie algebras $\mathfrak{g}$, let $\mathfrak{h}_{\mathfrak{g}}$ be the Lie sub-algebra of diagonal matrices. Let $r = \lfloor \frac{\theta}{2} \rfloor$. We have the descriptions:

$$\begin{aligned} \mathfrak{h}_{\mathfrak{gl}(\theta)} &= \{\mathrm{diag}[x_1, \dots, x_\theta] \ : \ x_i \in \mathbb{C}\}; \\ \mathfrak{h}_{\mathfrak{so}(\theta=2r+1)} &= \{\mathrm{diag}[x_1, \dots, x_r, 1, -x_r, \dots, -x_1] \ : \ x_i \in \mathbb{C}\}; \\ \mathfrak{h}_{\mathfrak{so}(\theta=2r)} &= \{\mathrm{diag}[x_1, \dots, x_r, -x_r, \dots, -x_1] \ : \ x_i \in \mathbb{C}\}; \end{aligned} \tag{2.18}$$

the latter two cases differentiating between the cases $\theta = 2r+1$ odd, and $\theta = 2r$ even. For each $\mathfrak{h}_{\mathfrak{g}}$, let $\mathfrak{h}_{\mathfrak{g}}^*$ be the dual space. Fix a basis $\varepsilon_i(x) = x_i$ of $\mathfrak{h}_{\mathfrak{g}}$ (producing the i^{th} entry on the diagonal), for $i = 1, \dots, \theta$ in the $\mathfrak{gl}(\theta)$ case and $i = 1, \dots, r$ in the $\mathfrak{so}(\theta)$ case. Note in the $\mathfrak{so}(\theta)$ case, $\varepsilon_i = -\varepsilon_{\theta+1-i}$ for all $1 \le i \le \theta$.

For $\mathfrak{g} = \mathfrak{gl}(\theta), \mathfrak{so}(\theta)$, M a module of $\mathfrak{g}$, and $\lambda \in \mathfrak{h}_{\mathfrak{g}}^*$, then a vector $v \in M$ is called a weight vector for $\mathfrak{g}$ of weight λ if for all $x \in \mathfrak{h}_{\mathfrak{g}}$, $x \cdot v = \lambda(x)v$. For each of the groups $GL(\theta)$, $SO(\theta)$, we define a set called its dominant integral weights; these are the weights

$$\begin{aligned} P_+(GL(\theta)) &= \{\lambda = (\lambda_1\varepsilon_1, \dots, \lambda_\theta\varepsilon_\theta) \in \mathfrak{h}^*_{\mathfrak{gl}(\theta)} \ : \ \lambda_i \in \mathbb{Z}, \ \lambda_1 \ge \dots \ge \lambda_\theta\}; \\ P_+(SO(\theta = 2r+1)) &= \{\lambda = (\lambda_1\varepsilon_1, \dots, \lambda_r\varepsilon_r) \in \mathfrak{h}^*_{\mathfrak{so}(\theta)} \ : \ \lambda_i \in \mathbb{Z}, \ \lambda_1 \ge \dots \ge \lambda_r \ge 0\}; \\ P_+(SO(\theta = 2r)) &= \{\lambda = (\lambda_1\varepsilon_1, \dots, \lambda_r\varepsilon_r) \in \mathfrak{h}^*_{\mathfrak{so}(\theta)} \ : \ \lambda_i \in \mathbb{Z}, \ \lambda_1 \ge \dots \ge |\lambda_r| \ge 0\}; \end{aligned} \tag{2.19}$$

the latter two cases differentiating between the cases $\theta = 2r+1$ odd, and $\theta = 2r$ even. We will often identify a weight λ with its θ-tuple or r-tuple which gives λ in the basis of functions ε_i. Let us define the simple root vectors of $\mathfrak{gl}(\theta)$ to be the matrices

$$\{E_{j,j+1} \ : \ j = 1, \dots, \theta - 1\}, \tag{2.20}$$

(where $E_{i,j}$ is the matrix which has all entries zero except for 1 in the i, j entry), and those of $\mathfrak{so}(\theta)$ to be

$$\{E_{j,j+1} \ : \ j = 1, \dots, r-1\} \cup \begin{cases} E_{r,r+1} - E_{r+1,r+2}, & \text{for } \theta = 2r+1 \\ E_{r,r+2} - E_{r-1,r+1}, & \text{for } \theta = 2r. \end{cases} \tag{2.21}$$

We can now turn to irreducible rational representations of $G = GL(\theta)$ and $SO(\theta)$. The theorem of the highest weight tells us the following.

Theorem 2.1.16 (Theorem of the highest weight)**.**

1. *For each of the groups $G = GL(\theta)$, $SO(\theta)$, every irreducible (finite-dimensional) rational representation of G has a unique vector v_0, known as the highest weight vector, which is a weight vector for $\mathfrak{g}$ (the Lie algebra of G) with some dominant weight $\lambda \in P_+(G)$, and is killed by every simple root vector. We say that such a representation has highest weight λ.*

2. *For each of $G = GL(\theta)$, $SO(\theta)$, and for each $\lambda \in P_+(G)$, there is a unique irreducible representation ψ_λ^G with highest weight λ.*

Remark 2.1.17. To summarise, the irreducible rational representations of $GL(\theta)$ are indexed by θ-tuples of non-increasing integers, and those of $SO(\theta)$ are indexed by r-tuples of non-increasing, non-negative integers (where $r = \lfloor \frac{\theta}{2} \rfloor$), with the exception that in the case $\theta = 2r$, the r^{th} entry can be negative.

We can also characterise the irreducible polynomial representations of $GL(\theta)$ (recall for $SO(\theta)$, polynomial and rational representations are the same).

Lemma 2.1.18. *An irreducible rational representation of $GL(\theta)$ with highest weight λ is polynomial if and only if all the entries λ_i, $i = 1, \dots, \theta$, are non-negative.*

Hence one can say that the irreducible polynomial representations of $GL(\theta)$ are indexed by partitions of any size, with at most θ parts. For each irreducible rational representation ψ_π^G of each of $G = GL(\theta)$ and $SO(\theta)$, let us denote its character and dimension by χ_π^G and d_π^G respectively.

There is a second way to index the irreducibles of the general linear group $GL(\theta)$. One can biject the set of θ-tuples of integers which are non-increasing with the set of pairs of partitions (λ, μ) of any size, satisfying $\lambda_1^\intercal + \mu_1^\intercal \leq \theta$. Indeed, given such a pair, let $[\lambda, \mu]$ be a θ-tuple defined as

$$[\lambda, \mu]_i = \lambda_i - \mu_{\theta+1-i}, \tag{2.22}$$

for each $i = 1, \dots, \theta$. Conversely, given a θ-tuple ρ, let $\lambda_i = \rho_i$ for all i where $\rho_i > 0$, and $\lambda_i = 0$ otherwise, and let $\mu_i = -\rho_{\theta+1-i}$ whenever $\rho_{\theta+1-i} < 0$, and $\mu_i = 0$ otherwise.

Lemma 2.1.19 (Theorem 3.2.13 of [44])**.** *If $\psi_{[\lambda,\mu]}^{GL(\theta)}$ is an irreducible representation of $GL(\theta)$ with highest weight $[\lambda, \mu]$, then its dual is the irreducible representation $\psi_{[\mu,\lambda]}^{GL(\theta)}$. In particular, the dual of an irreducible polynomial representation $\psi_\lambda^{GL(\theta)}$ is $\psi_{[\varnothing,\lambda]}^{GL(\theta)}$, $\varnothing$ denoting the empty partition.*

It remains to characterise the irreducible representations of the orthogonal group $G = O(\theta)$. These are based on those of $SO(\theta)$. The irreducible rational representations of $O(\theta)$ are indexed by partitions λ of any size satisfying $\lambda_1^\intercal + \lambda_2^\intercal \leq \theta$. For any such λ, let λ' be the same partition, with its first column $\lambda_1^\intercal$ replaced by $\theta - \lambda_1^\intercal$. Notice that $\lambda'' = \lambda$, and $\lambda = \lambda'$ if and only if $\theta = 2r$ and $\lambda_1^\intercal = r$, that is, λ has r parts. In the case $\lambda \neq \lambda'$, one of the pair λ, λ' has its first column strictly shorter than r, and one strictly longer; let us label these by λ^+, λ^- respectively. Let us also define a specific element $g_0 \in O(\theta) \smallsetminus SO(\theta)$ in the case $\theta = 2r$: g_0 fixes each basis vector f_i of V, $i \neq r, r+1$, and exchanges the basis vectors f_r, f_{r+1}, where recall $\{f_i\}_{1 \leq i \leq \theta}$ is the basis of V with $f_i^* = f_{\theta+1-i}$ for all $1 \leq i \leq \theta$.

Theorem 2.1.20. *The irreducible rational representations $\psi_\lambda^{O(\theta)}$ of $O(\theta)$ are indexed by partitions λ of any size satisfying $\lambda_1^\intercal + \lambda_2^\intercal \leq \theta$. For $\lambda \neq \lambda'$ (that is, θ odd or $\theta = 2r$ even and $\lambda_1^\intercal \neq r$), we have the restriction*

$$\mathrm{res}_{SO(\theta)}^{O(\theta)} \psi_{\lambda^+}^{O(\theta)} = \mathrm{res}_{SO(\theta)}^{O(\theta)} \psi_{\lambda^-}^{O(\theta)} = \psi_{\lambda^+}^{SO(\theta)}. \tag{2.23}$$

Further, when θ is odd, the element $-\boldsymbol{id} \in O(\theta) \smallsetminus SO(\theta)$ acts on $\psi_\lambda^{O(\theta)}$ as $(-1)^{|\lambda|}\boldsymbol{id}$, and when $\theta = 2r$ even, the element $g_0 \in O(\theta) \smallsetminus SO(\theta)$ acts on the ($SO(\theta)$) highest weight vector in $\psi_{\lambda^\pm}^{O(\theta)}$ as $\pm\boldsymbol{id}$. In the case $\lambda = \lambda'$ (this implies $\theta = 2r$ even), we have the restriction

$$\mathrm{res}_{SO(\theta)}^{O(\theta)} \psi_\lambda^{O(\theta)} = \psi_\lambda^{SO(\theta)} + \psi_{\lambda^\circ}^{SO(\theta)}, \tag{2.24}$$

where λ° is λ with λ_r replaced with $-\lambda_r$.

For each irreducible representation $\psi_\lambda^{O(\theta)}$ of $O(\theta)$, let us denote its character and dimension by $\chi_\lambda^{O(\theta)}$, and $d_\lambda^{O(\theta)}$ respectively.

2.1.6 Branching rules

In the rest of this section we give useful results about restrictions of representations, known as branching rules. Let A_1, A_2 be algebras over $\mathbb{C}$, with A_2 a subalgebra of A_1, and let (ψ, M), be a representation of A_1, with character χ. Then we define the restriction of (ψ, M) to A_2, $(\mathrm{res}_{A_2}^{A_1}[\psi], M)$, to be simply the function ψ restricted to A_2. We similarly define the restricted character $\mathrm{res}_{A_2}^{A_1}[\chi] = \mathrm{Tr}\big(\mathrm{res}_{A_2}^{A_1}[\psi]\big)$.

The decomposition of restrictions of representations of $\mathbb{C}S_n$ to $\mathbb{C}S_{n-1}$ and $\mathbb{B}_{n,\theta}$ to $\mathbb{B}_{n-1,\theta}$ are well studied. Let $\rho \vdash n$, $\lambda \vdash n-2k$, $0 \le k \le \lfloor \frac{n}{2} \rfloor$. Recall we denoted the irreducible representations of $\mathbb{C}S_n$ and $\mathbb{B}_{n,\theta}$ with partitions ρ and λ by $\psi_\rho^{S_n}$ and $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ respectively, and the cell module of $\mathbb{B}_{n,\theta}$ with partition λ by $\Delta_\lambda^{\mathbb{B}_{n,\theta}}$. We have the following (see, for example, Sections 4 and 5 (and Figures 1 and 2) of [32], and Proposition 1.3 of [78]):

$$\begin{aligned} \mathrm{res}_{S_{n-1}}^{S_n}[\psi_\rho^{S_n}] &= \sum_{\underline{\rho}=\rho-\square} \psi_{\underline{\rho}}^{S_{n-1}}; \\ \mathrm{res}_{\mathbb{B}_{n-1,\theta}}^{\mathbb{B}_{n,\theta}}[\Delta_\lambda^{\mathbb{B}_{n,\theta}}] &= \sum_{\underline{\lambda}=\lambda\pm\square} \Delta_{\underline{\lambda}}^{\mathbb{B}_{n-1,\theta}}; \end{aligned} \tag{2.25}$$

and if λ further satisfies $\lambda_1^\intercal + \lambda_2^\intercal \le \theta$,

$$\mathrm{res}_{\mathbb{B}_{n-1,\theta}}^{\mathbb{B}_{n,\theta}}[\psi_\lambda^{\mathbb{B}_{n,\theta}}] = \sum_{\substack{\underline{\lambda}=\lambda\pm\square \\ \underline{\lambda}_1^\intercal+\underline{\lambda}_2^\intercal\le\theta}} \chi_{\underline{\lambda}}^{\mathbb{B}_{n-1,\theta}}, \tag{2.26}$$

where in the first equality the sum is over all $\underline{\rho} \vdash n-1$ whose Young diagram can be obtained from that of ρ by removing a box; in the second the sum is over $\underline{\lambda} \vdash n-1-2r$, $0 \le r \le \lfloor \frac{n-1}{2} \rfloor$, whose Young diagram can be obtained from that of λ by removing or adding a box; and in the third the sum is the same as the second, except we are restricted to those $\underline{\lambda}$ with $\underline{\lambda}_1^\intercal + \underline{\lambda}_2^\intercal \le \theta$.

We now describe how cell modules of $\mathbb{B}_{n,\theta}$ decompose when restricted to $\mathbb{C}S_n$; first we need to describe the Littlewood-Richardson rule. See section I.9 of Macdonald [68] for a detailed exposition. The Littlewood-Richardson rule has several forms; we will describe three of them here.

Firstly, we need to define induction of a representation - this is in some sense the inverse process to restriction. Let A_1, A_2 be algebras over $\mathbb{C}$, with A_2 a subalgebra of A_1, and let (ψ, M), be a representation of A_2, with character χ. Then we define the induction of (ψ, M) to A_1, $(\mathrm{ind}_{A_2}^{A_1}[\psi], M)$, to be the tensor product $A_1 \otimes_{A_2} M$, with action $a \cdot (b \otimes v) = (ab) \otimes v$ for all $a, b \in A_1$, $v \in M$ (recall $A_1 \otimes_{A_2} M$ is the vector space $A_1 \otimes M$ quotiented by the relations $a_1 a_2 \otimes v = a_1 \otimes a_2 v$. We similarly define the induced character $\mathrm{ind}_{A_2}^{A_1}[\chi]$ to be the character of the induced representation.

Given a pair of representations ψ_r of $\mathbb{C}S_r$, and ψ_s of $\mathbb{C}S_s$, there exists a product representation which we call $\psi_1 \times \psi_2$ of $\mathbb{C}S_{r+s}$ (given by the induction from $\mathbb{C}S_r \otimes \mathbb{C}S_s$ to $\mathbb{C}S_{r+s}$ of the box-tensor product $\psi_1 \boxtimes \psi_2$). When the two representations are irreducible, their product is given by the Littlewood-Richardson rule:

$$\psi_\pi^{S_r} \times \psi_\mu^{S_s} = \sum_{\xi \vdash r+s} c^\xi_{\pi,\mu} \psi_\xi^{S_{r+s}}, \tag{2.27}$$

where $c^\xi_{\pi,\mu}$ is the Littlewood-Richardson coefficient. For two Young diagrams ξ, μ of any size, such that $\mu \leq \xi$ (that is, $\mu_i \leq \xi_i$ for all rows i of both diagrams), the skew-diagram $\xi \setminus \mu$ is the Young diagram of ξ with the boxes of μ removed from it. For our purposes we need only note that the Littlewood-Richardson coefficient $c^\xi_{\pi,\mu}$ is non-zero only if $\pi \leq \xi$ and $\mu \leq \xi$, and it is determined by π and the skew-diagram $\xi \setminus \mu$. The formula (2.27) is equivalent to the statement

$$\mathrm{res}^{S_{r+s}}_{S_r \times S_s}[\psi_\xi^{S_{r+s}}] = \sum_{\substack{\pi \vdash r \\ \mu \vdash s}} c^\xi_{\pi,\mu} \psi_\pi^{S_r} \boxtimes \psi_\mu^{S_s}, \tag{2.28}$$

the equivalence being a consequence of a general theorem called Frobenius reciprocity, see Section 5 of [34].

The product (2.27) is equivalent to, and often thought of as, the ordinary product of symmetric polynomials. Let $\rho \vdash n$ have at most θ parts. Define the Schur polynomial s_ρ (on θ symbols) to be the symmetric polynomial in the variables $x_1, \ldots, x_\theta$ as

$$s_\rho(x_1, \ldots, x_\theta) = \frac{\det\left[x_i^{\rho_j+\theta-j}\right]_{i,j=1}^\theta}{\det\left[x_i^{\theta-j}\right]_{i,j=1}^\theta} = \frac{\det\left[x_i^{\rho_j+\theta-j}\right]_{i,j=1}^\theta}{\prod_{1 \leq i < j \leq \theta}(x_i - x_j)}. \tag{2.29}$$

The Schur polynomials are, in fact, the polynomial characters of the general linear group.

Lemma 2.1.21. *Let $g \in GL(\theta)$ with eigenvalues $x_1, \ldots, x_\theta$. Then for any partition ρ with at most θ parts,*

$$\chi_\rho^{GL(\theta)}(g) = s_\rho(x_1, \ldots, x_\theta). \tag{2.30}$$

If s_π, s_μ are the Schur polynomials associated with π and μ, then the second form of the Littlewood-Richardson rule is

$$s_\pi s_\mu = \sum_{\xi \vdash r+s} c^\xi_{\pi,\mu} s_\xi. \tag{2.31}$$

Now since the Schur polynomials are the irreducible characters of $GL(\theta)$, we have by Lemma 2.1.21 that as a representation of $GL(\theta)$,

$$\psi_\pi^{GL(\theta)} \otimes \psi_\mu^{GL(\theta)} = \bigoplus_{\xi \vdash r+s} c^\xi_{\pi,\mu} \psi_\xi^{GL(\theta)}, \tag{2.32}$$

for all partitions π, μ with at most θ parts, which is our third form of the Littlewood-Richardson rule. Let us note that the Schur polynomials, along with the characters of the special orthogonal group, can be expressed in terms of tableaux. See, for example, Sundaram [94]. Let U be a finite, totally ordered set, and let $\lambda \vdash |U|$. Then τ is a semistandard tableau with shape λ and entries from U if τ is the Young diagram of λ with each box filled with an element of U, such that the entries are non-decreasing along rows and strictly increasing down columns. (Note in contrast to standard tableaux, one can have repeated entries in a semistandard tableau). Let $\mathcal{SS}_\lambda(U)$ be the set of semistandard tableaux with entries from U with shape λ, and $\mathcal{SS}_n(U) = \bigcup_{\lambda \vdash n} \mathcal{SS}_\lambda(U)$. Then, letting $\Theta = \{1, \dots, \theta\}$, (see Theorem 2.2 of [94]) we have

$$s_\rho(x_1, \dots, x_\theta) = \sum_{\tau \in \mathcal{SS}_\rho(\Theta)} \prod_{i=1}^{\theta} x_i^{m_i}, \tag{2.33}$$

where m_i is the number of times $i \in \Theta$ appears in τ. There are several analogous formulae for the characters of the special orthogonal group. One of them is a formula due to King (Theorem 2.5 of [94]), which is stated in full in Chapter 5, see (5.48).

A special case of the Littlewood-Richardson rule, known as the Pieri rule, deals with the case when one of the factors $\psi_\pi^{S_r}$ has Young diagram π with only one row, or only one column, that is, when $\pi = (s)$ or (1^s), for some s. In this case we have

$$\psi_{(r)}^{S_r} \times \psi_\mu^{S_s} = \sum_{\xi \vdash r+s} \psi_\xi^{S_{r+s}}, \qquad \psi_{(1^r)}^{S_r} \times \psi_\mu^{S_s} = \sum_{\xi \vdash r+s} \psi_\xi^{S_{r+s}}, \tag{2.34}$$

where in the first equation, the sum is over all ξ whose Young diagram can be obtained from that of μ by adding r boxes, no two of which are in the same column, and the first equation the sum is over the same ξ except that no two of the boxes one adds can be in the same row. A special case of this is the case when $\pi = (1^\theta)$ (here $\psi_\pi^{GL(\theta)}$ is the 1-dimensional determinant representation of $GL(\theta)$), in which case we have

$$\psi_{(1^\theta)}^{GL(\theta)} \otimes \psi_\mu^{GL(\theta)} = \psi_{\xi+\underline{1}}^{GL(\theta)}, \tag{2.35}$$

where $\xi + \underline{1}$ is the θ-tuple with i^{th} entry equal to $\xi_i + 1$, for all $i = 1, \dots, \theta$. There exists an equivalent Pieri rule for the orthogonal group (see Okada [81]); we will note only the special case when $s = 1$, that is, when $\pi = (1)$:

$$\psi_\lambda^{O(\theta)} \otimes \psi_{(1)}^{O(\theta)} = \bigoplus_{\substack{\underline{\lambda} = \lambda \pm \square \\ \underline{\lambda}_1^\top + \underline{\lambda}_2^\top \le \theta}} \psi_{\underline{\lambda}}^{O(\theta)}; \tag{2.36}$$

where the sum is over all partitions $\underline{\lambda}$ satisfying $\underline{\lambda}_1^\top + \underline{\lambda}_2^\top \le \theta$ which can be obtained from λ by adding or removing a box; in fact one appication of Schur-Weyl duality is that this is equivalent to (2.26) - see Lemma 5.7.1.

We can now give the restriction of cell modules of $\mathbb{B}_{n,\theta}$ to $\mathbb{C}S_n$. This result is from Theorem 4.1 of Hanlon and Wales [52], and is a special case of Theorem 5.1 of [84]. We call a partition π even if all its parts π_i are even. Let $\lambda \vdash n - 2k$, $0 \le k \le \lfloor \frac{n}{2} \rfloor$. Then

$$\mathrm{res}_{S_n}^{\mathbb{B}_{n,\theta}} [\Delta_\lambda^{\mathbb{B}_{n,\theta}}] = \bigoplus_{\rho \vdash n} (\psi_\rho^{S_n})^{\oplus \tilde{b}_{\lambda,\rho}^{n,\theta}} = \psi_\lambda^{S_{n-2k}} \times \bigoplus_{\substack{\pi \vdash 2k \\ \pi \text{ even}}} \psi_\pi^{S_{2k}}, \tag{2.37}$$

or,

$$\tilde{b}^{n,\theta}_{\lambda,\rho} = \sum_{\substack{\pi \vdash 2k \\ \pi \text{ even}}} c^{\rho}_{\lambda,\pi}.$$

Note that as a consequence, $\tilde{b}^{n,\theta}_{\lambda,\rho} = 0$ if $\lambda \nsubseteq \rho$, and that $\tilde{b}^{n,\theta}_{\lambda,\rho}$ is fully determined by the skew-diagram $\rho \setminus \lambda$.

Now let us define the branching coefficients $b^{n,\theta}_{\lambda,\rho}$ by

$$\operatorname{res}^{\mathbb{B}_{n,\theta}}_{S_n}[\psi^{\mathbb{B}_{n,\theta}}_{\lambda}] = \bigoplus_{\rho \vdash n} (\psi^{S_n}_{\rho})^{\oplus b^{n,\theta}_{\lambda,\rho}}. \tag{2.38}$$

There is no concise formula for the branching coefficients $b^{n,\theta}_{\lambda,\rho}$ of the form (2.37), however in the case $\lambda^{\mathsf{T}}_1 + \lambda^{\mathsf{T}}_2 \leq \theta$ and $\rho^{\mathsf{T}}_1 \leq \theta$, Okada (Proposition 2.5 and the orthogonal group version of equation (5.1) of [81]) gives an explicit algorithm for calculating the coefficients $b^{n,\theta}_{\lambda,\rho}$ in terms of the coefficients $\tilde{b}^{n,\theta}_{\lambda,\rho}$. Note that the results in [81] are given in terms of restriction from the General Linear group to the Orthogonal group - in Lemma 5.7.1 we show that this is equivalent. This equivalence is a consequence of the Schur-Weyl duality.

The branching coefficients $b^{n,\theta}_{\lambda,\rho}$ will prove to be crucial in the results on quantum spin systems in Chapter 5; in particular we will need a condition for $b^{n,\theta}_{\lambda,\rho} > 0$. We give results on this problem in Section 5.7.

Similar branching rules hold for the restriction of cell modules and irreducibles of the walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ to the symmetric group algebra $\mathbb{C}(S_m \times S_{n-m})$. Corollary 7.24 of Halverson [47], restated in Lemma 4.1 of [23], gives

$$\operatorname{res}^{\mathbb{B}_{n,m,\theta}}_{S_m \times S_{n-m}}[\Delta^{\mathbb{B}_{n,m,\theta}}_{\lambda,\mu}] = \bigoplus_{(\rho,\xi) \vdash (m,n-m)} (\psi^{S_m}_{\rho} \boxtimes \psi^{S_{n-m}}_{\xi})^{\oplus \tilde{b}^{n,m,\theta}_{(\lambda,\mu),(\rho,\xi)}}, \tag{2.39}$$

where

$$\tilde{b}^{n,m,\theta}_{(\lambda,\mu),(\rho,\xi)} = \sum_{\pi \vdash k} c^{\rho}_{\lambda,\pi} c^{\xi}_{\mu,\pi}.$$

Let us similarly define the branching coefficients $b^{n,m,\theta}_{(\lambda,\mu),(\rho,\xi)}$ by

$$\operatorname{res}^{\mathbb{B}_{n,m,\theta}}_{S_m \times S_{n-m}}[\psi^{\mathbb{B}_{n,m,\theta}}_{(\lambda,\mu)}] = \bigoplus_{(\rho,\xi) \vdash (m,n-m)} (\psi^{S_m}_{\rho} \boxtimes \psi^{S_{n-m}}_{\xi})^{\oplus b^{n,m,\theta}_{(\lambda,\mu),(\rho,\xi)}}. \tag{2.40}$$

There is no general, concise formula for the branching coefficients $b^{n,m,\theta}_{(\lambda,\mu),(\rho,\xi)}$; however, in contrast the the Brauer algebra case, in the case $\lambda^{\mathsf{T}}_1 + \mu^{\mathsf{T}}_1 \leq \theta$, an expression does exist. This will be crucial to part of our work on quantum spin systems in Chapter 6. Let us summarise the result here. For a θ-tuple (which could be a partition) λ, and an integer s, let us define the θ-tuple $\lambda + \underline{s}$ as having i^{th} entry equal to $\lambda_i + s$, for each $1 \leq i \leq \theta$. Recall that for two partitions λ, μ with length at most θ, we define $[\lambda, \mu]$ to be the θ-tuple $[\lambda, \mu]_i = \lambda_i - \mu_{\theta+1-i}$, and lastly recall that the irreducible representations of $GL(\theta)$ can be indexed by pairs of partitions λ, μ satisfying $\lambda^{\mathsf{T}}_1 + \mu^{\mathsf{T}}_1 \leq \theta$. The following lemma is from Section 6.2.2.

Lemma 2.1.22. *Let ρ, ξ be partitions with at most θ parts, and let*

$$\psi_{\rho}^{GL(\theta)} \otimes \psi_{[\varnothing,\xi]}^{GL(\theta)} = \bigoplus_{\substack{\lambda,\mu \\ \lambda_1^T+\mu_1^T\leq\theta}} \hat{b}^{n,m,\theta}_{[\lambda,\mu],(\pi,\xi)} \psi_{[\lambda,\mu]}^{GL(\theta)}. \tag{2.41}$$

Then $\hat{b}^{n,m,r}_{[\lambda,\mu],(\pi,\xi)} = b^{n,m,r}_{(\lambda,\mu),(\pi,\xi)} = c^{[\lambda,\mu]+\underline{\xi_1}}_{\pi,[\varnothing,\xi]+\underline{\xi_1}}$, where the latter term is the Littlewood-Richardson coefficient.

Chapter 3

Schur-Weyl duality

Schur-Weyl duality is a powerful theory which connects and explores the representation theory of several pairs of groups, or pairs of algebras. In each instance, it is essentially an example of a general theorem of representation theory, the double centraliser theorem. Let us state this theorem.

Theorem 3.0.1 (The double centraliser theorem)**.** *Let W be a complex vector space. Let $A \subset End(W)$ be a semisimple subalgebra and let $A' = End_A(W) := \{b \in End(W) \ : \ ab = ba \ \forall a \in A\}$, its centraliser. Then:*

1. *A' is also a semisimple algebra, and $A'' = A$;*

2. *As a representation of $A \otimes A'$, W decomposes as*

$$W = \bigoplus_{i=1}^{k} U_i \boxtimes V_i, \tag{3.1}$$

where U_i (resp. V_i), $i = 1, \ldots, k$ is an exhaustive list of pairwise non-isomorphic irreducible representations of A (resp. A').

The original version of Schur Weyl duality is for the pair of groups $GL(\theta)$ and S_n. Let V be a complex vector space of dimension θ. Briefly, the space $V^{\otimes n}$ is a representation for both $GL(\theta)$ and S_n, and these actions are each others' centralisers. The double centraliser theorem then decomposes the space $V^{\otimes n}$ as a module for $GL(\theta) \times S_n$, into a direct sum of irreducibles $U_i \boxtimes V_i$; this is the first half of the proof that we will give in this section. The second half consists of using the Specht modules and highest weight theory from Chapter 2 to identify $U_i \boxtimes V_i$ as the irreducible $\psi_\rho^{GL(\theta)} \boxtimes \psi_\rho^{S_n}$, where, note, the partition ρ indexing the irreducible of S_n is the same as the θ-tuple ρ indexing the highest weight of the irreducible of $GL(\theta)$. The irreducibles $\psi_\rho^{GL(\theta)} \boxtimes \psi_\rho^{S_n}$ that appear are all those indexed by $\rho \vdash n$ with at most θ parts and each appears once each.

One reason Schur-Weyl duality is powerful is that it gives a concrete realisation of every irreducible polynomial representation of $GL(\theta)$ in some tensor power of V, and gives a concrete realisation of every irreducible representation of S_n in $V^{\otimes n}$, so long as $\theta \geq n$. In particular, one can even define the irreducibles of $GL(\theta)$ and S_n as those that appear in Schur-Weyl duality (once one has proved the first part of the theorem, as described above, using the double centraliser theorem) and then work towards the highest weight theory and Specht module theory. As noted above, in our case, we will go the other way around, and identify the irreducibles in $V^{\otimes n}$ using these theories. A second reason that Schur-Weyl duality is powerful is that it intimately links the representation theories of the two groups $GL(\theta)$ and S_n. Statements about one group can often be transferred into statements about the other. We will see examples of this in Sections 5 and 6, for example Lemma 5.7.1 and Lemma 6.2.3.

There are many other versions of Schur-Weyl duality. We will study, (in addition to the $GL(\theta) - S_n$ version), two of them. Richard Brauer [19] proved the version linking the orthogonal group $O(\theta)$ and the Brauer algebra $\mathbb{B}_{n,\theta}$, where note, the dimension θ of V is the same as the multiplicative parameter θ of the Brauer algebra. The statement is essentially the same as in the $GL(\theta)-S_n$ version. The irreducibles $\psi_\lambda^{O(\theta)} \boxtimes \psi_\lambda^{\mathbb{B}_{n,\theta}}$ appearing in $V^{\otimes n}$ are those $\lambda \vdash n-2k$, $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$, such that $\lambda_1^\intercal + \lambda_2^\intercal \leq \theta$. Note that in this case, for fixed n and θ, often not all irreducibles of $\mathbb{B}_{n,\theta}$ appear in tensor space. Let us note that another version of Schur-Weyl duality analogously links the symplectic group $Sp(\theta)$ and the Brauer algebra $\mathbb{B}_{n,-\theta}$ for $\theta > 0$ even. Since we will not use this version in our applications, we omit it from our treatment in this work; however we note that its proof follows almost identical lines to that of the $O(\theta) - \mathbb{B}_{n,\theta}$ version presented in this section.

The third version that we will study links $GL(\theta)$ and the walled Brauer algebra $\mathbb{B}_{n,m,\theta}$, proved by [60] and [8]. The statement is essentially the same as in the $GL(\theta)-S_n$ version, although the action of $GL(\theta)$ is different. The irreducibles $\psi_{[\lambda,\mu]}^{GL(\theta)} \boxtimes \psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}$ appearing in $V^{\otimes n}$ are those $\lambda \vdash m-k, \mu \vdash n-m-k$, $0 \leq k \leq \min\{m, n-m\}$, such that $\lambda_1^\intercal + \mu_1^\intercal \leq \theta$. Note that as in the Brauer algebra case, for fixed n, m and θ, often not all irreducibles of $\mathbb{B}_{n,m,\theta}$ appear in tensor space.

Let us recall that one of the objectives of this section is to present as unified, streamlined and self-contained an account of these Schur-Weyl dualities as possible, in particular the $O(\theta) - \mathbb{B}_{n,\theta}$ and $GL(\theta) - \mathbb{B}_{n,m,\theta}$ versions, which often appear in partial forms in the literature.

This section is organised into three parts. In the first part we will prove the double centraliser theorem 3.0.1, and state our the three versions of Schur-Weyl duality precisely, Theorems 3.0.2, 3.0.3, and 3.0.5. As discussed above, in the second part we use the double centraliser theorem and a classical theory called invariant theory to prove each of the duality statements, up to the specific identification of the irreducibles concerned. Invariant theory is the study of functions on a vector space which are invariant under the action of a group, usually a classical group. That is, if $f : V \to \mathbb{C}$ lies in the algebra generated by V^*, and a group G acts on V, then f is an invariant if $f(gv) = f(v)$ for all $g \in G$, $v \in V$. We usually denote the set of invariants by $[V]^G$. In the third part we will prove Propositions 3.0.8, 3.0.11, and 3.0.12, which use the representation theory of the groups and algebras we explored in Chapter 2 to identify the irreducibles appearing in the tensor space in each statement.

3.0.1 Proof of the double centraliser theorem, and statements of Schur Weyl Duality

Let us prove the double centraliser theorem. We follow Kraft and Procesi [63].

Proof of the double centraliser theorem. Recall A is a semisimple subalgebra of $\mathrm{End}(W)$, W some complex vector space. Since A is semisimple, we can decompose W as a representation of A into:

$$W \cong \bigoplus_{i=1}^{k} U_i^{\oplus r_i}, \tag{3.2}$$

where U_i comprise an exhaustive list pairwise non-isomorphic irreducible representations of A. By Schur's lemma,

$$A' = \mathrm{End}_A(W) = \mathrm{End}_A(\oplus_{i=1}^{k} U_i^{\oplus r_i}) \cong \bigoplus_{i=1}^{k} \mathrm{Mat}_{r_i}(\mathbb{C}), \tag{3.3}$$

the last isomorphism coming from the fact that $\mathrm{Hom}_A(U_i, U_j) = \delta_{i,j}\mathbb{C}$, where $\mathrm{Hom}_A(U_i, U_j)$ is the set of algebra homomorphisms from U_i to U_j. Hence by Artin-Wedderburn, A' is semisimple, with a complete list of pairwise non-isomorphic irreducible representations given by $V_i \cong \mathbb{C}^{r_i}$, $i = 1, \dots, k$. Now

$$W = \bigoplus_{i=1}^{k} U_i \boxtimes V_i, \tag{3.4}$$

as a representation of $A \otimes A'$. The density theorem gives that $A = \bigoplus_{i=1}^{r} \mathrm{End}_{\mathbb{C}}(U_i) \cong \bigoplus_{i=1}^{k} \mathrm{Mat}_{s_i}(\mathbb{C})$, where $s_i = \dim(U_i)$. So as a representation of A',

$$W \cong \bigoplus_{i=1}^{k} V_i^{\oplus s_i}, \tag{3.5}$$

and then by a similar argument as before, $\mathrm{End}_B(W) = \mathrm{End}_B(\oplus_{i=1}^{k} V_i^{\oplus s_i}) \cong \bigoplus_{i=1}^{k} \mathrm{Mat}_{s_i}(\mathbb{C}) \cong A$, which completes the proof.

■

We can now begin to study the specific instances of the double centraliser theorem known as Schur-Weyl duality. Let us define precisely the actions of the relevant groups and algebras on tensor space, and state the theorems.

Let V be a complex vector space of dimension θ. Let $n \geq 1$. A basis $\{f_i\}_{i=1}^{\theta}$ of V gives a basis $\{f_{\underline{i}}\}$, $\underline{i} = (i_1, \dots, i_n)$, $1 \leq i_j \leq \theta$, for the space $V^{\otimes n}$. Recall in Section 2.1.5 we defined a standard basis $\{e_i\}_{i=1}^{\theta}$ of V as being its own dual $e_i = e_i^*$, $1 \leq i \leq \theta$, with respect to a non-degenerate, symmetric, bilinear product $(\cdot, \cdot)$. We call the basis $\{e_{\underline{i}}\}$ the standard basis of $V^{\otimes n}$. Consider the action $\mathfrak{p}^{GL(\theta)}$ of $GL(\theta)$ on the space $V^{\otimes n}$ given by n tensor copies of the natural representation, that is,

$$\mathfrak{p}^{GL(\theta)}(g)(v_1 \otimes \cdots \otimes v_n) = gv_1 \otimes \cdots \otimes gv_n. \tag{3.6}$$

Let us note that the corresponding action of the lie algebra $\mathfrak{gl}(\theta)$ is given by

$$X \cdot (v_1 \otimes \cdots \otimes v_n) = \sum_{i=1}^{n} v_1 \otimes \cdots \otimes Xv_i \otimes \cdots \otimes v_n, \tag{3.7}$$

for any $X \in \mathfrak{gl}(\theta)$. Consider the action $\mathfrak{p}^{S_n}$ of the symmetric group S_n on the same space as

$$\mathfrak{p}^{S_n}(\sigma)(v_1 \otimes \cdots \otimes v_n) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)}, \tag{3.8}$$

and extended linearly to $\mathbb{C}S_n$; or, written differently, (and in particular in the standard basis $\{e_{\underline{i}}\}$ of $V^{\otimes n}$), for a diagram $\sigma \in S_n$ written in edge notation (2.5), $\sigma = \prod_{j=1}^{n}(j^{\mathsf{N}}, \sigma(j)^{\mathsf{S}})$,

$$[\mathfrak{p}^{S_n}(\sigma)]_{\underline{i}_{\mathsf{N}}, \underline{i}_{\mathsf{S}}} = \prod_{j=1}^{n} \delta_{i_{j^{\mathsf{N}}}, i_{\sigma(j)^{\mathsf{S}}}}, \tag{3.9}$$

where $\underline{i}_{\mathsf{N}} = (i_{1^{\mathsf{N}}}, \ldots, i_{n^{\mathsf{N}}})$, $\underline{i}_{\mathsf{S}} = (i_{1^{\mathsf{S}}}, \ldots, i_{n^{\mathsf{S}}})$ and $1 \leq i_{t^{\xi}} \leq \theta$ for each $\xi = \mathsf{N}, \mathsf{S}$, $1 \leq t \leq n$.

Theorem 3.0.2 (Schur-Weyl Duality for the symmetric and general linear groups)**.** *The actions of $GL(\theta)$ and S_n on $W = V^{\otimes n}$ centralise each other, that is, $End_{\mathbb{C}GL(\theta)}W = \mathfrak{p}^{S_n}(\mathbb{C}S_n)$, and $End_{\mathbb{C}S_n}W = \mathfrak{p}^{GL(\theta)}(\mathbb{C}GL(\theta))$. Moreover, as a representation of $GL(\theta) \times S_n$,*

$$V^{\otimes n} \cong \bigoplus_{\substack{\rho \vdash n \\ \rho_1^{\mathsf{T}} \leq \theta}} \psi_\rho^{GL(\theta)} \boxtimes \psi_\rho^{S_n}, \tag{3.10}$$

where $\psi_\rho^{GL(\theta)}$ is the irreducible representation of $GL(\theta)$ with highest weight ρ, and $\psi_\rho^{S_n}$ is the irreducible representation of S_n corresponding to ρ.

Let $(\cdot,\cdot)$ be a non-degenerate, symmetric bilinear form on V, and recall the definition of $O(\theta)$ from Section 2.1.5. Recall also that the standard basis $\{e_i\}_{i=1}^{\theta}$ of V is its own dual, $e_i = e_i^*$, $1 \leq i \leq \theta$. We will also sometimes need the basis $\{f_i\}_{i=1}^{\theta}$ of V satisfying $f_i^* = f_{\theta+1-i}$, $1 \leq i \leq \theta$, since we used it to construct the irreducibles of the orthogonal group.

One can check that if the standard basis $\{e_i\}_{i=1}^{\theta}$ and another basis $\{f_i\}_{i=1}^{\theta}$ are related by the change of basis matrix M (i.e. $f_i = \sum_{j=1}^{\theta} M_{i,j} e_j$ for all i), then $\{f_i^*\}_{i=1}^{\theta}$ and $\{e_i\}_{i=1}^{\theta} = \{e_i^*\}_{i=1}^{\theta}$ are related by $M^{-\mathsf{T}}$, the inverse transpose of M. Then one can prove the equality

$$\sum_{i=1}^{\theta} e_i \otimes e_i = \sum_{i=1}^{\theta} f_i \otimes f_i^* \in V^{\otimes 2}; \tag{3.11}$$

or in other words, the vector $\sum_{i=1}^{\theta} f_i \otimes f_i^* \in V^{\otimes 2}$ is basis-independent.

Consider the action $\mathfrak{p}^{O(\theta)}$ of $O(\theta)$ on the space $V^{\otimes n}$ given by $\mathfrak{p}^{GL(\theta)}$ (3.6) restricted to $O(\theta)$. Note that this induces an action of $\mathfrak{so}(\theta)$ on $V^{\otimes n}$ too, as (3.7) restricted to $\mathfrak{so}(\theta)$. Consider the action $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ of the Brauer algebra $\mathbb{B}_{n,\theta}$ on the same space as $\mathfrak{p}^{\mathbb{B}_{n,\theta}}((i,j)) = T_{i,j}$, $\mathfrak{p}^{\mathbb{B}_{n,\theta}}((\overline{i,j})) = Q_{i,j}$, where

$$\begin{aligned} T_{i,j}(v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_n) &= (v_1 \otimes \cdots \otimes v_j \otimes \cdots \otimes v_i \otimes \cdots \otimes v_n) \\ Q_{i,j}(v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_n) &= (v_i, v_j) \sum_{i=1}^{\theta} (v_1 \otimes \cdots \otimes f_i \otimes \cdots \otimes f_i^* \otimes \cdots \otimes v_n), \end{aligned} \tag{3.12}$$

where we have written $Q_{i,j}$ in terms of a general basis $f_{\underline{i}}$ of $V^{\otimes n}$, and it is well defined by our observation (3.11). In particular, $Q_{i,j}$ projects $V^{\otimes 2}$ onto the one-dimensional space spanned by the vector (3.11). Notice that when restricted to $\mathbb{C}S_n$, the function $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ becomes $\mathfrak{p}^{S_n}$. Written differently (and in particular in the standard basis $e_{\underline{i}}$ of $V^{\otimes n}$), for a diagram $b \in B_n$ written in edge notation (2.5), $b = \prod_{j=1}^{n}(b_j, b'_j)$, where $b_j, b'_j \in \{t^\xi : \xi = \mathsf{N}, \mathsf{S},\ 1 \le t \le n\}$,

$$[\mathfrak{p}^{\mathbb{B}_{n,\theta}}(b)]_{\underline{i}_\mathsf{N}, \underline{i}_\mathsf{S}} = \prod_{j=1}^{n} \delta_{i_{b_j}, i_{b'_j}}, \tag{3.13}$$

where $\underline{i}_\mathsf{N} = (i_{1^\mathsf{N}}, \dots, i_{n^\mathsf{N}})$, $\underline{i}_\mathsf{S} = (i_{1^\mathsf{S}}, \dots, i_{n^\mathsf{S}})$ and $1 \le i_{t^\xi} \le \theta$ for each $\xi = \mathsf{N}, \mathsf{S}$, $1 \le t \le n$.

Theorem 3.0.3 (Schur-Weyl Duality for the Brauer algebra and Orthogonal group)**.** *The actions of $O(\theta)$ and $\mathbb{B}_{n,\theta}$ on $W = V^{\otimes n}$ centralise each other, that is, $End_{\mathbb{C}O(\theta)} W = \mathfrak{p}^{\mathbb{B}_{n,\theta}}(\mathbb{B}_{n,\theta})$, and $End_{\mathbb{B}_{n,\theta}} W = \mathfrak{p}^{O(\theta)}(\mathbb{C}O(\theta))$. Moreover, as a representation of $\mathbb{C}O(\theta) \otimes \mathbb{B}_{n,\theta}$,*

$$V^{\otimes n} \cong \bigoplus_{k=0}^{\lfloor \frac{n}{2} \rfloor} \bigoplus_{\substack{\lambda \vdash n-2k \\ \lambda_1^\mathsf{T} + \lambda_2^\mathsf{T} \le \theta}} \psi_\lambda^{O(\theta)} \boxtimes \psi_\lambda^{\mathbb{B}_{n,\theta}}, \tag{3.14}$$

where $\psi_\lambda^{O(\theta)}$, $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ are pairwise non-isomorphic irreducible representations of $O(\theta)$ and $\mathbb{B}_{n,\theta}$ respectively, each corresponding to the partition λ.

Remark 3.0.4. Let us note that in this work we will not quite prove that the representation $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ that appears in Theorem 3.0.3 is the irreducible of $\mathbb{B}_{n,\theta}$ defined in Theorem 2.1.11; we will only prove that the $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ in Theorem 3.0.3 are pairwise non-isomorphic irreducibles of $\mathbb{B}_{n,\theta}$, and that each $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ is a quotient of the cell module $\Delta_\lambda^{\mathbb{B}_{n,\theta}}$. This will be enough for the purposes of the applications of Schur-Weyl duality in Sections 5 and 6.

Recall that for a linear map $A : V \to V$, the adjoint of A with respect to $(\cdot, \cdot)$ is the linear map A^* satisfying $(Av, u) = (v, A^* u)$ for all $v, u \in V$. If A has matrix A_f in the basis $\{f_i\}_{i=1}^{\theta}$, then A^* has matrix A_f^T (the transpose of A_f) in the basis $\{f_i^*\}_{i=1}^{\theta}$. So, with both written in the standard basis, the matrix of A^* is the transpose of the matrix of A.

Let $n, m \in \mathbb{N}$ with $0 \le m < n$. Consider the action $\mathfrak{q}^{GL(\theta)}$ of $GL(\theta)$ on the space $V^{\otimes m} \otimes (V^*)^{\otimes n-m}$ given by m tensor copies of the natural representation, and $n-m$ tensor copies of the dual of the natural representation, that is,

$$\mathfrak{q}^{GL(\theta)}(g)(v_1 \otimes \cdots \otimes v_m \otimes v_{m+1} \otimes \cdots \otimes v_n) = (gv_1 \otimes \cdots \otimes gv_m \otimes g^{-*} v_{m+1} \otimes \cdots \otimes g^{-*} v_n), \tag{3.15}$$

where g^{-*} is the adjoint of g^{-1}. In the standard basis this is just $g^{-\mathsf{T}}$, the inverse transpose of g (or, if we assume that g^{-*} is written in the dual basis of whichever basis we are using for V, then g^{-*} is the inverse-transpose of the matrix of g). Note that the action of the lie algebra $\mathfrak{gl}(\theta)$ is therefore the same as (3.7), but with $X \in \mathfrak{gl}(\theta)$ acting as $-X^*$ on the tensor factors from V^*. For example, in a given basis $\{f_i\}_{i=1}^{\theta}$, under the action of a diagonal matrix $H = (h_{i,i})_{i=1}^{\theta} \in \mathfrak{gl}(\theta)$, the basis vector f_i is scaled by $h_{i,i}$, and the vector f_i^* is scaled by $-h_{i,i}$. Consider the action $\mathfrak{p}^{\mathbb{B}_{n,m,\theta}}$ of the walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ on the same space as the restriction of $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ from the Brauer algebra $\mathbb{B}_{n,\theta}$ to the walled Brauer algebra.

Theorem 3.0.5 (Schur-Weyl Duality for the walled Brauer algebra)**.** *The actions of $GL(\theta)$ and $\mathbb{B}_{n,m,\theta}$ on $W = V^{\otimes m} \otimes (V^*)^{\otimes n-m}$ centralise each other, that is, we have $End_{\mathbb{C}GL(\theta)} W = \mathfrak{p}^{\mathbb{B}_{n,m,\theta}}(\mathbb{B}_{n,m,\theta})$, and $End_{\mathbb{B}_{n,m,\theta}} W = \mathfrak{q}^{GL(\theta)}(\mathbb{C}GL(\theta))$. Moreover, as a rep-*

resentation of $\mathbb{C}GL(\theta) \otimes \mathbb{B}_{n,m,\theta}$*,*

$$V^{\otimes m} \otimes (V^*)^{\otimes n-m} \cong \bigoplus_{k=0}^{\min\{m,n-m\}} \bigoplus_{\substack{\lambda \vdash m-k,\ \mu \vdash n-m-k \\ \lambda_1^\top + \mu_1^\top \le \theta}} \psi^{GL(\theta)}_{[\lambda,\mu]} \boxtimes \psi^{\mathbb{B}_{n,m,\theta}}_{(\lambda,\mu)}, \tag{3.16}$$

where $\psi^{GL(\theta)}_{[\lambda,\mu]}$ *is the irreducible representation of* $GL(\theta)$ *with highest weight* $[\lambda,\mu]$*, and* $\psi^{\mathbb{B}_{n,m,\theta}}_{(\lambda,\mu)}$ *is the irreducible representation of* $\mathbb{B}_{n,m,\theta}$ *corresponding to the pair* (λ,μ)*.*

Remark 3.0.6. Let us note that, similarly to the Brauer algebra case, in this work we will not quite prove that the representation $\psi^{\mathbb{B}_{n,m,\theta}}_{(\lambda,\mu)}$ that appears in Theorem 3.0.5 is the irreducible of $\mathbb{B}_{n,m,\theta}$ defined in Theorem 2.1.13; we will only prove that the $\psi^{\mathbb{B}_{n,m,\theta}}_{(\lambda,\mu)}$ in Theorem 3.0.5 are pairwise non-isomorphic irreducibles of $\mathbb{B}_{n,m,\theta}$, and that each $\psi^{\mathbb{B}_{n,m,\theta}}_{(\lambda,\mu)}$ is a quotient of the cell module $\Delta^{\mathbb{B}_{n,m,\theta}}_{(\lambda,\mu)}$. This will be enough for the purposes of the applications of Schur-Weyl duality in Sections 5 and 6.

As noted earlier in this section, we will present proofs of each of these Theorems 3.0.2, 3.0.3, and 3.0.5, each coming in two parts. In each case, the first part, in Section 3.0.2, proves a decomposition of the form (3.1) using the double centraliser theorem and the specific irreducible representations in the decomposition are not identified; we only prove that they are irreducible and pairwise non-isomorphic. In order to employ the double centraliser theorem, it needs to be proved in each case that the action of one of the algebras (resp. groups) centralises the other. This is proved using invariant theory, in particular (the multilinear versions of) the First Fundamental Theorems of invariant theory for $GL(\theta)$ and $O(\theta)$, Theorems 3.0.9 and 3.0.10. Theorem 3.0.10 is the only statement which we do not prove ourselves, since it would require a lengthy detour into invariant theory.

The second part of each of the proofs explicitly constructs and thereby identifies the irreducibles, making use of Young symmetrisers, the Specht modules, and the highest weight theory of Chapter 2. These second parts will be proved in Section 3.0.3. For the invariant theory part we mainly follow Kraft and Procesi [63], and for the second part we follow Benkart, Britten and Lemire [43], and Benkart et al. [8], with some ideas from Goodman and Wallach [44].

3.0.2 Invariant theory proofs of Schur-Weyl duality

Proof of Theorem 3.0.2. In order to use the double centraliser theorem 3.0.1, we need to show that $\mathrm{End}_{S_n} W = \mathfrak{p}^{GL(\theta)}(\mathbb{C}GL(\theta))$. By inspection, it is clear that the right hand side is contained in the left; it remains to prove that this containment is equality. As noted above, the rest of Theorem 3.0.2 (minus the identification of the irreducibles) follows from the double centraliser theorem. The following lemma follows the lemma in Section 3.1 of [63].

Lemma 3.0.7. *Let* V *be a finite dimensional complex vector space. Then the linear span of the tensors* $v \otimes \cdots \otimes v \in V^{\otimes n}$ *is the subspace* Σ_n *of all tensors invariant under the action* $\mathfrak{p}^{S_n}$ *of* S_n*.*

Proof. Let $\{f_i\}_{i=1}^{\theta}$ be a basis of V; then $\{f_{\underline{i}} = f_{(i_1,\dots,i_n)} \ :\ 1 \le i_j \le \theta,\ 1 \le j \le n\}$ is a basis of $V^{\otimes n}$ which is stable under the action of S_n. Each orbit of this action has a unique representative of the form $f_1^{h_1} \otimes \cdots \otimes f_\theta^{h_\theta}$, where $h_1 + \cdots + h_\theta = n$. Let $r_{h_1,\ldots,h_\theta}$ denote the sum of the elements in this orbit; the set of these $r_{h_1,\ldots,h_\theta}$ is a basis Σ_n. We need to show that the tensors $v \otimes \cdots \otimes v$ span this space. It suffices to show that any linear functional $\eta : \Sigma_n \to \mathbb{C}$ which is zero on all tensors $v \otimes \cdots \otimes v$ is the zero functional. Let $v = \sum_{i=1}^{\theta} v_i f_i$, then $v \otimes \cdots \otimes v = \sum_{\sum h_i = n} v_1^{h_1} \cdots v_\theta^{h_\theta} r_{h_1,\ldots,h_\theta}$, so

$$\eta(v^{\otimes n}) = \sum_{\sum h_i = n} a_{h_1,\ldots,h_\theta} v_1^{h_1} \cdots v_\theta^{h_\theta}, \tag{3.17}$$

where $a_{h_1,\ldots,h_\theta} := \eta(r_{h_1,\ldots,h_\theta})$. Hence $\eta(v^{\otimes n})$ can be viewed as a polynomial in the coefficients $v_1, \ldots, v_\theta$; by assumption it vanishes on all $v^{\otimes n}$, so it must be the zero polynomial, so each $a_{h_1,\ldots,h_\theta}$ must be zero, so η must be the zero functional. ∎

Now note that the algebra $\mathrm{End}(V^{\otimes n})$ is canonically isomorphic to $\mathrm{End}(V)^{\otimes n}$. This isomorphism induces a map from $GL(\theta)$ to $\mathrm{End}(V)^{\otimes n}$: $g \in GL(\theta)$ sent to $g^{\otimes n}$; and an action of S_n on $\mathrm{End}(V)^{\otimes n}$: $\sigma \in S_n$ acts as $\sigma(A_1 \otimes \cdots \otimes A_n) = A_{\sigma^{-1}(1)} \otimes \cdots \otimes A_{\sigma^{-1}(n)}$. Now the isomorphism of $\mathrm{End}(V^{\otimes n})$ and $\mathrm{End}(V)^{\otimes n}$ induces an isomorphism between the set Σ_n of maps in $\mathrm{End}(V)^{\otimes n}$ invariant under this action of S_n, and $\mathrm{End}_{S_n} W$. Then by the lemma, the span of the image of the action of $GL(\theta)$ is exactly Σ_n. The statement $\mathrm{End}_{\mathbb{C}GL(\theta)} W = \mathfrak{p}^{S_n}(\mathbb{C}S_n)$ follows.

Now by the double centraliser theorem 3.0.1, we have

$$V^{\otimes n} = \bigoplus_{i=1}^{k} \psi_i^{GL(\theta)} \boxtimes \psi_i^{S_n}, \tag{3.18}$$

where $\psi_i^{GL(\theta)}$ (resp. $\psi_i^{S_n}$) are an exhaustive list of pairwise non-isomorphic representations of $\mathfrak{p}^{GL(\theta)}(\mathbb{C}GL(\theta))$ (resp. $\mathfrak{p}^{S_n}(\mathbb{C}S_n)$). These algebras are quotients of $\mathbb{C}GL(\theta)$ and $\mathbb{C}S_n$ respectively, so $\psi_i^{GL(\theta)}$ (resp. $\psi_i^{S_n}$) are a (possibly not exhaustive) list of pairwise non-isomorphic representations of $GL(\theta)$ (resp. S_n). Theorem 3.0.2 now follows from the following proposition, which identifies the specific irreducible representations $\psi_i^{GL(\theta)}$ and $\psi_i^{S_n}$. ∎

Proposition 3.0.8. *The irreducible representations $\psi_i^{GL(\theta)}$ (resp. $\psi_i^{S_n}$) appearing in (3.18) are the irreducible representations $\psi_\rho^{GL(\theta)}$ with highest weight ρ (resp. $\psi_\rho^{S_n}$) where ρ runs over all partitions of n with at most θ parts.*

This proposition is the second part of the proof of Theorem 3.0.2. We will prove it in Section 3.0.3.

Our working so far allows us to prove (the multilinear version of) the First Fundamental Theorem of invariant theory (for $GL(\theta)$), which we will use in the proof of the walled Brauer algebra version of Schur-Weyl duality. As noted earlier, invariant theory studies functions on a vector space which are invariant under the action of a group. Specifically, let $G = GL(\theta)$ or $O(\theta)$ act on $W \oplus W^* = V^{\oplus n} \oplus (V^*)^{\oplus n}$ as n direct summands of the natural representation, and n direct summands of its dual. (Note the dual of the natural representation of $O(\theta)$ is itself). Then an invariant for G on $W \oplus W^*$ is a function $f : W \oplus W^* \to \mathbb{C}$ which lies in the algebra generated by the linear functionals $(W \oplus W^*)^*$

(i.e. polynomials of functionals), which satisfies $f(gv) = f(v)$ for all $g \in G$ and $v \in W \oplus W^*$. We denote the set of invariants by $[W \oplus W^*]^G$. An invariant is multilinear if it is linear in each of the $2n$ arguments from the $2n$ direct summands.

Theorem 3.0.9 (Multilinear FFT for $GL(\theta)$ [104])**.** *Let $n \in \mathbb{N}$. Let $GL(\theta)$ act on $V^{\oplus n} \oplus (V^*)^{\oplus n}$ as n direct summands of the natural representation, and n direct summands of its dual. Let $(\cdot,\cdot)_{i^N, j^S}(v_{1^N} \oplus \cdots \oplus v_{n^N} \oplus w_{1^S} \oplus \cdots \oplus w_{n^S}) = (v_{i^N}, w_{j^S}) := w_{j^S}(v_{i^N})$ for each $1 \le i \le n$, $1 \le j \le n$, $v_{i^N} \in V$, $w_{j^S} \in V^*$. Then the space of multilinear $GL(\theta)$ invariants on $V^{\oplus n} \oplus (V^*)^{\oplus n}$ is spanned by $\overline{\sigma} = \prod_{j=1}^n (\cdot,\cdot)_{j^N, \sigma(j)^S}$, where $\sigma \in S_n$.*

Proof. We follow Section 4 of [63]. Let $W = V^{\otimes n}$ and $W_+ = V^{\oplus n}$. Let $GL(\theta)$ act on $W \otimes W^*$ as n tensor copies of the natural representation and n tensor copies of its dual; let $GL(\theta)$ act on $W \oplus W^*$ the same, but with tensor products replaced with direct sums. We have two canonical isomorphisms of vector spaces:

$$\mathrm{End}_{\mathbb{C}GL(\theta)} W \cong [(W \otimes W^*)^*]^{GL(\theta)} \cong \mathbb{C}[W_+ \oplus W_+^*]^{GL(\theta)}_{\text{multi}}, \tag{3.19}$$

where $[(W \otimes W^*)^*]^{GL(\theta)}$ is the set of linear functionals on $W \otimes W^*$ that are invariant under the action of $GL(\theta)$, and $\mathbb{C}[W_+ \oplus W_+^*]^{GL(\theta)}_{\text{multi}}$ is the set of multilinear polynomials on $W_+ \oplus W_+^*$ invariant under the action of $GL(\theta)$. The first isomorphism is given by $\alpha_0 : \mathrm{End}_{\mathbb{C}GL(\theta)} W \to [(W \otimes W^*)^*]^{GL(\theta)}$ as $\alpha_0(A)(v \otimes w^*) = w^*(A(v))$. It is straightforward to check that α_0 is linear. With the standard basis $\{e_{\underline{i}}\}$ of V, we can write its inverse as $[\alpha_0^{-1}(f)]_{\underline{i}_N, \underline{i}_S} = f(e_{\underline{i}_N} \otimes e_{\underline{i}_S})$ (and so in particular, α_0 is a bijection). The second isomorphism is given by $\alpha_1 : \mathbb{C}[W_+ \oplus W_+^*]^{GL(\theta)}_{\text{multi}} \to [(W \otimes W^*)^*]^{GL(\theta)}$ as

$$\alpha_1(F)(v_1 \otimes \cdots \otimes v_n \otimes w_1 \otimes \cdots \otimes w_n) = F(v_1 \oplus \cdots \oplus v_n \oplus w_1 \oplus \cdots \oplus w_n); \tag{3.20}$$

once again α_1 is linear and its inverse is straightforward to write down. Now from our working above in the proof of Theorem 3.0.2, the leftmost expression in (3.19) is $\mathfrak{p}^{S_n}(\mathbb{C}S_n)$, and, using (3.9), it is straightforward to check that under the map $\alpha_1^{-1} \circ \alpha_0$, the elements $\mathfrak{p}^{S_n}(\sigma)$, $\sigma \in S_n$, are precisely the maps $\overline{\sigma} = \prod_{j=1}^n (\cdot,\cdot)_{j^N, \sigma(j)^S}$. This completes the proof of Theorem 3.0.9. ∎

Let us now state the equivalent statement for the Orthogonal group. This theorem is due to Weyl [104]. We will not give a proof for this theorem, as it would require too long a deviation into invariant theory; Kraft and Procesi [63] gives a full proof. Note that for the Orthogonal group, the natural representation is self-dual, as each $g \in O(\theta)$ satisfies, in the standard basis, $g^{-\mathsf{T}} = g$. Recall that we can identify V and V^* via the canonical isomorphism L (2.17), so the product (v, w) makes sense for any $v, w \in V$ or V^*.

Theorem 3.0.10 (FFT for $O(\theta)$ [104])**.** *Let $n \in \mathbb{N}$. Let $O(\theta)$ act on $V^{\oplus n} \oplus (V^*)^{\oplus n}$ as $2n$ direct summands of the natural representation. Let $(\cdot,\cdot)_{\xi,\eta}(v_{1^N} \oplus \cdots \oplus v_{n^N} \oplus v_{1^S} \oplus \cdots \oplus v_{n^S}) = (v_\xi, w_\eta)$ for each $\xi, \eta = i^{\mathsf{N}}$ or i^{S}, $1 \le i \le n$, $v_{i^N} \in V$, $v_{j^S} \in V^*$. Then the space of multilinear $O(\theta)$ invariants on $V^{\oplus n} \oplus (V^*)^{\oplus n}$ is spanned by $\overline{b} = \prod_{j=1}^n (\cdot,\cdot)_{b_j, b'_j}$, where $b \in B_n$, and $b = \prod_{j=1}^n (b_j, b'_j)$ written in edge notation (2.5).*

We can now prove (the first parts of) Theorems 3.0.3 and 3.0.5, the Schur-Weyl dualities for the Brauer and walled Brauer algebras. We follow Kraft and Procesi [63] and Lemma

1.2 of Koike [60].

Proof of Theorems 3.0.3 and 3.0.5. Let $W = V^{\otimes m} \otimes (V^*)^{\otimes n-m} = V^{\otimes m} \otimes V^{\otimes n-m}$ and $W_+ = V^{\oplus m} \oplus (V^*)^{\oplus n-m} = V^{\oplus m} \oplus V^{\oplus n-m}$. Recall that $GL(\theta)$ and $O(\theta)$ act on W (resp. W_+) as m tensor powers (resp. direct summands) of the natural representation and $n-m$ tensor powers (resp. direct summands) of its dual (where, note, the natural representation of $O(\theta)$ is its own dual).

The proofs follow the structure of that of Theorem 3.0.2: we first prove the statements $\mathrm{End}_{\mathbb{C}O(\theta)} W = \mathfrak{p}^{\mathbb{B}_{n,\theta}}(\mathbb{B}_{n,\theta})$ and $\mathrm{End}_{\mathbb{C}GL(\theta)} W = \mathfrak{p}^{\mathbb{B}_{n,m,\theta}}(\mathbb{B}_{n,m,\theta})$, and then the double centraliser theorem 3.0.1 gives us the rest, minus the identification of the irreducibles in the decompositions of W. Note that for each statement $\mathrm{End}_{\mathbb{C}O(\theta)} W = \mathfrak{p}^{\mathbb{B}_{n,\theta}}(\mathbb{B}_{n,\theta})$ and $\mathrm{End}_{\mathbb{C}GL(\theta)} W = \mathfrak{p}^{\mathbb{B}_{n,m,\theta}}(\mathbb{B}_{n,m,\theta})$, one can straightforwardly check that the right hand side is contained in the left - what remains is to show that this containment is equality. The proofs follow our proof above of the First Fundamental Theorem 3.0.9, but in reverse. Indeed, we have two canonical isomorphisms of vector spaces:

$$\mathrm{End}_{\mathbb{C}G} W \cong [(W \otimes W^*)^*]^G \cong \mathbb{C}[W_+ \oplus W_+^*]^G_{\text{multi}}, \tag{3.21}$$

where $G = GL(\theta)$ or $O(\theta)$, and the invariance is with respect to the actions of $GL(\theta)$ and $O(\theta)$ described above. The isomorphisms are the same as the α_0 and α_1 described in the proof of Theorem 3.0.2; in particular for $f \in [(W \otimes W^*)^*]^G$, $\alpha_0^{-1}(f) \in \mathrm{End}_{\mathbb{C}G} W$ as, in the standard basis, $[\alpha_0^{-1}(f)]_{\underline{i}_N, \underline{i}_S} = f(e_{\underline{i}_N} \otimes e_{\underline{i}_S})$. Now in the $O(\theta)$ case, the First Fundamental Theorem 3.0.10 gives a basis for the right hand side of (3.21) as $\bar{b} = \prod_{j=1}^n (\cdot,\cdot)_{b_j, b'_j}$, where $b \in B_n$. Now passing these functions through the isomorphism $\alpha_0^{-1} \circ \alpha_1$ gives precisely $\mathfrak{p}^{\mathbb{B}_{n,\theta}}(b)$, $b \in B_n$; see the edge notation version of $\mathfrak{p}^{\mathbb{B}_{n,\theta}}(b)$ (3.13); this gives $\mathrm{End}_{\mathbb{C}O(\theta)} W = \mathfrak{p}^{\mathbb{B}_{n,\theta}}(\mathbb{B}_{n,\theta})$. An identical proof can show $\mathrm{End}_{\mathbb{C}GL(\theta)} W = \mathfrak{p}^{\mathbb{B}_{n,m,\theta}}(\mathbb{B}_{n,m,\theta})$, one only needs to modify the First Fundamental Theorem 3.0.9 for $GL(\theta)$ by rearranging the summands V and V^* to obtain:

$$\mathbb{C}[W_+ \oplus W_+^*]^{GL(\theta)}_{\text{multi}} = \mathrm{span}\left\{ \bar{b} = \prod_{j=1}^{n} (\cdot,\cdot)_{b_j, b'_j} \ : \ b \in B_{n,m} \right\}. \tag{3.22}$$

Now the double centraliser theorem 3.0.1 gives us, in the two cases $(G, A) = (O(\theta), \mathbb{B}_{n,\theta})$ and $(GL(\theta), \mathbb{B}_{n,m,\theta})$,

$$V^{\otimes n} = \bigoplus_{i=1}^{k} \psi_i^G \boxtimes \psi_i^A, \tag{3.23}$$

where ψ_i^G (resp. ψ_i^A) are a (possibly not exhaustive) list of pairwise non-isomorphic representations of G (resp. A). Theorems 3.0.3 and 3.0.5 now follow from the following two propositions, which will be proved in the following Section 3.0.3. ∎

Proposition 3.0.11. *The irreducible representations $\psi_i^{O(\theta)}$ (resp. $\psi_i^{\mathbb{B}_{n,\theta}}$) appearing in (3.23) are the irreducible representations $\psi_\lambda^{O(\theta)}$ (resp. $\psi_\lambda^{\mathbb{B}_{n,\theta}}$) where λ runs over all partitions of $n-2k$, $0 \le k \le \lfloor \frac{n}{2} \rfloor$, and $\lambda_1^\top + \lambda_2^\top \le \theta$.*

Proposition 3.0.12. *The irreducible representations $\psi_i^{GL(\theta)}$ (resp. $\psi_i^{\mathbb{B}_{n,m,\theta}}$) appearing in (3.23) are the irreducible representations $\psi_{[\lambda,\mu]}^{GL(\theta)}$ with highest weight $[\lambda,\mu]$ (resp. $\psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}$) where λ,μ run over all partitions of $m-k, n-m-k$, $0 \le k \le \min\{m, n-m\}$, with $\lambda_1^\top + \mu_1^\top \le \theta$.*

3.0.3 Constructing irreducibles in Schur-Weyl duality

In this section, we present constructive proofs of Propositions 3.0.8, 3.0.11, and 3.0.12. We follow and adapt the arguments of Benkart, Britten and Lemire [43], Benkart et al. [8], and Goodman and Wallach [44]. In the first two of those works, in the Brauer and walled Brauer cases, the theorems are proved in the cases $\theta \geq n$. As noted in Section 2.1.3 and 2.1.4, in this range the Brauer and walled Brauer algebras are semisimple, and we have a full description of their irreducible representations - they are the cell modules from Lemmas 2.1.11 and 2.1.13. In the general case, the irreducibles are quotients of the cell modules. Goodman and Wallach prove statements for arbitrary θ, n, but only study the image of the Brauer algebra in tensor space, not the full algebra.

While the three proofs of the Propositions are very similar, we present the proof of Proposition 3.0.8 first, as it is simpler and serves as a prototype for the other two. For the rest of this section we will use the specific basis of V, $\{f_i\}_{i=1}^{\theta}$, where $f_i^* = f_{\theta+1-i}$ for all i, as this is the basis used when we defined our representations of the classical groups in Section 2.1.5.

Proof of 3.0.8. Recall the definition of the Young symmetriser z_τ (2.4), for $\tau \in \mathcal{ST}(\mathcal{N})$ the set of standard Young tableaux with entries $\mathcal{N} = \{1, \dots, n\}$ and some shape ρ, ρ a partition of n. Recall (2.1.9) we have

$$\mathbb{C}S_n = \bigoplus_{\tau \in \mathcal{ST}(\mathcal{N})} z_\tau \mathbb{C}S_n, \tag{3.24}$$

where each $z_\tau \mathbb{C}S_n$ is a minimal right ideal of $\mathbb{C}S_n$. Now,

$$V^{\otimes n} = \mathbb{C}S_n V^{\otimes n} = \sum_{\tau \in \mathcal{ST}(\mathcal{N})} z_\tau V^{\otimes n}. \tag{3.25}$$

Consider the vector

$$z_\tau \beta_\tau \in z_\tau V^{\otimes n}, \tag{3.26}$$

where $\beta_\tau = f_{i_1} \otimes \cdots \otimes f_{i_n}$, where i_j is equal to the row in which the number j appears in τ. Let us break down what we will prove.

Lemma 3.0.13. *1. If $z_\tau V^{\otimes n}$ is non-zero, then it is the irreducible module $\psi_\rho^{GL(\theta)}$ of $GL(\theta)$ with highest weight ρ, where ρ is the shape of τ;*

2. The space $z_\tau V^{\otimes n}$ is non-zero if and only if ρ has at most θ parts, where ρ is the shape of τ;

3. The vectors $z_\tau \beta_\tau$, as τ ranges over $\mathcal{ST}(\mathcal{N})$ of shape with at most θ parts, are linearly independent;

4. The space M_ρ spanned by the vectors $z_\tau \beta_\tau$ with τ shape ρ is a copy of the irreducible $\psi_\rho^{S_n}$ of S_n.

Together, Parts *1* to *3* Lemma 3.0.13 shows us that the sum in (3.25) is direct (over τ with shape having at most θ parts), and that it is the decomposition of $V^{\otimes n}$ into irreducibles of $GL(\theta)$. Adding part *4* gives us the decomposition of $V^{\otimes n}$ from Theorem 3.0.2.

Proof. Let us prove part *1* from the lemma. Indeed, since the actions of $GL(\theta)$ and S_n commute, $z_\tau V^{\otimes n}$ is indeed a $GL(\theta)$ module. If it is reducible, then there must be a projection u onto a non-zero submodule. This u commutes with the action of $GL(\theta)$, so since we have already proved $\mathrm{End}_{\mathbb{C}GL(\theta)}W = \mathfrak{p}^{S_n}(\mathbb{C}S_n)$, it must be the action of some $u' \in \mathbb{C}S_n$. Now

$$0 \neq uz_\tau V^{\otimes n} \subsetneq z_\tau V^{\otimes n}, \tag{3.27}$$

which implies, by applying the z_τ on the left, which is the identity map on $z_\tau V^{\otimes n}$,

$$0 \neq z_\tau u z_\tau V^{\otimes n} \subsetneq z_\tau V^{\otimes n}. \tag{3.28}$$

Then $0 \neq z_\tau u z_\tau \mathbb{C}S_n \subsetneq z_\tau \mathbb{C}S_n$, which contradicts $z_\tau \mathbb{C}S_n$ being a minimal right ideal of $\mathbb{C}S_n$. So $z_\tau V^{\otimes n}$ is indeed an irreducible of $GL(\theta)$.

Recall the highest weight theorem (2.1.16). We can show that $z_\tau \beta_\tau$ is a highest weight vector in the representation $z_\tau V^{\otimes n}$ with weight ρ. A simple calculation shows that the weight of β_τ is ρ, and since the action of $GL(\theta)$ commutes with z_τ, $z_\tau\beta_\tau$ also has weight ρ. To prove $z_\tau\beta_\tau$ is maximal, it suffices to show that it is killed by each weight vector $E_{i,j} \in \mathfrak{gl}(\theta)$. We have that $E_{i,j}\beta_\tau$ is zero, or a sum of tensors β'_τ that look like β_τ, but with one f_j changed to f_i; for each $\sigma \in \mathcal{R}(\tau)$, and each β', there is some $(x, y) \in \mathcal{C}(\tau)$ which leaves $\sigma\beta'_\tau$ invariant - this gives $E_{i,j}z_\tau\beta_\tau = z_\tau E_{i,j}\beta_\tau = \sum z_\tau\beta'_\tau = 0$. Hence $z_\tau\beta_\tau$ is a highest weight vector of weight ρ, and so by the highest weight theorem $z_\tau V^{\otimes n}$ is a copy of $\psi_\rho^{GL(\theta)}$.

Let us prove part *2* of the lemma, that $z_\tau V^{\otimes n} = 0$ if and only if τ has shape ρ, and ρ has more than θ parts. Indeed, notice that $z_\tau V^{\otimes n}$ is a set of tensors which are antisymmetric in the indices which appear in the first column of τ, so if ρ has more than θ parts, then $z_\tau V^{\otimes n} = 0$. Say ρ has at most θ parts. Notice that all the $\sigma \in \mathcal{R}(\tau)$ fix β_τ, and only $\mathrm{id} \in \mathcal{C}(\tau)$ does; the rest of $\sigma \in \mathcal{C}(\tau)$ sends β_τ to other basis vectors of $V^{\otimes n}$. Now expanding $z_\tau\beta_\tau$ in the basis, we see that the coefficient of β_τ is $|\mathcal{R}(\tau)|$, so $z_\tau\beta_\tau \neq 0$, and so $z_\tau V^{\otimes n} \neq 0$.

Let us prove part *3* of the lemma, that these highest weights $z_\tau\beta_\tau$ (the ones which are non-zero) are linearly independent. Assume

$$\sum_{\substack{\rho \vdash n \\ \rho_1^\top \leq \theta}} \sum_{\tau \in \mathcal{ST}(\rho)} a_\tau z_\tau \beta_\tau = 0, \tag{3.29}$$

where the a_τ are a collection of complex coefficients. If τ, τ' have different shapes, then β_τ and $\beta_{\tau'}$ have different tensor factors, and since the symmetrisers just permute the factors, we can see that the sums of terms with different shape diagrams are linearly independent. So assume

$$\sum_{\tau \in \mathcal{ST}(\rho)} a_\tau z_\tau \beta_\tau = 0, \tag{3.30}$$

for some $\rho \vdash n$, $\rho_1^\top \leq \theta$. Let τ' be the smallest τ of shape ρ in the ordering $<$ with $a_{\tau'} \neq 0$, and recall Lemma (2.1.8) regarding this ordering. Then $0 = z_{\tau'}0 = \sum_{\tau \in \mathcal{ST}(\rho)} a_\tau z_{\tau'} z_\tau \beta_\tau = a_{\tau'} z_{\tau'} \beta_{\tau'}$, since $z_\tau z_{\tau'} = 0$ if $\tau > \tau'$. Now $a_{\tau'} = 0$, which is a contradiction; hence all $a_\tau = 0$.

Part *4* of the lemma follows from the definition of the irreducible representation $\psi_\rho^{S_n}$ and part *3* of the lemma. ∎

Hence $\bigoplus_{\tau \in \mathcal{ST}(\rho)} z_\tau V^{\otimes n} \cong \psi_\rho^{GL(\theta)} \boxtimes \psi_\rho^{S_n}$, which gives the decomposition:

$$V^{\otimes n} = \bigoplus_{\substack{\rho \vdash n \\ \rho_1^\top \leq \theta}} \bigoplus_{\tau \in \mathcal{ST}(\rho)} z_\tau V^{\otimes n} \cong \bigoplus_{\substack{\rho \vdash n \\ \rho_1^\top \leq \theta}} \psi_\rho^{GL(\theta)} \boxtimes \psi_\rho^{S_n}, \tag{3.31}$$

which completes the proof of Proposition 3.0.8, and therefore also Theorem 3.0.2.

∎

The proofs of the Brauer and walled Brauer cases, Propositions 3.0.11 and 3.0.12, follow the basic structure of that of 3.0.8, with two exceptions: that the decomposition of $V^{\otimes n}$ is more complicated, and that the equivalent highest weight vectors to $z_\tau \beta_\tau$ are no longer necessarily linearly independent. The latter reflects the fact that for $\theta < n$, the cell modules of the Brauer algebra are sometimes no longer irreducible. The key step in the proofs is showing that the decomposition is essentially given by decomposing a subspace of tensor space known as the traceless or harmonic tensors. In order to define these tensors, and to describe how the decomposition of tensor space is different from the S_n case, let us introduce some notation.

Let $\mathcal{Q}(k)$ be the set of lists $(\underline{t}, \underline{t}')$, where $\underline{t} = (t_1, \dots, t_k)$, $\underline{t}' = (t_1', \dots, t_k')$, the $t_i, t_i' \in \mathcal{N} = \{1, \dots, n\}$ all distinct, t_i increasing in i and $t_i < t_i'$. Let $\mathcal{Q}'(k)$ be the set of $(\underline{t}, \underline{t}') \in \mathcal{Q}(k)$ such that $t_i \leq m < t_i'$, for all $i = 1, \dots, k$. For a pair $(\underline{t}, \underline{t}') \in \mathcal{Q}(k)$, let $(\underline{t}, \underline{t}')^c$ be the set of elements of $\mathcal{N}$ which do not appear in $\underline{t}$ or $\underline{t}'$. For $(\underline{t}, \underline{t}') \in \mathcal{Q}(k)$, define $Q_{\underline{t}, \underline{t}'} = Q_{t_1, t_1'} \cdots Q_{t_k, t_k'}$. This is the action of the diagram $(\overline{\underline{t}, \underline{t}'}) \coloneqq \prod_{i=1}^s (\overline{t_i, t_i'}) \in B_n$ on tensor space. Note if $t = (i)$, $t' = (j)$ then $Q_{\underline{t}, \underline{t}'} = Q_{i,j}$, and note that $Q_{\underline{t}, \underline{t}'}$ arises from the action of an element of the walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ if and only if $(\underline{t}, \underline{t}') \in \mathcal{Q}'(k)$.

For ease of notation, let $W_n = V^{\otimes n}$ and $W_{n,m} = V^{\otimes m} \otimes (V^*)^{\otimes n-m}$. Let W_n^k be the span of all $Q_{\underline{t}, \underline{t}'} W_n$ with $|t| = |t'| = k$. Note this is the image of $\mathbb{B}_{n,\theta}^k$ in W_n, where recall $\mathbb{B}_{n,\theta}^k$ is the span of diagrams in the Brauer algebra $\mathbb{B}_{n,\theta}$ with at least k bars. Then let $[W_n]^k$ be the subspace of W_n^k which is killed by any $Q_{\underline{t}, \underline{t}'}$ with $|t| = |t'| = k + 1$, or equivalently, any $Q_{i,j}$ with $i, j \in (\underline{t} \cup \underline{t}')^c$. Then $W_n^k = [W_n]^k \oplus W_n^{k+1}$. The sets $W_{n,m}^k$ and $[W_{n,m}]^k$ are defined similarly (i.e. only $(\underline{t}, \underline{t}') \in \mathcal{Q}'(k)$ are allowed to act on $W_{n,m}$), and we have $W_{n,m}^k = [W_{n,m}]^k \oplus W_{n,m}^{k+1}$. Note we define $[W_n]^{\lfloor \frac{n}{2} \rfloor}$ as just $W_n^{\lfloor \frac{n}{2} \rfloor}$ (and $[W_{n,m}]^{\min\{m,n-m\}}$ as $W_{n,m}^{\min\{m,n-m\}}$). This provides direct sum decompositions

$$\begin{aligned} W_n &= \bigoplus_{k=0}^{\lfloor \frac{n}{2} \rfloor} [W_n]^k, \\ W_{n,m} &= \bigoplus_{k=0}^{\min\{m,n-m\}} [W_{n,m}]^k, \end{aligned} \tag{3.32}$$

and importantly, each $[W_n]^k$ (resp. $[W_{n,m}]^k$) is invariant under the action of both $\mathbb{B}_{n,\theta}$ and $O(\theta)$ (resp. $\mathbb{B}_{n,m,\theta}$ and $GL(\theta)$), so this is a decomposition of W_n (resp. $W_{n,m}$) into submodules of $\mathbb{C}O(\theta) \otimes \mathbb{B}_{n,\theta}$ (resp. $\mathbb{C}GL(\theta) \otimes \mathbb{B}_{n,m,\theta}$). Combining this with (3.23), it suffices to identify the irreducibles $\psi_i^{O(\theta)} \boxtimes \psi_i^{\mathbb{B}_{n,\theta}}$ contained in each $[W_n]^k$ (resp. $\psi_i^{GL(\theta)} \boxtimes \psi_i^{\mathbb{B}_{n,m,\theta}}$ contained in each $[W_{n,m}]^k$).

Proofs of Propositions 3.0.11 and 3.0.12. The key idea in the proof is that all the information we need is in the decomposition of $[W_n]^0$ - this space is known as the traceless, or harmonic, tensors. The proof therefore comes in two parts. The first part (Lemma 3.0.14) will decompose the traceless tensors $[W_n]^0$ (resp. $[W_{n,m}]^0$). Then the second part will show that we can decompose the rest of the summands $[W_n]^k$ (resp. $[W_{n,m}]^k$) using the decomposition of the traceless tensors.

Let us begin the first part, decomposing $[W_n]^0 \subset W_n$ (resp. $[W_{n,m}]^0 \subset W_{n,m}$), which, recall, is the space killed by all $Q_{i,j}$ (resp. $Q_{i,j}$ with $1 \leq i \leq m < j \leq n$).

Proposition 3.0.14. *As a representation of $\mathbb{C}O(\theta) \otimes \mathbb{B}_{n,\theta}$ (resp. $\mathbb{C}GL(\theta) \otimes \mathbb{B}_{n,m,\theta}$), we have*

$$\begin{aligned}
[W_n]^0 &\cong \bigoplus_{\substack{\lambda \vdash n \\ \lambda_1^\top + \lambda_2^\top \leq \theta}} \psi_\lambda^{O(\theta)} \boxtimes \psi_\lambda^{\mathbb{B}_{n,\theta}}, \\
[W_{n,m}]^0 &\cong \bigoplus_{\substack{(\lambda,\mu) \vdash (m,n-m) \\ \lambda_1^\top + \mu_1^\top \leq \theta}} \psi_{(\lambda,\mu)}^{GL(\theta)} \boxtimes \psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}.
\end{aligned} \tag{3.33}$$

Proof. As noted above, $[W_n]^0 \subset W_n$ (resp. $[W_{n,m}]^0 \subset W_{n,m}$) is invariant under the action of $\mathbb{C}O(\theta) \otimes \mathbb{B}_{n,\theta}$ (resp. $\mathbb{C}GL(\theta) \otimes \mathbb{B}_{n,m,\theta}$). Moreover, since $[W_n]^0$ (resp. $[W_{n,m}]^0$) is killed by any $b \in B_n \smallsetminus S_n$ (resp. $b \in B_{n,m} \smallsetminus (S_m \times S_{n-m})$), we must have that $\mathbb{B}_{n,\theta}$ acts as $\mathbb{C}S_n$ on $[W_n]^0$ (resp. $\mathbb{B}_{n,m,\theta}$ acts as $\mathbb{C}(S_m \times S_{n-m})$ on $[W_{n,m}]^0$). Moreover from the definition of the cell and irreducible modules (Lemma 2.1.11) any irreducible of $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ of $\mathbb{B}_{n,\theta}$ appearing in $[W_n]^0$ must have $\lambda \vdash n$ (Similarly from Lemma 2.1.13, if $\psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}$ appears in $[W_{n,m}]^0$ we must have $(\lambda,\mu) \vdash (m, n-m)$).

Recall we already know from the invariant theory parts of the proofs (3.23) that W_n (resp. $W_{n,m}$) decomposes intro irreducibles $U_i \boxtimes V_i$. So, as a representation of $\mathbb{C}O(\theta) \otimes \mathbb{B}_{n,\theta}$ (resp. $\mathbb{C}GL(\theta) \otimes \mathbb{B}_{n,m,\theta}$),

$$\begin{aligned}
[W_n]^0 &\cong \bigoplus_{\lambda \in \Lambda} \psi_{\underline{\lambda}}^{O(\theta)} \boxtimes \psi_\lambda^{\mathbb{B}_{n,\theta}}, \\
[W_{n,m}]^0 &\cong \bigoplus_{(\lambda,\mu) \in \Lambda'} \psi_{\underline{(\lambda,\mu)}}^{GL(\theta)} \boxtimes \psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}},
\end{aligned} \tag{3.34}$$

where Λ is some set of partitions $\lambda \vdash n$, with $\underline{\lambda}$ some partition of some size satisfying $\lambda_1^\top + \lambda_2^\top \leq \theta$ (and respectively, Λ' is a set of pairs of partitions $(\lambda,\mu) \vdash (m, n-m)$, and $\underline{(\lambda,\mu)}$ is some θ-tuple of non-increasing integers). Our first job is to identify the sets Λ and Λ', that is, determine which irreducibles of $\mathbb{B}_{n,\theta}$ (resp. $\mathbb{B}_{n,m,\theta}$) lie in the traceless tensors. Let us recall from the $GL(\theta) - S_n$ duality:

$$\begin{aligned}
W_n &\cong \bigoplus_{\substack{\rho \vdash n \\ \rho_1^\top \leq \theta}} \psi_\rho^{GL(\theta)} \boxtimes \psi_\rho^{S_n}; \\
W_{n,m} &\cong \bigoplus_{\substack{(\lambda,\mu) \vdash (m,n-m) \\ \lambda_1^\top, \mu_1^\top \leq \theta}} \left(\psi_\lambda^{GL(\theta)} \otimes \psi_{[\varnothing,\mu]}^{GL(\theta)} \right) \boxtimes \left(\psi_{(\lambda,\mu)}^{S_m \times S_{n-m}} \right),
\end{aligned} \tag{3.35}$$

where the second identity has used the dual of the first as a representation of $GL(\theta)$ (recall from Lemma 2.1.19 that the dual of $\psi_\mu^{GL(\theta)}$ is $\psi_{[\varnothing,\mu]}^{GL(\theta)}$).

Notice that $\psi_\rho^{S_n}$ must either be a subspace of $[W_n]^0$, or satisfy $\psi_\rho^{S_n} \cap [W_n]^0 = 0$; indeed, since $[W_n]^0$ is preserved by S_n, we see that $\psi_\rho^{S_n} \cap [W_n]^0$ is a subrepresentation of $\psi_\rho^{S_n}$, so it must be all of $\psi_\rho^{S_n}$, or zero. Similarly, $\psi_{(\lambda,\mu)}^{S_m\times S_{n-m}} \subset [W_{n,m}]^0$ or their intersection is zero. The following lemma determines which $\psi_\rho^{S_n}$ lie in $[W_n]^0$ (resp. $\psi_{(\lambda,\mu)}^{S_m\times S_{n-m}}$ lie in $[W_{n,m}]^0$).

Lemma 3.0.15. *1. The space $\psi_\rho^{S_n}$ is a subset of $[W_n]^0$ if $\rho_1^\top + \rho_2^\top \leq \theta$, and otherwise $\psi_\rho^{S_n} \cap [W_n]^0 = 0$;*

2. The space $\psi_{(\lambda,\mu)}^{S_m\times S_{n-m}}$ is a subset of $[W_{n,m}]^0$ if $\lambda_1^\top + \mu_1^\top \leq \theta$, and otherwise $\psi_\rho^{S_n} \cap [W_n]^0 = 0$.

Proof. We follow the proof of Theorem 10.2.5 of [44]. For part *1*, for each $\rho \vdash n$, $\rho_1^\top \leq \theta$, we take the (highest weight) vector $z_\tau\beta_\tau$ (3.26) in $\psi_\rho^{S_n}$ (for some τ shape ρ), and show that it lies in $[W_n]^0$ if and only if $\rho_1^\top + \rho_2^\top \leq \theta$. Then by our remark above, the lemma follows. Similarly for part *2*, we show the (highest weight) vector $z_\tau\beta_\tau \otimes z_\pi\beta_\pi \in \psi_{(\lambda,\mu)}^{S_m\times S_{n-m}}$ lies in $[W_n]^0$ if and only if $\lambda_1^\top + \mu_1^\top \leq \theta$; this suffices.

Notice that if we pick τ, shape $\lambda \vdash n$, with first column numbered $1, 2, \ldots, \lambda_1^\top$, second column numbered $\lambda_1^\top + 1, \ldots$, etc, then we can write the highest weight vector $z_\tau\beta_\tau$ (3.26) as

$$z_\tau\beta_\tau = (f_1 \wedge \cdots \wedge f_{\lambda_1^\top}) \otimes \cdots \otimes (f_1 \wedge \cdots \wedge f_{\lambda_s^\top}), \tag{3.36}$$

where λ has s columns. Similarly if τ shape $\lambda \vdash m$, π shape $\mu \vdash n-m$ are defined analogously to the τ above, then

$$z_\tau\beta_\tau \otimes z_\pi\beta_\pi = (f_1 \wedge \cdots \wedge f_{\lambda_1^\top}) \otimes \cdots \otimes (f_1 \wedge \cdots \wedge f_{\lambda_s^\top}) \otimes (f_1 \wedge \cdots \wedge f_{\lambda_1^\top}) \otimes \cdots \otimes (f_1 \wedge \cdots \wedge f_{\lambda_s^\top}). \tag{3.37}$$

Notice that for $1 \leq i < j \leq p$, using $Q_{i,j} = Q_{i,j}T_{i,j}$ and the antisymmetry of the wedge product,

$$Q_{i,j}(f_1 \wedge \cdots \wedge f_p) = Q_{i,j}T_{i,j}(f_1 \wedge \cdots \wedge f_p) = -Q_{i,j}(f_1 \wedge \cdots \wedge f_p), \tag{3.38}$$

so both sides are zero. Then consider $Q_{i,j}(f_1 \wedge \cdots \wedge f_p) \otimes (f_1 \wedge \cdots \wedge f_q)$, with $1 \leq i \leq p < j \leq p+q$. If $p+q \leq \theta$ then no pair of indices appears in the wedge product which sum to $\theta+1$, so the result is zero by the definition of $Q_{i,j}$ and the basis $\{f_i\}_{i=1}^\theta$. Now say $p+q \geq \theta+1$. Then there is a pair i, j such that the indices at positions i and j sum to $\theta+1$. Calculations then yield that

$$Q_{i,j}(f_1 \wedge \cdots \wedge f_p) \otimes (f_1 \wedge \cdots \wedge f_q) = \sum_{l=\theta+1-q}^{p} (f_1 \wedge \cdots \wedge \hat{f}_l \wedge \cdots \wedge f_p) \otimes (f_1 \wedge \cdots \wedge \hat{f}_{\theta+1-l} \wedge \cdots \wedge f_q), \tag{3.39}$$

where the factors $\hat{f}_l$ and $\hat{f}_{\theta+1-l}$ mean that the factors f_l and $f_{\theta+1-l}$ are omitted from the wedge product, and instead just tensor multiplied by the remaining wedge product. This sum is non-zero, since the summands are linearly independent, and at least one summand is non-zero. Now applying this to (3.36) and (3.37), we see that $Q_{i,j}z_\tau\beta_\tau = 0$ for all i, j if and only if no two of the columns of λ sum to more than θ, and $Q_{i,j}z_\tau\beta_\tau \otimes z_\pi\beta_\pi^* = 0$ for all $1 \leq i \leq m < j \leq n$ if and only if no two columns, one from λ and one from μ, sum to more than θ. ∎

By Lemma 3.0.15, the irreducibles of $\mathbb{B}_{n,\theta}$ (resp. $\mathbb{B}_{n,m,\theta}$) that appear in $[W_n]^0$ (resp. $[W_{n,m}]^0$) are exactly those with partition $\lambda \vdash n$ with $\lambda_1^\top + \lambda_2^\top \leq \theta$ (resp. $(\lambda, \mu) \vdash (m, n-m)$

with $\lambda_1^\top + \mu_1^\top \leq \theta$). That is,

$$\begin{aligned}
[W_n]^0 &\cong \bigoplus_{\substack{\lambda \vdash n \\ \lambda_1^\top + \lambda_2^\top \leq \theta}} \psi_{\underline{\lambda}}^{O(\theta)} \boxtimes \psi_{\lambda}^{\mathbb{B}_{n,\theta}}, \\
[W_{n,m}]^0 &\cong \bigoplus_{\substack{(\lambda,\mu) \vdash (m,n-m) \\ \lambda_1^\top + \mu_1^\top \leq \theta}} \psi_{\underline{(\lambda,\mu)}}^{GL(\theta)} \boxtimes \psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}},
\end{aligned} \tag{3.40}$$

where, recall, we have not yet identified the orthogonal group or general linear group irreducibles. To that end, analogously to part *1* of Lemma 3.0.13, we can show straightforwardly that $z_\tau \beta_\tau \otimes z_\pi \beta_\pi \in \psi_{(\lambda,\mu)}^{S_m \times S_{n-m}}$ is a highest weight vector for $GL(\theta)$ of weight $[\lambda, \mu]$, so $\underline{(\lambda,\mu)} = (\lambda,\mu)$, that is, the component of $[W_{n,m}]^0$ in (3.40) is exactly $\psi_{[\lambda,\mu]}^{GL(\theta)} \boxtimes \psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}$. This completes our decomposition of the traceless tensors, and the proof of Proposition 3.0.14, in the $GL(\theta) - \mathbb{B}_{n,m,\theta}$ case:

$$[W_{n,m}]^0 \cong \bigoplus_{\substack{(\lambda,\mu) \vdash (m,n-m) \\ \lambda_1^\top + \mu_1^\top \leq \theta}} \psi_{(\lambda,\mu)}^{GL(\theta)} \boxtimes \psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}. \tag{3.41}$$

The orthogonal group case takes a little more work. Recall it remains to prove that the $\underline{\lambda}$ appearing in (3.40) is actually λ, that is, the irreducible representation of $O(\theta)$ paired with $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ in (3.40) is $\psi_\lambda^{O(\theta)}$.

Let us recall some notation. Let $r = \lfloor \frac{\theta}{2} \rfloor$. For a partition λ with $\lambda_1^\top + \lambda_2^\top \leq \theta$, then λ' is λ with its first column $\lambda_1^\top$ replaced with $\theta - \lambda_1^\top$. Recall $\lambda'' = \lambda$. Note that $\lambda \neq \lambda'$ if and only if θ is odd, or θ is even and $\lambda_1^\top \neq r$ (λ does not have r parts). Pair up the partitions λ with $\lambda_1^\top + \lambda_2^\top \leq \theta$ into the pairs λ and λ'. If $\lambda \neq \lambda'$, call λ^+ the one of the pair λ, λ' with at most r parts, and λ^- the one with more than r parts. Recall from Remark 2.1.17 that the irreducibles of $SO(\theta)$ are indexed by partitions (of any size) with at most r parts, except in the case θ even, where we can allow the r^{th} part of the partition to be negative. Recall from (2.23) that in the case $\lambda \neq \lambda'$,

$$\operatorname{res}_{SO(\theta)}^{O(\theta)} \psi_{\lambda^+}^{O(\theta)} = \operatorname{res}_{SO(\theta)}^{O(\theta)} \psi_{\lambda^-}^{O(\theta)} = \psi_{\lambda^+}^{SO(\theta)}. \tag{3.42}$$

In the case $\lambda = \lambda'$ (θ even and λ with r parts), we have from (2.24) that

$$\operatorname{res}_{SO(\theta)}^{O(\theta)} \psi_{\lambda}^{O(\theta)} = \psi_{\lambda}^{SO(\theta)} + \psi_{\lambda^\circ}^{SO(\theta)}, \tag{3.43}$$

where λ° is the r-tuple λ with λ_r replaced with $-\lambda_r$. It is straightforward to show that the vector $z_\tau \beta_\tau \in \psi_\lambda^{\mathbb{B}_{n,\theta}}$ is a highest weight vector under the action of $SO(\theta)$, with weight λ^+. In the case $\lambda = \lambda'$ (θ even and λ with r parts), this is enough to show that $\mathbb{C}O(\theta) z_\tau \beta_\tau$ is a copy of the irreducible $\psi_\lambda^{O(\theta)}$, which is what we wanted.

In the other case, $\lambda \neq \lambda'$, equation (3.42) tells us that $\mathbb{C}O(\theta) z_\tau \beta_\tau$ is either $\psi_\lambda^{O(\theta)}$ or $\psi_{\lambda'}^{O(\theta)}$ (i.e. $\psi_{\lambda^+}^{O(\theta)}$ or $\psi_{\lambda^-}^{O(\theta)}$). In order to tell which it is, Theorem 2.1.20 tells us we need to ascertain how $O(\theta) \smallsetminus SO(\theta)$ acts on $\mathbb{C}O(\theta) z_\tau \beta_\tau$. For θ odd, it suffices to show how $-\mathbf{id} \in O(\theta)$ acts. Recall $-\mathbf{id}$ acts on $\psi_\lambda^{O(\theta)}$ as $(-1)^{|\lambda|}\mathbf{id}$ (since θ is odd, $|\lambda|$ and $|\lambda'|$ have different parity). Now indeed $-\mathbf{id}$ does act on $z_\tau \beta_\tau$ as $(-1)^{|\lambda|}\mathbf{id}$, since $-\mathbf{id}$ acts on V as itself, so it acts on $W_n = V^{\otimes n}$ as $(-1)^n \mathbf{id}$, and $\lambda \vdash n$.

In the remaining case, θ even and $\lambda_1^\intercal \neq r$, it suffices to find how g_0 acts on the highest weight vector $z_\tau\beta_\tau$, where $g_0 \in O(\theta) \setminus SO(\theta)$ acts on V by fixing each basis vector f_i, $i \neq r, r+1$, and exchanging f_r and f_{r+1}. Recall from Theorem 2.1.20 that g_0 multiplies the highest weight vector in $\psi_{\lambda^\pm}^{O(\theta)}$ by ± 1. In the case $\lambda = \lambda^+$ (that is, λ has less than r parts), notice that from the definition of β_τ, f_r and f_{r+1} do not appear as tensor factors in β_τ, so g_0 fixes β_τ. As g_0 commutes with the action of B_n, g_0 also fixes $z_\tau\beta_\tau$. Now let $\lambda = \lambda^-$ (that is, λ has more than r parts). If λ has s columns, then up to rearrangement of tensor factors,

$$z_\tau\beta_\tau = (f_1 \wedge \cdots \wedge f_{\lambda_1^\intercal}) \otimes \cdots \otimes (f_1 \wedge \cdots \wedge f_{\lambda_s^\intercal}). \tag{3.44}$$

Note f_r and f_{r+1} appear here only in the leftmost wedge product, since all columns of λ except for the first have length strictly less than r. Now g_0 acts on this vector the same as the transposition $(r, r+1)$, which, by the antisymmetry of the wedge product, multiplies the vector by (-1). This shows that when τ has shape $\lambda^\pm$, indeed g_0 multiplies $z_\tau\beta_\tau$ by ± 1, so we can conclude in all cases that $\mathbb{C}O(\theta)z_\tau\beta_\tau \cong \psi_\lambda^{O(\theta)}$. This completes the proof of Proposition 3.0.14, our decomposition of the traceless tensors in the $O(\theta) - \mathbb{B}_{n,\theta}$ case. ∎

This, in turn, completes the first part of the proof of Propositions 3.0.11 and 3.0.12.

Now for part two of the proofs of Propositions 3.0.11 and 3.0.12, we move on to the summand $[W_n]^k$ (resp. $[W_{n,m}]^k$) in (3.32), $k > 0$.

Proposition 3.0.16. *As a representation of $\mathbb{C}O(\theta) \otimes \mathbb{B}_{n,\theta}$ (resp. $\mathbb{C}GL(\theta) \otimes \mathbb{B}_{n,m,\theta}$), we have*

$$\begin{aligned} [W_n]^k &\cong \bigoplus_{\substack{\lambda \vdash n-2k \\ \lambda_1^\intercal + \lambda_2^\intercal \leq \theta}} \psi_\lambda^{O(\theta)} \boxtimes \psi_\lambda^{\mathbb{B}_{n,\theta}}, \\ [W_{n,m}]^k &\cong \bigoplus_{\substack{(\lambda,\mu) \vdash (m-k, n-m-k) \\ \lambda_1^\intercal + \mu_1^\intercal \leq \theta}} \psi_{(\lambda,\mu)}^{GL(\theta)} \boxtimes \psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}. \end{aligned} \tag{3.45}$$

Notice the dependence on k on both sides of the equations (3.45).

Proof. As noted earlier, we aim to show that all the information we need is contained in our decomposition of $[W_n]^0$ (resp. $[W_{n,m}]^0$) from (3.34). First, the image of a contraction $Q_{1,2}$ on the tensor space W_2 (resp. $W_{2,1}$), (which is spanned by the vector $\sum_{i=1}^\theta f_i \otimes f_i^* = \sum_{i=1}^\theta f_i \otimes f_{\theta+1-i}$), is a copy of the trivial representation of $O(\theta)$ (resp. $GL(\theta)$) under the action $g \mapsto g \otimes g$ (resp. $g \otimes g^{-*}$). So, $Q_{i,j}W_n \cong W_{n-2}$ as a representation of $O(\theta)$, and $Q_{i,j}W_{n,m} \cong W_{n-2,m-1}$ as a representation of $GL(\theta)$, for all $1 \leq i \leq m < j \leq n$. Repeating this argument we see that for any $(\underline{t}, \underline{t}') \in \mathcal{Q}(k)$ (resp. $\mathcal{Q}'(k)$), as representations of $O(\theta)$ (resp. $GL(\theta)$),

$$\begin{aligned} Q_{\underline{t},\underline{t}'}W_n &\cong W_{n-2k}, \\ Q_{\underline{t},\underline{t}'}W_{n,m} &\cong W_{n-2k,m-k}, \end{aligned} \tag{3.46}$$

and moreover

$$\begin{aligned}[][Q_{\underline{t},\underline{t}'}W_n]^0 &\cong [W_{n-2k}]^0,\\ [Q_{\underline{t},\underline{t}'}W_{n,m}]^0 &\cong [W_{n-2k,m-k}]^0,\end{aligned} \tag{3.47}$$

where $[Q_{\underline{t},\underline{t}'}W_n]^0$ is the set of vectors in $Q_{\underline{t},\underline{t}'}W_n$ which are killed by any $Q_{i,j}$ with $i,j \in (\underline{t} \cup \underline{t}')^c$, $[Q_{\underline{t},\underline{t}'}W_{n,m}]^0$ similar.

Second, we have a (perhaps not direct sum) decomposition of the space

$$[W_n]^k = \sum_{(\underline{t},\underline{t}')\in\mathcal{Q}(k)} [Q_{\underline{t},\underline{t}'}W_n]^0. \tag{3.48}$$

Indeed, if $[Q_{\underline{t},\underline{t}'}W_n]^1$ is those vectors in $Q_{\underline{t},\underline{t}'}W_n$ which are in the image of some $Q_{i,j}$ with $i,j \in (\underline{t}\cup\underline{t}')^c$, then

$$\begin{aligned} W_n^k &= \sum_{(\underline{t},\underline{t}')\in\mathcal{Q}(k)} Q_{\underline{t},\underline{t}'}W_n = \left(\sum_{(\underline{t},\underline{t}')\in\mathcal{Q}(k)} [Q_{\underline{t},\underline{t}'}W_n]^0\right) \oplus \left(\sum_{(\underline{t},\underline{t}')\in\mathcal{Q}(k)} [Q_{\underline{t},\underline{t}'}W_n]^1\right)\\ &= \left(\sum_{(\underline{t},\underline{t}')\in\mathcal{Q}(k)} [Q_{\underline{t},\underline{t}'}W_n]^0\right) \oplus W_n^{k+1}, \end{aligned} \tag{3.49}$$

where we have used that $W_n^k = [W_n]^k \oplus W_n^{k+1}$. It then suffices by dimension count to show that $\sum_{(\underline{t},\underline{t}')\in\mathcal{Q}(k)}[Q_{\underline{t},\underline{t}'}W_n]^0 \subset [W_n]^k$. Any vector in $[Q_{\underline{t},\underline{t}'}W_n]^0$ is of the form $Q_{\underline{t},\underline{t}'}v$, and is killed by any $Q_{i,j}$ with $i,j \in (\underline{t}\cup\underline{t}')^c$; we need to prove that this vector is killed by any $Q_{\underline{s},\underline{s}'}$, $(\underline{s},\underline{s}') \in \mathcal{Q}(k+1)$. This follows from the fact that in $\mathbb{B}_{n,\theta}$, the product of diagrams $(\overline{\underline{s},\underline{s}'})(\overline{\underline{t},\underline{t}'})$ is (some scalar multiple of) a diagram with a southern bar connecting i^{S} and j^{S}, with $i,j \in (\underline{t}\cup\underline{t}')^c$. In a very similar manner, we can prove a (not necessarily direct) decomposition $[W_{n,m}]^k = \sum_{(\underline{t},\underline{t}')\in\mathcal{Q}'(k)}[Q_{\underline{t},\underline{t}'}W_{n,m}]^0$.

Now combining equation (3.48) with equation (3.47), we have, as a representation of $O(\theta)$ (resp. $GL(\theta)$):

$$\begin{aligned}[][W_n]^k &= \sum_{(\underline{t},\underline{t}')\in\mathcal{P}(k)} [Q_{\underline{t},\underline{t}'}W_n]^0 \cong \sum_{(\underline{t},\underline{t}')\in\mathcal{P}(k)} [W_{n-2k}]^0,\\ [W_{n,m}]^k &= \sum_{(\underline{t},\underline{t}')\in\mathcal{P}'(k)} [Q_{\underline{t},\underline{t}'}W_{n,m}]^0 \cong \sum_{(\underline{t},\underline{t}')\in\mathcal{P}'(k)} [W_{n-2k,m-k}]^0.\end{aligned} \tag{3.50}$$

But now, we know from the first part of our proof, Proposition 3.0.14, how the spaces $[Q_{\underline{t},\underline{t}'}W_n]^0$ (resp. $[Q_{\underline{t},\underline{t}'}W_{n,m}]^0$) decompose as representations of $O(\theta)$ (resp. $GL(\theta)$). The irreducibles of $O(\theta)$ (resp. $GL(\theta)$) appearing are all of the $\psi_\lambda^{O(\theta)}$ with $\lambda \vdash n-2k$, $\lambda_1^{\mathsf{T}}+\lambda_2^{\mathsf{T}} \le \theta$ (resp. all of the $\psi_{[\lambda,\mu]}^{GL(\theta)}$ with $(\lambda,\mu) \vdash (m-k, n-m-k)$ with $\lambda_1^{\mathsf{T}}+\mu_1^{\mathsf{T}} \le \theta$). Formally, we have

$$\begin{aligned}[][W_n]^k &\cong \bigoplus_{\substack{\lambda\vdash n-2k\\ \lambda_1^{\mathsf{T}}+\lambda_2^{\mathsf{T}}\le\theta}} \psi_\lambda^{O(\theta)} \boxtimes \psi_{\hat{\lambda}}^{\mathbb{B}_{n,\theta}},\\ [W_{n,m}]^k &\cong \bigoplus_{\substack{(\lambda,\mu)\vdash(m-k,n-m-k)\\ \lambda_1^{\mathsf{T}}+\mu_1^{\mathsf{T}}\le\theta}} \psi_{(\lambda,\mu)}^{GL(\theta)} \boxtimes \psi_{\widehat{(\lambda,\mu)}}^{\mathbb{B}_{n,m,\theta}},\end{aligned} \tag{3.51}$$

where it now only remains to check that $\hat{\lambda} = \lambda$ and $\widehat{(\lambda,\mu)} = (\lambda,\mu)$.

Recall $\mathcal{ST}((\underline{t}\cup\underline{t}')^c)$ is the set of standard Young tableaux with entries in $\{1,\dots,n\}\setminus(\underline{t}\cup\underline{t}')$. For fixed $0\le m\le n$ and $(\underline{t},\underline{t}')\in\mathcal{Q}'(k)$, define similarly the set $\mathcal{ST}'((\underline{t}\cup\underline{t}')^c)$ as the set of pairs of standard Young tableaux (τ,π) with τ having entries in $\{1,\dots,m\}\setminus(\underline{t}\cup\underline{t}')$, and π having entries in $\{m+1,\dots,n\}\setminus(\underline{t}\cup\underline{t}')$.

Using (3.34) and (3.50), the highest weight vector for each instance of $\psi_\lambda^{O(\theta)}$ (resp. $\psi_{[\lambda,\mu]}^{GL(\theta)}$) is the vector

$$\begin{aligned} y_{\tau,\underline{t},\underline{t}'} &:= z_\tau Q_{\underline{t},\underline{t}'}\beta_{\tau,\underline{t},\underline{t}'};\\ y_{(\tau,\pi),\underline{t},\underline{t}'} &:= z_{(\tau,\pi)}Q_{\underline{t},\underline{t}'}\beta_{(\tau,\pi),\underline{t},\underline{t}'}, \end{aligned} \tag{3.52}$$

for $(\underline{t},\underline{t}')\in\mathcal{Q}(k)$ (resp. $\mathcal{Q}'(k)$), and $\tau\in\mathcal{ST}((\underline{t}\cup\underline{t}')^c)$ (resp. $(\tau,\pi)\in\mathcal{ST}'((\underline{t}\cup\underline{t}')^c)$). Here $z_{(\tau,\pi)}=z_\tau z_\pi$, and $\beta_{\tau,\underline{t},\underline{t}'}$ is the basis vector $f_{i_1}\otimes\cdots\otimes f_{i_n}$ such that $f_{t_i}=f^*_{t'_i}=f_1$ for all $i=1,\dots,k$, and if $j\in(\underline{t},\underline{t}')^c$, then i_j is the index of the row that j lies in in τ. The vector $\beta_{(\tau,\pi),\underline{t},\underline{t}'}$ is defined similarly: it is the basis vector $f_{i_1}\otimes\cdots\otimes f_{i_n}$ such that $f_{t_i}=f^*_{t'_i}=f_1$ for all $i=1,\dots,k$, and otherwise $f_{i_j}=f_p$ if j lies in the p^{th} row of τ, or $f_{i_j}=f_{\theta+1-p}$ if j lies in the p^{th} row of π.

Let M_λ be the space spanned by the vectors $y_{\tau,\underline{t},\underline{t}'}$, $(\underline{t},\underline{t}')\in\mathcal{Q}(k)$ and $\tau\in\mathcal{ST}_\lambda((\underline{t}\cup\underline{t}')^c)$, and $M_{\lambda,\mu}$ similar. Then we have

$$\begin{aligned} [W_n]^k &= \sum_{\substack{\lambda\vdash n-2k\\ \lambda_1^\intercal+\lambda_2^\intercal\le\theta}} \psi_\lambda^{O(\theta)}\otimes M_\lambda\\ [W_{n,m}]^k &= \sum_{\substack{(\lambda,\mu)\vdash(m-k,n-m-k)\\ \lambda_1^\intercal+\mu_1^\intercal\le\theta}} \psi_{[\lambda,\mu]}^{GL(\theta)}\otimes M_{\lambda,\mu}. \end{aligned} \tag{3.53}$$

It now remains to prove that M_λ is the irreducible representation $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ of $\mathbb{B}_{n,\theta}$, and $M_{\lambda,\mu}$ is the irreducible representation $\psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}$ of $\mathbb{B}_{n,m,\theta}$. By the definitions of the cell modules Δ_λ of $\mathbb{B}_{n,\theta}$ (resp. $\Delta_{\lambda,\mu}$ of $\mathbb{B}_{n,m,\theta}$), it is clear that M_λ is a quotient of Δ_λ (and $M_{\lambda,\mu}$ is a quotient of $\Delta_{\lambda,\mu}$). Let us show M_λ is irreducible; the proof for $M_{\lambda,\mu}$ is almost identical. We follow Theorem 4.5 of Benkart et al. [8]. Take $0\ne v\in M_\lambda$. We want to show $\mathbb{B}_{n,\theta}v=M_\lambda$. Since $\mathbb{B}_{n,\theta}$ acts transitively on the vectors $y_{\tau,\underline{t},\underline{t}'}$, it suffices to show that one $y_{\tau,\underline{t},\underline{t}'}$ lies in $\mathbb{B}_{n,\theta}v$.

There exists a $Q_{\underline{t},\underline{t}'}$ with $(\underline{t},\underline{t}')\in\mathcal{Q}(k)$ such that $Q_{\underline{t},\underline{t}'}v\ne0$. Indeed, if there were not, then v would lie in $[W_n]^{k-1}$ by definition. But by definition, $y_{\tau,\underline{t},\underline{t}'}\in W_n^k$, and $[W_n]^{k-1}\cap W_n^k=0$. So, say

$$0\ne Q_{\underline{t},\underline{t}'}v=\sum_{\pi\in\mathcal{ST}_\lambda((\underline{t}\cup\underline{t}')^c)} a_\pi y_{\pi,\underline{t},\underline{t}'}. \tag{3.54}$$

The right hand side follows from calculations using the relations in $\mathbb{B}_{n,\theta}$. Now we use the same trick we used in the $S_n-GL(\theta)$ case: let τ be minimal with respect to the ordering $<$ such that $a_\tau\ne0$. Multiplying (3.54) by z_τ kills all terms apart from the τ one by our working with Young symmetrisers; hence $y_{\tau,\underline{t},\underline{t}'}\in\mathbb{B}_{n,\theta}v$, which completes the proof. This completes the proof of Proposition 3.0.16.

■

Now putting together this result, along with equations (3.50), and (3.32), we have that

$$\begin{aligned} W_n &\cong \bigoplus_{k=0}^{\lfloor \frac{n}{2} \rfloor} \bigoplus_{\substack{\lambda \vdash n-2k \\ \lambda_1^\top + \lambda_2^\top \leq \theta}} \psi_\lambda^{O(\theta)} \boxtimes \psi_\lambda^{\mathbb{B}_{n,\theta}}, \\ W_{n,m} &\cong \bigoplus_{k=0}^{\min\{m,n-m\}} \bigoplus_{\substack{(\lambda,\mu) \vdash (m-k,n-m-k) \\ \lambda_1^\top + \mu_1^\top \leq \theta}} \psi_{[\lambda,\mu]}^{GL(\theta)} \boxtimes \psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}, \end{aligned} \tag{3.55}$$

which completes the proofs of Propositions 3.0.11 and 3.0.12, and thereby the proofs of Schur-Weyl duality in the $O(\theta)-\mathbb{B}_{n,\theta}$ and $GL(\theta)-\mathbb{B}_{n,m,\theta}$ cases, Theorems 3.0.3 and 3.0.5.

■

Chapter 4

The Manhattan and Lorentz Mirror Models

In this section, we present the results of the paper "The Manhattan and Lorentz Mirror Models - A result on the Cylinder with low density of mirrors" [90]. This paper studies two random walks on the two-dimensional lattice $\mathbb{Z}^2$, where the central question is whether the walk is bounded or not. The Brauer and walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ is made use of, by viewing the walk on the cylinder as a Markov chain on its basis $B_{n,m}$. The main result bounds the distance the walk can travel on the cylinder of width n, given that the probability of mirrors (see below) decays at least as order n^{-1}.

4.1 Introduction

The Manhattan and Lorentz mirror models [7], [62], are two very similar models, each describing a random walk on the $\mathbb{Z}^2$ lattice. Let $0 \le p \le 1$. The walker is a particle of light which bounces off mirrors placed at each vertex at 45°, independently with probability p. For the Lorentz mirror model, the orientation of the mirror (NW or NE) is chosen independently with probability $\frac{1}{2}$. For the Manhattan model, the lattice is a priori given Manhattan directions (see Figure 4.1), and the orientation of the mirror is determined by its location (i.e. a NW mirror if the sum of the coordinates of the point is odd, and NE if the sum is even), so that the walker always follows the directions of the lattice. The main questions of interest in both models are whether the paths remain bounded or not, and the nature of the motion of the walker.

We study the models on an infinite cylinder $\mathbb{Z} \times (\mathbb{Z}/n\mathbb{Z})$ of finite even width n. We are interested in how the length of the paths vary with p. Note that on the cylinder, paths are bounded with probability 1 - indeed, there is a positive probability that a horizontal row is filled with mirrors such that no path can pass the row; one has to wait an expected p^{-n} rows for this event. It is natural to hope that this bound can be improved. The result of this paper, Theorem 4.1.1, shows that for both models, when $p \le Cn^{-1}$, C a constant, the highest row reached by a path on the n-cylinder is order p^{-2}. We wonder whether this is true for all p.

We observe an underlying algebraic structure (valid for any value of p). The models on the cylinder can be thought of as Markov chains on the Brauer algebra (in the mirror case), or its subalgebra the walled Brauer algebra (in the Manhattan case). While the result of this paper can be obtained without these algebras, we suggest that the models' association with different algebraic structures reflects their different behaviours. A third model, on the L-lattice (see [6]), with different behaviour from the other two models, is solved using percolation, and can be similarly thought of as a Markov chain on the (extended) Temperley Lieb algebra.

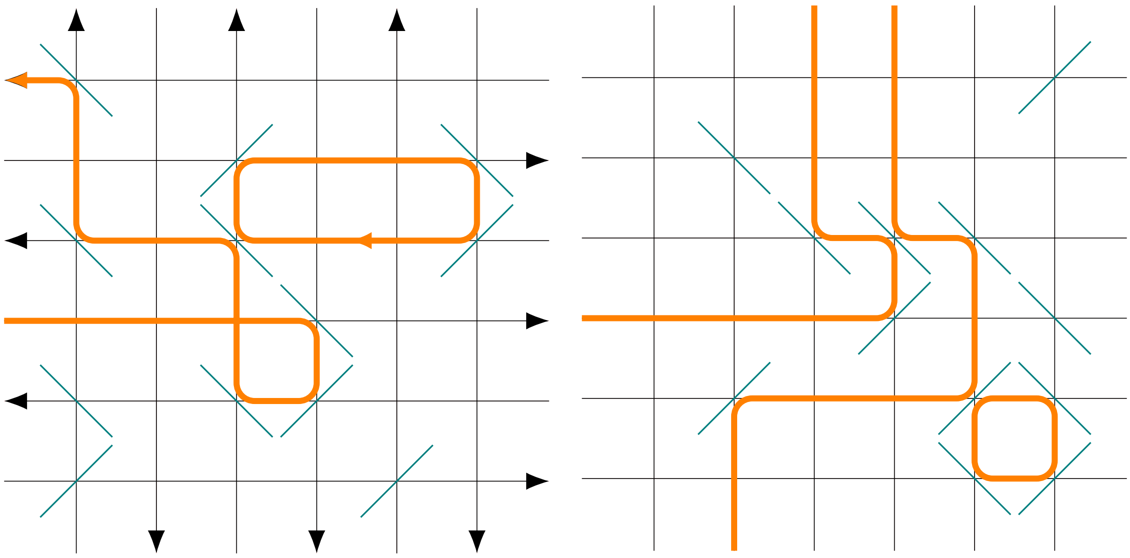


Figure 4.1: Examples of the Manhattan model (left) and Mirror model (right), with mirrors in blue, and a few paths highlighted in orange. Note that the orientation of a mirror in the Manhattan case is determined by the Manhattan directions of the lattice.

Let us recap the existing results on both models (which are on $\mathbb{Z}^2$, unless otherwise specified). The Mirror model was introduced by Ruijgrok and Cohen [88] as a lattice version of the Ehrenfest wind-tree model. Grimmett [46] proved with a straightforward argument that on $\mathbb{Z}^2$, if $p = 1$, then the path of the walker is bounded with probability 1. It is conjectured that this is also true for $0 < p \leq 1$. This is supported by numerical simulations, for example, in [106]. More recently, Kozma and Sidoravicius [62] showed that, for any $0 < p \leq 1$, the probability the walker reaches the boundary of the n-box $[-n, n]^2$ is at least $\frac{1}{2n+1}$. To obtain this result, they study the model on an infinite cylinder of finite odd width, where there is deterministically always an infinite path. The Manhattan model cannot be neatly defined on a cylinder of odd width (it cannot remain rotation-invariant), so this method cannot be applied (and indeed, the result is not true in the Manhattan case - see below).

The Mirror model on the cylinder (often under the name the O(1) loop model) has been studied using the Brauer algebra before, in several papers relating to a conjecture (and variations thereof) by Razumov and Stroganov [85], [28], [27], which gives the entries of the limiting distribution in terms of combinatorial objects such as alternating sign matrices. A generalised mirror model (the $O(\theta)$ loop model), where the distribution on configurations is weighted by $\theta^{\#loops}$, $\theta \in \mathbb{C}$, is studied in [70], [79]; this is the model on the Brauer algebra with parameter θ, $\mathbb{B}_{n,\theta}$. In these papers, the requirement of a Yang-Baxter equation restricts the permissible values of the parameters - in our specific setup, only

$p = \frac{8}{9}$ qualifies (see the end of [79]).

The Manhattan model shares features of quantum disordered systems. The model was introduced by Beamond, Cardy and Owczarek [6], in close relation to a quantum network model on the Manhattan lattice. The quantum model has random $Sp(2) = SU(2)$ matrices on each edge of the lattice, and the classical model arises on averaging over this disorder. In most classical models in two dimensions, localisation (bounded paths) is not observed, whereas in the Manhattan (and Mirror) model, it is expected (see below). It is not clear if the mirror model has a similar explicit relationship with a quantum model. For more detail on the connection to quantum models, see Spencer's review [92].

An argument from [6] for tackling the Manhattan model uses percolation. The placement of the mirrors is exactly a Bernoulli percolation on the edges of $\mathbb{Z}^2$, rotated 45° and scaled. The path of the walker stays within $\frac{1}{\sqrt{2}}$ of its closest dual cluster (see Figure 4.2). The dual clusters are finite with probability 1 for $p \geq \frac{1}{2}$, so so are the Manhattan paths.

For $p > \frac{1}{2}$, the probability that two points are in the same dual cluster decays exponentially in the distance, which gives the same for connection by a Manhattan path. This is markedly different from the Mirror model's polynomial decay. For $p < \frac{1}{2}$, this argument is wholly inconclusive, since dual clusters are almost surely infinite. Recently, Li [65] gave exponential decay in connection probabilities for $p > \frac{1}{2} - \epsilon$, for some $\epsilon > 0$. Numerical simulations in [7] indicate that paths are finite for $0 < p < \frac{1}{2}$, with exponential decay in connection probabilities. Clearly for $p = 0$, the paths escape in straight lines to infinity.

On the cylinder, for both models, there is first the crude, simple bound given above. Notably, after this paper was originally posted, Li [66] showed the following: for both models, on the cylinder of even width n, and for fixed p, the walker reaches at most $O(n^{10})$ rows from its startpoint, with probability exponentially close to 1. Let us note that the results on the (even) cylinder (including this paper) are the same for the two models, but on $\mathbb{Z}^2$ they are different. To analyse the $\mathbb{Z}^2$ case via the cylinder, one must look at a cylinder of equal height and width (as in [62]); it is here that the models differ.

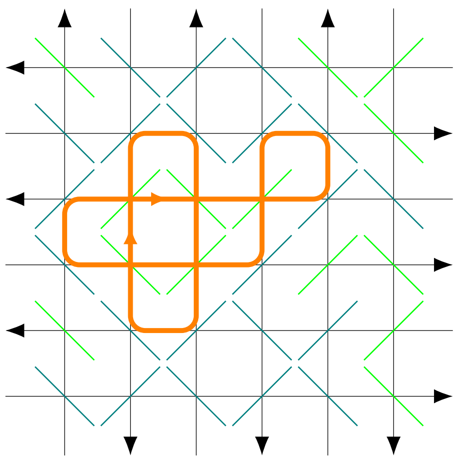


Figure 4.2: The mirrors (in blue) in the Manhattan model as edges in Bernoulli percolation. The green edges form the dual clusters. The two paths shown are restricted to stay within $\frac{1}{\sqrt{2}}$ of one dual cluster.

Let us now state our result more precisely. Consider the models on the n-cylinder $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z} = \{(i,t) \ : \ i,t \in \mathbb{Z}, \ 1 \leq i \leq n\}$, with n even. We label s_t the horizontal row

$\{(i,t) \;:\; 1 \le i \le n\}$ - the "t^{th} street". For the Mirror model, let $V^{mir}_{\frac{n}{2}}$ be the random variable given by the smallest t such that s_t has no path connecting it to the first street, s_1. In other words, the highest street a path from s_1 reaches is exactly $V^{mir}_{\frac{n}{2}} - 1$. Let $V^{mat}_{\frac{n}{2}}$ be defined identically for the Manhattan model.

Theorem 4.1.1. *Let $*$ represent mat or mir.*

a) Let $p \le Cn^{-1}$, $C > 0$ a constant. For all $\alpha > 0$,

$$\mathbb{P}[V^*_{\frac{n}{2}} \ge \alpha p^{-2}] \le 2A_* e^{-\frac{1}{8e^C}\alpha},$$

where $A_{mir} = cosh(\pi)$, and $A_{mat} = \frac{sinh(\pi)}{\pi}$.

b) For any $p \le \frac{1}{2}$ (not necessarily constrained by $p < Cn^{-1}$), and for all $\alpha > 0$,

$$\mathbb{P}\left[V^*_{\frac{n}{2}} \le \alpha p^{-2}\right] \le 2\alpha.$$

Let us give an informal overview of the proof of part a). Our argument uses the streets which have at most two mirrors. As $n \to \infty$, for all $p \le Cn^{-1}$, and C small, the probability of mirrors is small, and in particular, the probability that each street s_t has at most two mirrors is large. We show that the model is not changed too much if we actually condition on each s_t having at most two mirrors. This conditioning simplifies the model greatly, in essence removing the cylindrical geometry, making the interactions on each street mean-field (in the sense that if the particle arrives at street s_t at the point (i,t) and leaves from (j,t), $j \ne i$, then j is uniformly distributed). This allows us to do explicit computations. For C not small, the theorem still holds, but the bounds are less sharp; one needs to set α exponentially large in C to bring the bound to less than 1. Part b) is more straightforward; it is proved by coupling $V^*_{\frac{n}{2}}$ with a geometric random variable G with parameter p^2.

In section 4.2, we give key definitions, including the Brauer and walled Brauer algebras. In section 4.3 we study the model assuming at most two mirrors per street, and obtain the results needed to prove Theorem 4.1.1.

4.2 Definitions, and the Brauer algebra

Let us recall the algebraic structures and notation from Chapter 2 that we will use in this section. The Brauer algebra $\mathbb{B}_{n,1}$ (the special case of $\mathbb{B}_{n,\theta}$ from Section 2.1.3 with $\theta = 1$) (see, for example, [19], [20], [24], [103]) is the (formal) complex span of the set of pairings of $2n$ vertices. We think of pairings as graphs, which we will call diagrams, with each vertex having degree exactly 1. We arrange the vertices in two horizontal rows, labelling the upper row (the northern vertices) $1^{\mathsf{N}}, 2^{\mathsf{N}}, \ldots, n^{\mathsf{N}}$, and the lower (southern) $1^{\mathsf{S}}, \ldots, n^{\mathsf{S}}$. We call an edge connecting two northern vertices (or two southern) a bar. The number of bars in the north and south is always the same, and we refer to either simply as the number of bars in the diagram. We call an edge connecting a northern and southern vertex a NS-path.

Multiplication of two diagrams is given by concatenation. If b, c are two diagrams, we align the northern vertices of b with the southern of c, and the result is obtained by removing these middle vertices. See Figure 4.3. We let bc denote the product (occasionally $b \cdot c$ for clarity). This defines $\mathbb{B}_{n,1}$ as an algebra. Of course this is a special case of $\mathbb{B}_{n,\theta}$ from Section 2.1.3, but for our purposes in this section we only need $\theta = 1$, which gives the multiplication described above.

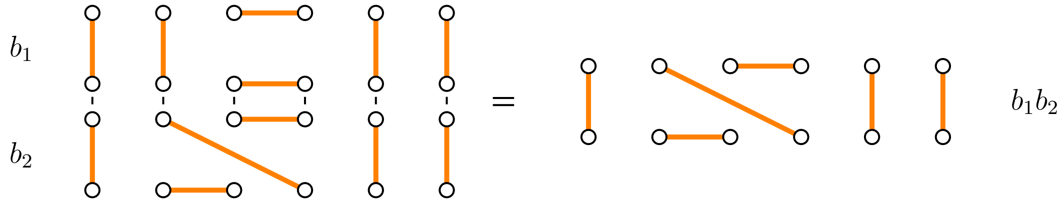


Figure 4.3: Two diagrams b_1 and b_1 (left), concatenated to produce their product (right).

We call the set of all diagrams B_n. We call the set of diagrams with exactly k bars $B_n\langle k\rangle$, and the set of diagrams with at least k bars B_n^k. Notice that $B_n\langle 0\rangle$ is exactly the symmetric group S_n, and the concatenation multiplication exactly reduces to the multiplication in S_n. So $\mathbb{C}S_n$ is a subalgebra of $\mathbb{B}_{n,1}$.

We write id for the identity in S_n - its diagram has all its edges vertical. We denote the transposition S_n swapping i and j by (i,j), and we write $(\overline{i,j})$ for the diagram with i^{N} connected to j^{N}, and i^{S} connected to j^{S}, and all other edges vertical. See Figure 4.4.

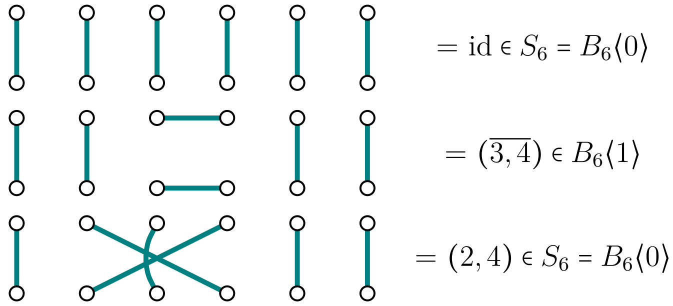


Figure 4.4: The identity element, the element $(\overline{3,4}) \in B_6\langle 1\rangle$, and the transposition $(2,4) \in S_6 = B_6\langle 0\rangle$.

Finally, we remark that if b has k bars, and c is any diagram in B_n, then bc must have at least k bars:

$$b \in B_n\langle k\rangle \;\Rightarrow\; bc \in B_n^k. \tag{4.1}$$

Let us now see how the Brauer algebra can be used to describe the models. Let n be even from hereon in. Observe that given a configuration σ_t of mirrors on a street s_t on the n-cylinder, the paths through the street form a diagram $b(\sigma_t) \in B_n$. See Figure 4.5 for an illustration. Moreover, on any section of the cylinder, say, from street s_{t_1} to s_{t_2}, given a configuration of mirrors $\sigma_{t_1\to t_2}$, the paths through those streets form a diagram $b(\sigma_{t_1\to t_2})$. Crucially, we see that $b(\sigma_{t_1\to t_2}) = b(\sigma_{t_1})\cdots b(\sigma_{t_2})$, where the multiplication on the right hand side is in the Brauer algebra. See Figure 4.6.

Let each σ_t, the configuration of mirrors on the t^{th} street, be distributed according to the Manhattan or Mirror model. Then $b(\sigma_t)$ is a random diagram in B_n. We can think of the distribution of this random diagram as a (deterministic) element $Z_{(t)}$ of the algebra:

$$Z_{(t)} = \sum_{g\in B_n} \mathbb{P}[b(\sigma_t) = g] \cdot g.$$

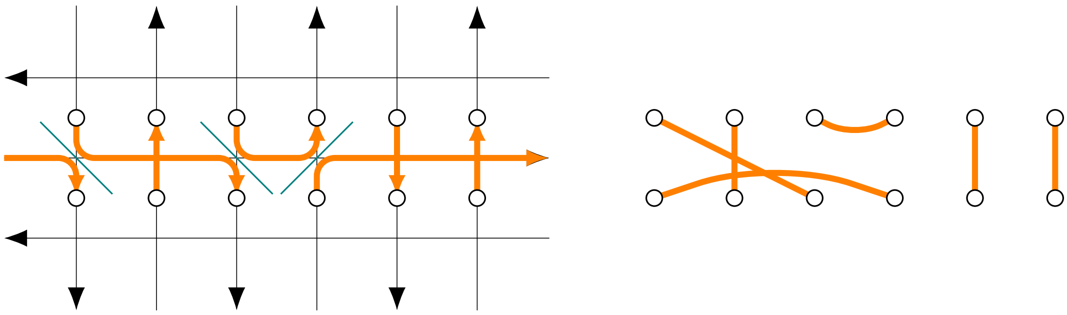


Figure 4.5: An example of a configuration of mirrors σ_t on street s_t in the Manhattan model (left), and the resulting diagram $b(\sigma_t)$ (right).

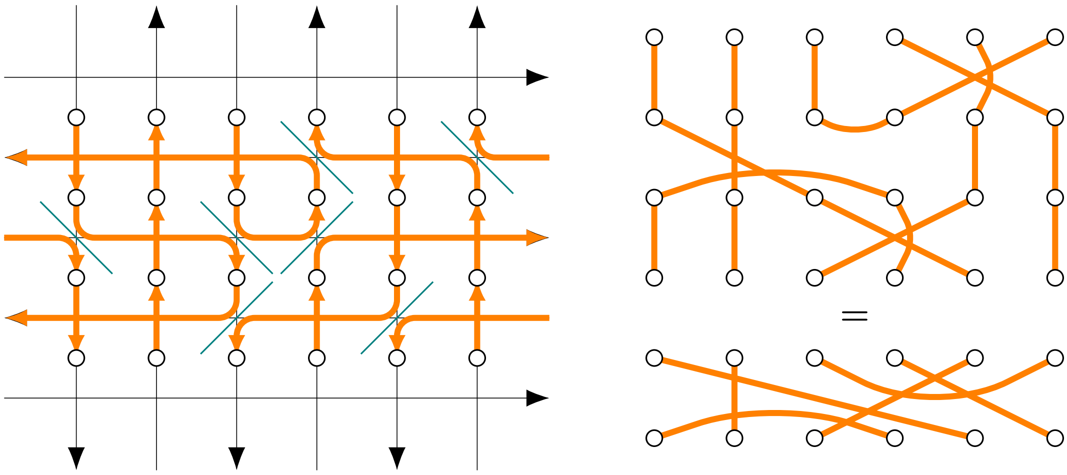


Figure 4.6: The paths through three consecutive streets in the Manhattan model (left), the three resulting diagrams (upper right), and their product (lower right), which gives the paths through the union of the three streets.

The following lemma lets us describe the paths through any number of consecutive streets. Note that it does not use the specific distributions of the random diagrams given by different streets, it only uses that they are independent.

Lemma 4.2.1. *The distribution of the random diagram $b(\sigma_{t_1 \to t_2})$ produced by the paths through streets $s_{t_1}, \dots s_{t_2}$ is given by the Brauer algebra element:*

$$Z_{(t_1)} \cdots Z_{(t_2)} = \sum_{g \in B_n} \mathbb{P}[b(\sigma_{t_1 \to t_2}) = g] \cdot g,$$

where the multiplication on the left is in the Brauer algebra.

Proof. We see that

$$\begin{aligned}
\sum_{g \in B_n} \mathbb{P}[b(\sigma_{t_1 \to t_2}) = g] \cdot g &= \sum_{g \in B_n} \mathbb{P}[b(\sigma_{t_1}) \cdots b(\sigma_{t_2}) = g] \cdot g \\
&= \sum_{g \in B_n} \sum_{g_{t_1} \cdots g_{t_2} = g} \mathbb{P}[b(\sigma_{t_1}) = g_{t_1}] \cdots \mathbb{P}[b(\sigma_{t_2}) = g_{t_2}] \cdot g_{t_1} \cdots g_{t_2} \\
&= Z_{(t_1)} \cdots Z_{(t_2)},
\end{aligned}$$

where we use that the configurations on each street are independent. ∎

We are interested in the highest (or most northerly) street reached by the paths starting at the first street s_1. One more than this is the first street which has no path connecting it to s_1. Using the notation above, this is the smallest t such that the random diagram $b(\sigma_{1\to t})$ has no NS-paths, ie:

$$b(\sigma_{1\to t}) \in B_n\langle\frac{n}{2}\rangle.$$

Let $*$ represent mat or mir. Let $\sigma_{*,t}$ (resp. $\sigma_{*,t_1\to t_2}$) denote the random configuration of mirrors on the street s_t (resp. the streets $s_{t_1},\dots,s_{t_2}$), in the corresponding model. Now, in the Mirror model, the random configuration of mirrors $\sigma_{mir,t}$ is iid for each street s_t. Let Z_{mir} be the distribution of the random diagram $b(\sigma_{mir,t})$ (as an element of the Brauer algebra) produced by the paths through this random configuration on one street. (Since the $\sigma_{mir,t}$ are iid, Z_{mir} is independent of t). We note that, from Lemma 4.2.1,

$$Z^t_{mir} = \sum_{g\in B_n} \mathbb{P}[b(\sigma_{mir,1\to t}) = g]\cdot g.$$

Let V^{mir}_k be the random variable given by the smallest t such that $b(\sigma_{mir,1\to t}) \in B_n\langle k\rangle$ (this is the first street which has at most $n-2k$ paths reaching it from the first street s_1). We are primarily interested in $V^{mir}_{\frac{n}{2}}$.

The Manhattan model is almost identical in this regard, with two differences. The first is that the random configuration of mirrors $\sigma_{mat,t}$ on a street s_t is dependent on whether the street is directed eastbound or westbound. We can let $Z_{(mat,E)}, Z_{(mat,W)}$ the corresponding elements of the Brauer algebra (similar to the mirror case, each only dependent on eastbound or westbound).

Secondly, the diagrams that arise in the Manhattan case actually live in a subalgebra of $\mathbb{B}_{n,1}$. Note that each vertical column of the cylinder $\mathbb{Z}\times\{i\}, i = 1,\dots,n$, is southbound for i odd, northbound for i even. This means that on a chosen street, the vertices i^{N} for i odd, and i^{S} for i even, can be thought of as "entrypoints" to the street. Similarly, each j^{N} for j even, j^{S} for j odd can be thought of as "exitpoints" to the street. In particular, in the diagram which results from the street, *exitpoints must be connected to entrypoints.* This condition can also be thought of as: a NS-path must connect vertices of the same parity, and a bar must connect vertices of different parity. See Figure 4.7.

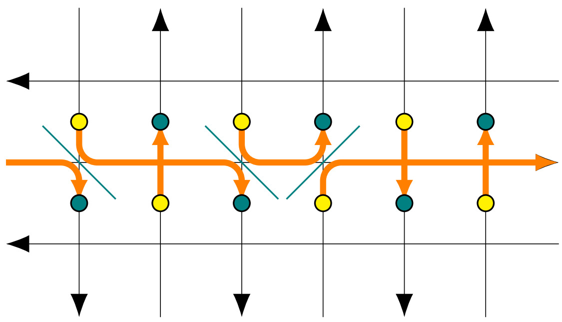


Figure 4.7: An example of paths through a street in the Manhattan model, with entrypoints coloured in yellow, and exitpoints in blue.

Let M_n be the set of diagrams which satisfy the requirement that exitpoints are only connected to entrypoints, and let $\mathbb{M}_{n,1}$ be the (formal) complex span of M_n. This space

$\mathbb{M}_{n,1}$ is a subalgebra of $\mathbb{B}_{n,1}$, indeed it is just the walled Brauer algebra $\mathbb{B}_{n,\frac{n}{2},1}$ from Section 2.1.4, where we have re-ordered the vertices of the diagram so that the vertices "to the left of the wall" (see Figure 2.4) are now those with odd index, and those "to the right of the wall" have even index. It is a straightforward exercise to prove that this reordering is an isomorphism of algebras. Similar to the full Brauer algebra, let $M_n\langle k\rangle$ be the set of diagrams in M_n with k bars, and let M_n^k be those with at least k bars.

Let us assume that the first street, s_1, is eastbound. Now let V_k^{mat} be the random variable given by the smallest t such that $b(\sigma_{mat,1\to t}) \in M_n\langle k\rangle$. Note that the distribution of the random diagram $b(\sigma_{mat,1\to t})$ is described by the element of $\mathbb{M}_{n,1}$:

$$\sum_{g\in B_n} \mathbb{P}[b(\sigma_{mat,1\to t}) = g]\cdot g = \underbrace{Z_{(mat,E)}Z_{(mat,W)}Z_{(mat,E)}\cdots}_{\text{t terms}}$$
$$= \begin{cases} \left(Z_{(mat,E)}Z_{(mat,W)}\right)^{\frac{t}{2}} & t \text{ even} \\ \left(Z_{(mat,E)}Z_{(mat,W)}\right)^{\frac{t-1}{2}} Z_{(mat,E)} & t \text{ odd,} \end{cases}$$

where the equality is included for clarity. We are primarily interested in $V_{\frac{n}{2}}^{mat}$.

Now recall that our method is to condition on there being at most two mirrors per street. Let $U_{\le 2}^{(t)}$ be the event that there are at most two mirrors on a street s_t (this event has probability $\mathbb{P}[U_{\le 2}]$ independent of the street, and the model we are considering). Let $\sigma_{*,t}^{\le 2}$ (resp. $\sigma_{*,t_1\to t_2}^{\le 2}$) be the random configuration of mirrors on the street t (resp. the streets $s_{t_1},\dots,s_{t_2}$) when conditioning on $U_{\le 2}$. Let X_{mir}, X_{mat} be the elements Z_{mir}, $Z_{(mat,E/W)}$, produced when conditioning on $U_{\le 2}$, respectively. (In the Manhattan case, it actually does not matter whether the street is eastbound or westbound). That is, X_{mir} and X_{mat} describe the distributions of $b(\sigma_{mir,t}^{\le 2})$ and $b(\sigma_{mat,t}^{\le 2})$, respectively; for $*$ denoting mir or mat,

$$X_*^t = \sum_{g\in B_n} \mathbb{P}[b(\sigma_{*,1\to t}^{\le 2}) = g]\cdot g.$$

We can write these elements explicitly:

$$X_{mir} = \frac{(1-p)^{n-2}}{\mathbb{P}[U_{\le 2}]}\left[\left(np(1-p)+(1-p)^2\right)\cdot \mathrm{id} + \frac{p^2}{2}\left(\sum_{1\le i<j\le n}(i,j)+(\overline{i,j})\right)\right],$$

and very similarly:

$$X_{mat} = \frac{(1-p)^{n-2}}{\mathbb{P}[U_{\le 2}]}\left[\left(np(1-p)+(1-p)^2\right)\cdot \mathrm{id} + p^2\left(\sum_{j-i \text{ even}}(i,j)+\sum_{j-i \text{ odd}}(\overline{i,j})\right)\right],$$

where we recall that the diagrams (i,j), $(\overline{i,j})$, and id are given in Figure 4.4. Note that Z_{mir}, $Z_{(mat,E/W)}$ can also be explicitly written down as elements of the Brauer algebra (for any p), they are just far more unwieldy.

Similar to above, let $*$ denote mir or mat, and define W_k^* to be the random variable given by the smallest t such that $b(\sigma_{1\to t}^{\le 2}) \in B_n\langle\frac{n}{2}\rangle$. In the next section, we give bounds on how large $W_{\frac{n}{2}}^*$ can be, and then we transfer these bounds to $V_{\frac{n}{2}}^*$.

4.3 Results

Let us first prove part $b)$ of Theorem 4.1.1. Let $*$ denote mir or mat. Let G be a geometric random variable with parameter p^2. We first show that $\mathbb{P}[V^*_{\frac{n}{2}} \le x] \le \mathbb{P}[G \le x]$, for all $x \ge 0$.

Assume that $b(\sigma_{*,1\to t}) \notin B_n\langle \frac{n}{2} \rangle$, that is, after t streets, there are at least two remaining NS-paths. Consider the probability $\mathbb{P}[b(\sigma_{*,1\to t+1}) \in B_n\langle \frac{n}{2} \rangle]$, that after the next street, no NS-paths remain. In order for $b(\sigma_{*,1\to t+1}) \in B_n\langle \frac{n}{2} \rangle$ to hold, there certainly must be a mirror on s_{t+1} reflecting each of the remaining NS-paths - since there are at least two of these, the probability of this is at most p^2. Hence we can say that, given that $b(\sigma_{*,1\to t}) \notin B_n\langle \frac{n}{2} \rangle$,

$$\mathbb{P}[b(\sigma_{*,1\to t+1}) \in B_n\langle \frac{n}{2} \rangle] \le p^2.$$

Now we can easily couple the process with one which enters $B_n\langle \frac{n}{2} \rangle$ at each step with probability exactly p^2. The time taken for this process to enter $B_n\langle \frac{n}{2} \rangle$ can be described by G, and our claim $\mathbb{P}[V^*_{\frac{n}{2}} \le x] \le \mathbb{P}[G \le x]$ follows. Now for $p \le \frac{1}{2}$,

$$\mathbb{P}[V^*_{\frac{n}{2}} \le \alpha p^{-2}] \le \mathbb{P}[G \le \alpha p^{-2}] = 1 - (1-p^2)^{\alpha p^{-2}} \le 2\alpha,$$

the last inequality following from both functions taking the value 0 at $\alpha = 0$, and the differential of the first function being $-(1-p^2)^{\alpha p^{-2}} \log\Big((1-p^2)^{p^{-2}}\Big)$, whose value is less than 2 at $\alpha = 0$ and decreasing as α increases. This completes the proof of part $b)$.

The rest of this section proves part $a)$ of Theorem 4.1.1. We return to our simplified model, assuming at most two mirrors on each street. Observe that if the random diagram $b(\sigma^{\le 2}_{*,t})$ is multiplied with a diagram g which has k bars, the probability that the result has $k+1$ bars is *independent of the chosen diagram* b. This is made precise in the following lemma.

Lemma 4.3.1. *a) Let $g \in B_n\langle k \rangle$, a diagram with k bars. Then $g \cdot b(\sigma^{\le 2}_{*,t}) \in B_n\langle k \rangle \cup B_n\langle k+1 \rangle$, and*

$$g^{mir}_{n,p,k} := \mathbb{P}[g \cdot b(\sigma^{\le 2}_{mir,t}) \in B_n\langle k+1 \rangle] = \frac{1}{\mathbb{P}[U_{\le 2}]}\frac{p^2}{2}(1-p)^{n-2}\binom{n-2k}{2}.$$

b) Let $g \in M_n\langle k \rangle$, a diagram with k bars. Then $g \cdot b(\sigma^{\le 2}_{mat,t}) \in M_n\langle k \rangle \cup M_n\langle k+1 \rangle$, and

$$g^{mat}_{n,p,k} := \mathbb{P}[g \cdot b(\sigma^{\le 2}_{mat,t}) \in M_n\langle k+1 \rangle] = \frac{1}{\mathbb{P}[U_{\le 2}]}p^2(1-p)^{n-2}(\frac{n}{2} - k)^2.$$

Proof. Let us do part $a)$ first. Let $g \in B_n\langle k \rangle$. It is clear that $g(i,j) \in B_n\langle k \rangle$. Further, $g(\overline{i,j}) \in B_n\langle k+1 \rangle$ iff the vertices i^{S} and j^{S} in g lie on NS-paths. There are $\binom{n-2k}{2}$ such pairs. So,

$$\mathbb{P}[b(\sigma^{\le 2}_{mir,t}) = (\overline{i,j}),\ i^{\mathsf{S}}, j^{\mathsf{S}} \text{ on NS paths in } g] = \frac{p^2}{2}\frac{(1-p)^{n-2}}{\mathbb{P}[U_{\le 2}]}\binom{n-2k}{2} = g^{mir}_{n,p,k}.$$

Part $b)$ follows very similarly. Let $g \in M_n\langle k \rangle$. Then $g(\overline{i,j}) \in M_n\langle k+1 \rangle$ iff the vertices

i^{S} and j^{S} in g lie on NS-paths. There are $(\frac{n}{2} - k)^2$ such pairs. So,

$$\mathbb{P}[b(\sigma^{\leq 2}_{mat,t}) = (\overline{i,j}),\ i^{\mathsf{S}}, j^{\mathsf{S}} \text{ on NS paths in } g] = \frac{p^2(1-p)^{n-2}}{\mathbb{P}[U_{\leq 2}]}(\frac{n}{2} - k)^2 = g^{mat}_{n,p,k}.$$

∎

Let $*$ denote *mir* or *mat*. Let $w^*_k = W^*_{k+1} - W^*_k$; this is the number of streets we have to wait between the k^{th} and the $k+1^{th}$ bar being added to the random diagram. Lemma 4.3.1 shows that w^*_k is a geometric random variable, with parameter $g^*_{n,p,k}$. Note that $W^*_{\frac{n}{2}} = \sum_{k=0}^{\frac{n}{2}-1} w^*_k$. The next theorem bounds the probability that $W^*_{\frac{n}{2}}$ is large.

Theorem 4.3.2. *Let $*$ represent mat or mir. Let $p \leq Cn^{-1}$, C a constant. Then for all $\alpha > 0$,*

$$\mathbb{P}[W^*_{\frac{n}{2}} \geq \alpha p^{-2}] \leq A_* e^{-\frac{1}{4C_2}\alpha},$$

where $A_{mir} = cosh(\pi)$, and $A_{mat} = \frac{sinh(\pi)}{\pi}$, and $C_2 = \frac{1}{2}C^2 + C + 1$.

Proof of Theorem 4.3.2. Let us look at the Manhattan case. We first note that, using $p < Cn^{-1}$,

$$\begin{aligned} g^{mat}_{n,p,k} &= \frac{(\frac{n}{2} - k)^2}{\frac{1}{2}n^2 - \frac{1}{2}n + np^{-1} - n + p^{-2} - 2p^{-1} + 1} \\ &\geq \frac{(\frac{n}{2} - k)^2}{(\frac{1}{2}C^2 + C + 1)p^{-2}} \\ &= (\frac{n}{2} - k)^2 C_2^{-1} p^2. \end{aligned}$$

Now, recall that $W^{mat}_{\frac{n}{2}} = \sum_{k=0}^{\frac{n}{2}-1} w^{mat}_k$, and that w^{mat}_k are independent and geometrically distributed with parameter $g^{mat}_{n,p,k}$. Recall also that the moment generating function of a geometric random variable G with parameter λ is given by

$$\mathbb{E}[e^{tG}] = \frac{\lambda}{1 - (1-\lambda)e^t},$$

for $t < -\log(1-\lambda)$. This inequality holds when we set $t = \frac{p^2}{4C_2}$ and $\lambda = p^2$, since $C_2 = \frac{1}{2}C^2 + C + 1 \geq \frac{1}{2}$. We have, using Chebyshev's exponential inequality,

$$\begin{aligned} \mathbb{P}[W^{mat}_{\frac{n}{2}} \geq \alpha p^{-2}] &\leq e^{-\frac{1}{4C_2}\alpha}\mathbb{E}\left[e^{\frac{p^2}{4C_2}W^{mat}_{\frac{n}{2}}}\right] \\ &= e^{-\frac{1}{4C_2}\alpha}\prod_{k=0}^{\frac{n}{2}-1}\left(\frac{g^{mat}_{n,p,k}}{1 - (1 - g^{mat}_{n,p,k})e^{\frac{p^2}{4C_2}}}\right) \\ &= e^{-\frac{1}{4C_2}\alpha}\prod_{k=0}^{\frac{n}{2}-1}\left(1 + \frac{e^{\frac{p^2}{4C_2}} - 1}{1 - (1 - g^{mat}_{n,p,k})e^{\frac{p^2}{4C_2}}}\right). \end{aligned}$$

Using $e^t \leq 2t + 1$, (which holds for $t = \frac{p^2}{4C_2} < 1$, which in turn always holds, since $C_2 \geq \frac{1}{2}$), we have

$$
\begin{aligned}
1-(1-g_{n,p,k}^{mat})e^{\frac{p^2}{4C_2}} &\geq 1-(1-(\frac{n}{2}-k)^2C_2^{-1}p^2)(\frac{1}{2}C_2^{-1}p^2+1)\\
&\geq C_2^{-1}p^2((\frac{n}{2}-k)^2-\frac{1}{2}),
\end{aligned}
$$

which gives:

$$
\begin{aligned}
\mathbb{P}[W_{\frac{n}{2}}^{mat}\geq \alpha p^{-2}] &\leq e^{-\frac{1}{4C_2}\alpha}\prod_{k=0}^{\frac{n}{2}-1}\left(1+\frac{\frac{1}{2}C_2^{-1}p^2}{C_2^{-1}p^2((\frac{n}{2}-k)^2-\frac{1}{2})}\right)\\
&= e^{-\frac{1}{4C_2}\alpha}\prod_{k=1}^{\frac{n}{2}}\left(1+\frac{1}{2k^2-1}\right)\\
&\leq e^{-\frac{1}{4C_2}\alpha}\prod_{k=1}^{\frac{n}{2}}\left(1+\frac{1}{k^2}\right)\\
&\leq \frac{\sinh\pi}{\pi}e^{-\frac{1}{4C_2}\alpha},
\end{aligned}
$$

as desired. In the last inequality we used the product formula $\sin(\pi z)=\pi z\prod_{\nu=1}^{\infty}(1-\frac{z^2}{\nu^2})$, with $z=i$.

The Mirror model case is almost identical; all the above working is the same except the expression $(\frac{n}{2}-k)^2$ is replaced with $\frac{1}{2}\binom{n-2k}{2}$. This gives

$$
\begin{aligned}
\mathbb{P}[W_{\frac{n}{2}}^{mat}\geq \alpha p^{-2}] &\leq e^{-\frac{1}{4C_2}\alpha}\prod_{k=0}^{\frac{n}{2}-1}\left(1+\frac{\frac{1}{2}}{\frac{1}{2}\binom{n-2k}{2}-\frac{1}{2}}\right)\\
&\leq e^{-\frac{1}{4C_2}\alpha}\prod_{k=1}^{\frac{n}{2}}\left(1+\frac{4}{2k(2k-1)}\right)\\
&\leq \cosh(\pi)e^{-\frac{1}{4C_2}\alpha}
\end{aligned}
$$

as desired, where for the last equality we used the product formula $\cos(\pi z)=\prod_{\nu=1}^{\infty}(1-\frac{4z^2}{(2\nu-1)^2})$, with $z=i$.

■

We can now compare the full models with the models assuming at most two mirrors per street. Let $t\in\mathbb{N}$. Let $\tau(t)$ be the random variable given by the number of the first t streets which have at most 2 mirrors. We see that $\tau(t)$ is binomially distributed with parameters $(t,\mathbb{P}[U_{\leq 2}])$. Essentially what we would like to say is that if we omit each street which has more than 2 mirrors, we do not, in distribution, add any bars.

This sounds like it should follow from the remark (4.1), but it is more subtle. Let us illustrate why: certainly if the product of two diagrams ab has k bars, then we can conclude that each of a and b have no more than k bars. However, if abc has k bars, it is very possible that ac has more than k bars. So, when removing factors from the middle of a product, there is more to be proved.

Lemma 4.3.3. *Let $*$ denote mir or mat. Then $\mathbb{P}[V_k^*\leq t]\geq\mathbb{P}[W_k^*\leq\tau(t)]$.*

Recall that $g\in B_n^k$ iff g has at least k bars, and $g\in M_n^k$ similar. Note that $V_k^*\leq t$ iff

$b(\sigma_{*,1\to t}) \in B_n^k$; similar for W_k^*. So Lemma 4.3.3 can be rewritten as:

$$\mathbb{P}[b(\sigma_{*,1\to t}) \in B_n^k] \geq \mathbb{P}[b(\sigma^{\leq 2}_{*,1\to\tau(t)}) \in B_n^k]. \tag{4.2}$$

We postpone the proof of Lemma 4.3.3, and first see how it is implemented, combining with Theorem 4.3.2 in proving part a) of Theorem 4.1.1.

Proof of part a) of Theorem 4.1.1. Recall that we assume $p \leq Cn^{-1}$, C a constant. We approximate $b(\sigma_{*,1\to t})$ with $b(\sigma^{\leq 2}_{*,1\to\tau(t)})$, that is, we approximate by ignoring streets which have more than two mirrors. Since the expected number of mirrors per street is at most C, we expect (at least for C small) the proportion of streets with at most two mirrors to be large. Indeed:

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}[U_{\leq 2}] &= \lim_{n\to\infty}(1-p)^{n-2}\left((1-p)^2 + np(1-p) + \binom{n}{2}p^2\right)\\ &\geq \lim_{n\to\infty}(1-p)^{cp^{-1}-2}\left((1-p)(1-p+C) + \frac{C}{2}(C-p)\right),\end{aligned}$$

the limit of which is $(\frac{1}{e})^C(1 + C + \frac{C^2}{2}) =: C_3$. We can pick $n \in \mathbb{N}$ such that $\mathbb{P}[U_{\leq 2}] > \frac{3C_3}{4}$. Recalling $\tau(t)$ is binomially distributed with parameters $(t, \mathbb{P}[U_{\leq 2}])$, by Hoeffding's inequality,

$$\mathbb{P}\left[\tau(t) \leq \frac{C_3 t}{2}\right] \leq \exp\left[-2t\left(\mathbb{P}[U_{\leq 2}] - \frac{C_3}{2}\right)^2\right] \leq \exp\left[-2t\left(\frac{C_3}{4}\right)^2\right], \tag{4.3}$$

for n large enough.

Let $t = \alpha p^{-2}$. Now using Lemma 4.3.3,

$$\begin{aligned}\mathbb{P}[V_k^* \geq \alpha p^{-2}] &\leq \mathbb{P}[W_k^* \geq \tau(\alpha p^{-2})]\\ &\leq \mathbb{P}\left[W_k^* \geq \frac{C_3}{2}\alpha p^{-2}\right] + \mathbb{P}\left[\tau(t) \leq \frac{1}{2}\alpha p^{-2}\right]\\ &\leq A_* \exp\left[-\frac{C_3}{8C_2}\alpha\right] + \exp\left[-2\alpha p^{-2}\left(\frac{C_3}{4}\right)^2\right]\\ &\leq 2A_* \exp\left[-\frac{1}{8e^C}\alpha\right],\end{aligned}$$

where the second to last inequality is from Theorem 4.3.2 and equation (4.3), and the last is for p small enough. ∎

It remains to prove Lemma 4.3.3.

Proof of Lemma 4.3.3. We prove the inequality (4.2). We fix n, and work by induction on t and k. The inequality is trivially true for $k = 0$ (and any t), and for $t = 1$ (and any k).

Assume the Lemma holds for the parameters $(t-1, k)$, $(t-1, k-1)$, and $(t, k-1)$. The left hand side of equation (4.2) is:

$$\begin{aligned}\mathbb{P}[b(\sigma_{*,1\to t}) \in B_n^k] &= \mathbb{P}\left[b(\sigma_{*,1\to t-1})b(\sigma_{*,t}) \in B_n^k \mid U_{>2}^{(t)}\right] \cdot (1 - \mathbb{P}[U_{\le 2}]) \\ &\quad + \mathbb{P}\left[b(\sigma_{*,1\to t-1})b(\sigma_{*,t}) \in B_n^k \mid U_{\le 2}^{(t)}\right] \cdot \mathbb{P}[U_{\le 2}] \\ &\ge \mathbb{P}\left[b(\sigma_{*,1\to t-1}) \in B_n^k \mid U_{>2}^{(t)}\right] \cdot (1 - \mathbb{P}[U_{\le 2}]) \\ &\quad + \mathbb{P}\left[b(\sigma_{*,1\to t-1})b(\sigma_{*,t}^{\le 2}) \in B_n^k \mid U_{\le 2}^{(t)}\right] \cdot \mathbb{P}[U_{\le 2}],\end{aligned}$$

where we have noted that the number of bars in the product $b(\sigma_{*,1\to t-1})b(\sigma_{*,t})$ cannot be less that in $b(\sigma_{*,1\to t-1})$, and that $b(\sigma_{*,t})$ is equal to $b(\sigma_{*,t}^{\le 2})$ when conditioned on $U_{\le 2}^{(t)}$. Now the above is at least:

$$\begin{aligned}&\ge \mathbb{P}\left[b(\sigma_{*,1\to\tau(t-1)}^{\le 2}) \in B_n^k \mid U_{>2}^{(t)}\right] \cdot (1 - \mathbb{P}[U_{\le 2}]) + \mathbb{P}[b(\sigma_{*,1\to\tau(t)}^{\le 2}) \in B_n^k \mid U_{\le 2}^{(t)}] \cdot \mathbb{P}[U_{\le 2}] \\ &= \mathbb{P}[b(\sigma_{*,1\to\tau(t)}^{\le 2}) \in B_n^k],\end{aligned}$$

where in the inequality we used the inductive assumption on t and the final Lemma below, and in the equality we used the fact that under $U_{>2}^{(t)}$, $\tau(t) = \tau(t-1)$. The proof of the final Lemma therefore concludes the whole proof.

Lemma 4.3.4. *We have that* $\mathbb{P}\left[b(\sigma_{*,1\to t-1})b(\sigma_{*,t}^{\le 2}) \in B_n^k \mid U_{\le 2}^{(t)}\right] \ge \mathbb{P}[b(\sigma_{*,1\to\tau(t)}^{\le 2}) \in B_n^k \mid U_{\le 2}^{(t)}]$.

Proof. To prove the claim, we split the left hand term based on whether or not $b(\sigma_{*,t}^{\le 2})$ adds a bar to $b(\sigma_{*,1\to t-1})$:

$$\begin{aligned}\text{LHS} &= \mathbb{P}\left[b(\sigma_{*,1\to t-1}) \in B_n^{k-1} \mid U_{\le 2}^{(t)}\right] \cdot g_{n,p,k}^* + \mathbb{P}\left[b(\sigma_{*,1\to t-1}) \in B_n^k \mid U_{\le 2}^{(t)}\right] \\ &\ge \mathbb{P}\left[b(\sigma_{*,1\to\tau(t-1)}^{\le 2}) \in B_n^{k-1} \mid U_{\le 2}^{(t)}\right] \cdot g_{n,p,k}^* + \mathbb{P}\left[b(\sigma_{*,1\to\tau(t-1)}^{\le 2}) \in B_n^k \mid U_{\le 2}^{(t)}\right] \\ &= \mathbb{P}\left[b(\sigma_{*,1\to\tau(t-1)}^{\le 2})b(\sigma_{*,t}^{\le 2}) \in B_n^k \mid U_{\le 2}^{(t)}\right],\end{aligned}$$

where in the inequality we used the inductive assumption on t and k. Now recalling that under $U_{\le 2}$, $\tau(t) = \tau(t-1) + 1$, the result follows. This concludes the proof of Lemma 4.3.3 and part a) of Theorem 4.1.1.

■

Chapter 5

Quantum Spin Systems on the complete graph

5.1 Classical and quantum spin systems

We give a short introduction here to the area of classical and quantum spins systems. We follow Ruelle [86] and Friedli and Velenik [37]. Classical statistical mechanics attempts to derive the macroscopic laws of nature (such as thermodynamics) from laws of the interactions of particles on a microscopic scale. One of the remarkable and most studied features of these systems is that there are sometimes abrupt changes in the behaviour of the system, called phase transitions, as the parameters (such as temperature) are varied. Practically, one uses mathematical models of such systems which are simplifications (to varying degrees) of the reality, where the laws of interactions of the particles are given, from which one attempts to derive macroscopic behaviours. Even with these simplifications, in many models one can derive results, including rigorous ones, that show the models has phase transitions. This is not observed mathematically for finite systems, but in taking the limit of the models as the number of particles tends to infinity, phase transitions can be observed.

A very illustrative example of a classical spin system comes in the form of the classical Heisenberg model. This is a simplified model of ferromagentism, the phenomenon where some materials, under some conditions, retain a magnetism after an external magnetic field has been applied and then taken away. The model simplifies such a situation to describe the material as a large number of particles arranged in a lattice, which do not move, but interact via the directions in which they are magnetised. That is, particles close together want to be magnetised in the same direction. Mathematically, the model describes a probability measure on possible configurations of particles arranged on a lattice (i.e. a large box in $\mathbb{Z}^d$), where a configuration assigns each particle an orientation (called a spin) in 3D space (i.e. an element of $\mathbb{S}^2$).

More precisely, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph. It is illustrative to think of $\mathcal{G}$ as a finite box in $\mathbb{Z}^d$; note that the behaviour of the model depends on the dimension d. A particle at a site $i \in \mathcal{V}$ is given a spin $\sigma_i \in \mathbb{S}^2$, the two-sphere. Allow a parameter β to represent inverse temperature. Then, for a given β, the configurations $\sigma = (\sigma_i)_{1\leq i\leq n}$ occur with density

$$\phi_{H,\mathcal{G},\beta}(\sigma) = \phi_\beta(\sigma) = \frac{1}{Z(\mathcal{G},\beta)} e^{-\beta H(\sigma)}. \tag{5.1}$$

Here the function $Z(\mathcal{G},\beta) = \int d\sigma \ e^{-\beta H(\sigma)}$ is called the partition function and is the normalisation constant which makes the measure a probability measure, and

$$H(\sigma) = -\sum_{\{i,j\}\in\mathcal{E}} \sigma_i \cdot \sigma_j \tag{5.2}$$

is the Hamiltonian, describing the energy of the configuration. Note that the sum is over pairs of vertices i and j which are nearest neighbours in the underlying graph $\mathcal{G}$. This function appropriately describes a ferromagnetic interaction, that is, an interaction between spins where the spins want to be aligned: the more aligned the spins are, the lower the energy. The configurations with lowest energy are those with the spins at all vertices pointing in the same direction. These are relatively few in number, compared with the vast number of configurations which would give a large energy, with neighbouring spins being much less aligned. The Heisenberg model thereby models a (large) block of some ferromagnetic material in dimension d made of many particles (arranged in a lattice), where each particle is magnetised in some direction of 3-dimensional space. The particles then exert magnetic forces on one another, which one assumes to have short range, so that the assumption that only nearest neighbours interact is sensible.

For high temperatures (β small), the many configurations with high energy dominate the measure ϕ_β - their entropy overcomes the exponent β - and the system is said to be disordered. One of the central questions in studying spin systems is whether for low temperatures (β large), one finds that the low energy configurations dominate, in which case the system is said to be ordered. This is indeed the case in dimensions $d \geq 3$, but not for $d = 1, 2$; the former is due to Fröhlich, Simon and Spencer [38], and the latter is due to the Mermin-Wagner theorem. We will state these results more precisely later in this introduction. For $d \geq 3$, this is a heuristic description of an example of there being two regions of parameter space where the measure ϕ_β behaves very differently. Further, it turns out that these regions are separated by a critical temperature, say β_c, at which the properties of the measure ϕ_β abruptly change - a phase transition. As noted above, such abrupt changes are mathematically only observed when we take a limit of the measure ϕ_β as the graph becomes infinite (for example when the box in $\mathbb{Z}^d$ grows to become the whole lattice).

There are many ways one can observe whether a phase transition occurs, that is, many mathematical quantities one can derive from the model, which are in some sense sensible, which show abrupt changes at the critical temperature. Let us describe four important such ways. Note one can sometimes show that the various notions of a phase transition are equivalent, but this is not always possible. The first way of observing a phase transition is by studying how particles at large distances can affect one another. Imagine the classical Heisenberg model, $d \geq 3$ in its ordered region (low temperature) as heuristically described above, with it being very likely that all spins are aligned in some direction. It follows that if we change the spin at one site, then with high probability all the others follow it - in particular spins arbitrarily far away from it change. In contrast, in the disordered, high temperature region, it is perhaps natural to think that changing a spin at one vertex has little effect on spins far away. Mathematically, we can study the correlation of the spins at two distant particles. For the classical Heisenberg model, this is

$$\lim_{|x_1-x_2|\to\infty}\lim_{\mathcal{G}\nearrow\mathbb{Z}^d}\langle\sigma_{x_1}\cdot\sigma_{x_2}\rangle_{H,\mathcal{G},\beta}, \tag{5.3}$$

where $\langle f(\sigma)\rangle_{H,\mathcal{G},\beta}$ is the expectation of f with respect to our probability measure. This quantity being zero indicates the particles do not affect each other. Sometimes this quantity is positive for low temperatures, and the point at which it becomes positive marks a phase transition. This includes the case of the classical Heisenberg model, $d\geq 3$, and this is the form in which the transition is proved in [38]. When $d=2$ the limit (5.3) is zero for all $\beta>0$ (i.e. there is no phase transition of this type), and the decay to zero of (5.3) is at least polynomial in speed for all temperatures, proved by McBryan and Spencer [71]. We should note though that it is not yet proved whether or not there is a more subtle transition in dimension $d=2$ in the sense that for high temperatures, the decay of (5.3) to zero is exponential, and for lower temperatures, it is a power law. Such a transition, known as a Berezinskii-Kosterlitz-Thouless phase transition, does occur in other two-dimensional models, including models with continuous symmetry, most notably the "XY" model, which is the Heisenberg model with spins on the circle rather than the sphere.

The second method of observing a phase transition is magnetisation. The "ferromagnetism" described earlier is exactly an example of this. We study the model with an infinitesimal external magnetic field, (i.e. an external magnetic field whose strength is reduced to zero). To be precise, we can modify the classical Hamiltonian H (5.2):

$$H(\sigma)=-\sum_{\{i,j\}\in\mathcal{E}}\sigma_i\cdot\sigma_j-h\sum_{i\in\mathcal{V}}b\cdot\sigma_i, \tag{5.4}$$

where $b\in\mathbb{S}^2$ denotes some direction in 3-space, and $h\in\mathbb{R}$ is a strength parameter. This models the spin at each vertex being pulled by an external field in some particular direction b with strength h. One can imagine that turning on such a field pulls all the spins to point in the direction b, with high probability; the question of interest is whether this structure remains once the field is turned off. Mathematically, this is studied by analysing the quantity

$$\frac{\partial\Phi(\beta,h)}{\partial h}\Big|_{h=0}, \tag{5.5}$$

the magnetisation in the direction b, where $\Phi(\beta,h)=\lim_{|\mathcal{V}|\to\infty}\frac{1}{|\mathcal{V}|}\log Z(\mathcal{G},\beta,h)$ is the free energy (see below). In several models this can be shown to be equal to

$$\lim_{|\mathcal{V}|\to\infty}\lim_{h\searrow 0}\langle(\frac{1}{|\mathcal{V}|}\sum_{x\in\mathcal{V}}\sigma_x)\cdot b\rangle_{H(h),\mathcal{G},\beta}, \tag{5.6}$$

(where we have highlighted the dependence of H on h), which is far more intuitive - it is the expected average amount that all the spins are pointing in the direction b. If magnetisation does not occur (which happens at high temperatures) this average is zero, and then, at low temperatures, one can ask whether it is strictly positive, that is, whether magnetisation occurs.

A slightly less intuitive way to show a phase transition is through an expression called the free energy. It is a function of β, and a point where it is non-analytic indicates a phase transition at that value of β. For the graphs $\mathcal{G}$ we are most interested in, $\mathcal{G} = \mathbb{Z}^d$, we define it as

$$\Phi(\beta) = \lim_{\mathcal{G}\to\mathbb{Z}^d} \frac{1}{|\mathcal{V}|} \log(Z(\mathcal{G},\beta)), \tag{5.7}$$

for $\mathcal{G}$ taken to be successively large boxes in $\mathbb{Z}^d$. The free energy in some sense gives the energy which dominates the measure at the given temperature; let us heuristically explain how. Imagine that there are finitely many possible energies E_i that the Hamiltonian H can produce, and there are d_i many configurations with energy E_i. Then we can write $\frac{1}{|\mathcal{V}|}\log Z(\mathcal{G},\beta)$ as a sum of the form

$$\frac{1}{|\mathcal{V}|} \log Z(\mathcal{G},\beta) = \frac{1}{|\mathcal{V}|} \log \sum_i d_i e^{-\beta E_i} = \frac{1}{|\mathcal{V}|} \log \sum_i e^{-\beta E_i + \log d_i}. \tag{5.8}$$

Now taking the limit as $|\mathcal{V}| \to \infty$ essentially pulls out the largest $-\beta E_i + \log d_i$, which indeed in some sense marks the energy which dominates the measure at the given temperature, and the expression, suitably, takes into account both the energy E_i and entropy d_i. The free energy is the quantity which we study in detail for our specific models in this and the next chapters.

The fourth way we will describe to observe a phase transition is via Gibbs states. In the classical Heisenberg model, a boundary condition is a fixing of the spins on the boundary of $\mathcal{G} \subset \mathbb{Z}^d$. One can study the possible suitable limits of the measure ϕ_β, if one takes different boundary conditions or infinitesimal external magnetic fields (see Chapter 6 of [37] for formal definitions). These suitable limits are the infinite Gibbs measures, or Gibbs states. For high temperatures (β small), there is usually a unique Gibbs state, which corresponds to the idea that the system "forgets" the external magnetic field, or boundary conditions. One can say there is a phase transition if at low temperatures, (β large), there is more than one Gibbs state. This is what concerns the Mermin-Wagner theorem noted above, which shows that for all positive temperatures, all Gibbs states are invariant under the action of $SO(3)$, see Theorem 9.2 of [37] (i.e. states in which the spins tend to point in one direction are excluded). In the classical case defining Gibbs states rigorously is done through the DLR (Dobrushin–Lanford–Ruelle) equations: a Gibbs state is a measure ϕ on configurations on $\mathbb{Z}^d$, such that if one conditions on the spins outside some finite set $\mathcal{G}$, one obtains the measure (5.1) on $\mathcal{G}$, with boundary conditions induced by the spins outside $\mathcal{G}$.

The study of a quantum spin systems cover the same phenomena as a classical one, with the addition that quantum behaviour is accounted for. It should be noted that one of the postulates of quantum mechanics is that the theory should in some way contain the classical version, via some suitable limit.

Mathematically, instead of working with a probability measure, we work with Hermitian operators on some Hilbert space $\mathcal{H}$. The Hamiltonian H is such an operator, and the possible energies are the spectrum of H. In place of the measure on configurations, we study the operator $\frac{1}{Z(\beta)}e^{-\beta H}$, where $Z(\beta) = \mathrm{Tr}\left[e^{-\beta H}\right]$ is the partition function, the normalisation constant which makes the operator have trace 1. In keeping with our heuristics above, imagining that H has a finite spectrum $\{E_i\}$, with each eigenvalue having eigenspace of dimension d_i, one has the same form for the partition function

$$Z(\beta) = \sum_i d_i e^{-\beta E_i} = \sum_i e^{-\beta E_i + \log d_i}. \tag{5.9}$$

One interpretation of classical models lying within the quantum setup is the following. If the Hamiltonian is written as a sum of operators (see (5.10) for instance) which commute with one another, then the operators share eigenspaces, and one can think of the model as classical - as a probability measure on the eigenvalues with weights proportional to the dimension of the associated eigenspace. We will see an instance of this in Section 5.2.2 when the 'XXZ" model (5.15) becomes the classical Ising model when $K_1 = 0$. In contrast, when the Hamiltonian is a sum of operators which do not commute, we say the model is quantum.

Let $S \in \frac{1}{2}\mathbb{N}$. In the quantum Heisenberg model, the spins σ_i are replaced with Hermitian operators on a copy of $\mathbb{C}^{2S+1}$ at $i \in \mathcal{V}$, $\mathbb{C}_i^{2S+1}$. The Hamiltonian is an operator acting on tensor space $(\mathbb{C}^{2S+1})^{\otimes \mathcal{V}}$. For the quantum Heisenberg model, the Hamiltonian is:

$$H = -\sum_{\{i,j\}\in\mathcal{E}} (\boldsymbol{S}_i \cdot \boldsymbol{S}_j), \tag{5.10}$$

where $\boldsymbol{S}_i = (S_i^{(1)}, S_i^{(2)}, S_i^{(3)})$, $(\boldsymbol{S}_i \cdot \boldsymbol{S}_j) = (S_i^{(1)}S_j^{(1)} + S_i^{(2)}S_j^{(2)} + S_i^{(3)}S_j^{(3)})$ and $S_i^{(1)}, S_i^{(2)}, S_i^{(3)}$ are explicit Hermitian operators acting on $\mathbb{C}_i^{2S+1}$. These operators are analogies of the three co-ordinate components of the classical σ_i. The parameter S is called the spin quantum number. When it is unambiguous, we will just refer to it as the spin. The four methods of observing a phase transition described above for the classical case all have their analogues in the quantum case.

The free energy in this setting is $\lim_{|\mathcal{V}|\to\infty} \frac{1}{|\mathcal{V}|} \log Z(\mathcal{G}, \beta) = \frac{1}{|\mathcal{V}|} \log \mathrm{Tr}\left[e^{-\beta H}\right]$, which one can again think of as pulling out the largest of the values $-\beta E_i + \log d_i$, E_i in the spectrum of H, d_i the dimension of its eigenspace. For observing magnetisation, one amends the Hamiltonian (5.10) similarly to the classical case (5.4), as

$$H = -\sum_{\{i,j\}\in\mathcal{E}} (\boldsymbol{S}_i \cdot \boldsymbol{S}_j) - h\sum_{i\in\mathcal{V}} \boldsymbol{S}_i \cdot b, \tag{5.11}$$

where $b \in \mathbb{S}^2$ and $\boldsymbol{S}_i \cdot b = \sum_{k=1}^3 b_k S_i^{(k)}$. One then studies the quantity (5.5) using the quantum version of the free energy, or quantities of the form of (5.6), ie:

$$\lim_{|\mathcal{V}|\to\infty} \langle (\frac{1}{|\mathcal{V}|}\sum_{x\in\mathcal{V}} \boldsymbol{S}_x) \cdot b \rangle_{H,\mathcal{G},\beta}, \tag{5.12}$$

where for an operator A on the phase space $V^{\otimes n}$, $\langle A \rangle_{H,\mathcal{G},\beta} = \frac{1}{Z(\mathcal{G},H,\beta)} \mathrm{Tr}\left[A \cdot e^{-\beta H}\right]$. The spin-spin correlations (5.3) have their analogue in the quantum case too:

$$\lim_{|x_1 - x_2|\to\infty} \langle \boldsymbol{S}_{x_1} \cdot \boldsymbol{S}_{x_2} \rangle_{H,\mathcal{G},\beta}. \tag{5.13}$$

As in the classical case, in dimensions 1 and 2 this limit is zero for all $\beta > 0$, see the lecture notes of Ueltschi [100]. There is a notion of Gibbs states in the quantum case too. One needs the theory of C^*-algebras to define them rigorously. See [86] for formal definitions. In [14], the authors give a certain heuristic argument which points towards the structure of the Gibbs states of several models (including the quantum Heisenberg model). In both this Chapter 5 and 6, we observe that we can make analogous heuristic arguments for the models that we study in those chapters. See Sections 5.2.2, 5.2.3 and 6.1.5.

While models on $\mathbb{Z}^d$ are already simplifications of the real-life situations that they model, often working with them can be difficult. For example, the question of whether there is a phase transition for the quantum Heisenberg model on $\mathbb{Z}^d$, $d \geq 3$, is an open problem (let alone the nature of such a transition). Notice the difference from the classical case. One way of gaining intuition for these models is via the mean field approximation, where the effect of all particles on any one particle is approximated by a single, averaged effect.

Mathematically, this amounts to studying the models on the complete graph on n particles (i.e. all particles are neighbours of one another). Often, this makes computations (for example computing the free energy) easier. Note that on the complete graph, some of the methods for detecting phase transitions have workable analogues, and some do not. The notion of Gibbs states is not well-defined, and correlations such as (5.13) do not have meaning since the distance between any pair of particles is 1. However, an analogue of the free energy is well-defined (and one can sometimes study its analyticity properties), and expressions such as (5.12) make sense, and can often be computed.

In several models the mean-field approximation gives exactly the corresponding quantities for the model on $\mathbb{Z}^d$, and in some cases it is a good approximation, particularly when the interactions between particles are long range, or the dimension is high (both cases meaning that the valency of a vertex in the underlying graph is high). In this Chapter 5, we make such a mean field approximation, and study the free energy of a class of quantum spin systems (which includes the spin $S = \frac{1}{2}$ quantum Heisenberg model) on the complete graph. We also compute certain observables of the form (5.12). In Chapter 6 we make a similar approximation, studying models on the complete bipartite graph.

We will study three simple generalisations of the quantum Heisenberg model, which will appear in, and indeed be a large focus of, this chapter. Equivalent classical models exist for each of the three models, by amending the the Hamiltonians accordingly. Firstly, the quantum Heisenberg antiferromagnet has Hamiltonian

$$H = + \sum_{\{i,j\} \in \mathcal{E}} (\boldsymbol{S}_i \cdot \boldsymbol{S}_j), \tag{5.14}$$

the same as the ferromagnet but multiplied by (-1). This model favours adjacent spins which are anti-aligned. On a bipartite graph, it is not hard to see that configurations (in the classical model) with lowest energy are those with all spins aligned on one subgraph, and all spins on the other subgraph aligned in the opposite direction. Dyson, Lieb and Simon [33] showed a phase transition for a large class of models on $\mathbb{Z}^d$, $d \geq 3$, including the quantum Heisenberg antiferromagnet with spin $S \geq 1$, using the method known as reflection positivity; this was extended to $d \geq 3$, $S \geq \frac{1}{2}$ by Kennedy, Lieb and Shastry [59]. Notice the contrast with the ferromagnet remaining an open problem.

Secondly, the Heisenberg XXZ model is the Heisenberg model with Hamiltonian tweaked to become:

$$H = -\sum_{\{i,j\}\in\mathcal{E}} \left(K_1 S_i^{(1)} S_j^{(1)} + K_2 S_i^{(2)} S_j^{(2)} + K_1 S_i^{(3)} S_j^{(3)} \right), \tag{5.15}$$

that is, we give a certain weight K_2 to the interaction in the $S^{(2)}$ direction, and a second weight K_1 to the other two directions. For example, if $K_1 > 0$ and $K_2 < 0$, the system wants adjacent spins which point in the $1-3$ plane to be aligned, but those pointing in the 2-axis to be anti-aligned. The name "XXZ" is simply from there being two weights the same and one different. Notice that the ferromagnetic Heisenberg model is the special case $K_1 = K_2 = 1$, and the antiferromagnet is $K_1 = K_2 = -1$. One can perform a unitary transformation (conjugate the Hamiltonian by a unitary matrix) so that the K_2 weight appears in front of the $S^{(1)}$ or the $S^{(3)}$ term instead; another unitary transformation can replace K_1 with $-K_1$. See the lecture notes of Ueltschi [100] for details. Fröhlich and Lieb [39] and Kennedy [58] showed that for $K_2 > K_1 > 0$, and for dimensions $d \geq 2$ there is a phase transition in the spin $S = \frac{1}{2}$ model in the sense that for low temperatures there is long range order, that is, the limit $\lim_{|x_1-x_2|\to\infty} \langle S_{x_1}^{(2)} S_{x_2}^{(2)} \rangle_{H,\mathbb{Z}^d,\beta}$ (see (5.13)) is strictly positive. Notice the difference from the ferromagnetic model, where there is no transition in dimension $d = 2$.

Thirdly, the Heisenberg bilinear-biquadratic model has Hamiltonian

$$H = -\sum_{\{i,j\}\in\mathcal{E}} \left(J_1 (\boldsymbol{S}_i \cdot \boldsymbol{S}_j) + J_2 (\boldsymbol{S}_i \cdot \boldsymbol{S}_j)^2 \right). \tag{5.16}$$

This model is studied in spin $S = 1$, where it is the most general $SU(2)$-invariant model, where here the invariance is the action on the spin operators. The first term in (5.16) is essentially the ferromagnet for $J_1 > 0$ and the antiferromagnet for $J_1 < 0$, indeed the ferromagnet is the special case $J_1 = 1$, $J_2 = 0$, and the antiferromagnet is $J_1 = -1$, $J_2 = 0$. The second term, for $J_2 > 0$, prefers adjacent spins to be either aligned or anti-aligned, but not orthogonal to one another, and vice-versa for $J_2 < 0$. Ueltschi [101] showed that there is a phase transition for $d \geq 3$ and $0 \leq J_1 \leq \frac{1}{2} J_2$, in the sense that for low temperatures, there is a "nematic order", and on the line $J_1 = 0 < J_2$ there is a "Néel order", both types of long range order. Lees [64] showed that for $d \geq 3$, $J_1 \leq 0 \leq J_2$ and $-J_1/J_2 < \alpha = \alpha(d)$ some constant depending on dimension, there is Néel order for low temperatures. See Section 7 of [101], as well as [102], for a full description and the expected phase diagram on $\mathbb{Z}^d$, $d \geq 3$. See [36], [54], [95], [98], for further work on this model.

One interesting difference between the quantum and classical model is the following. In the nematic region in the quantum model, $0 < J_1 < J_2$, one expects the extremal Gibbs states to be indexed by $\mathbb{RP}^2$. Following the classical intuition, one would expect these extremal states to arise as limits, for $b \in \mathbb{S}^2$:

$$\langle \cdot \rangle_b = \lim_{h \searrow 0} \lim_{|\mathcal{V}|\to\infty} \langle \cdot \rangle_{b, H(h,\beta,J_1,J_2), \mathcal{G}}, \tag{5.17}$$

where

$$H(h,\beta,J_1,J_2) = -\sum_{\{i,j\}\in\mathcal{E}} \left(J_1(\boldsymbol{S}_i\cdot\boldsymbol{S}_j) + J_2(\boldsymbol{S}_i\cdot\boldsymbol{S}_j)^2\right) - \sum_{\{i\}\in\mathcal{V}} h(\boldsymbol{S}_i\cdot b)^2. \tag{5.18}$$

Here the magnetisation term encourages spins to be either aligned or anti-aligned in the direction of b. However, in the quantum case, the correct Gibbs state is expected [102] to arise with magnetisation term $+\sum_{\{i\}\in\mathcal{V}} h(\boldsymbol{S}_i\cdot b)^2$ (with the sign changed to a plus), which encourages spins to lie in the plane orthogonal to b.

We will study the latter two models in the rest of Chapter 5 as they are special cases of the more general model (5.20) studied there.

5.2 Introduction to quantum spin systems on the complete graph

The remainder of this Chapter 5 presents the results from the paper "The free energy of a class of $O_{2S+1}(\mathbb{C})$-invariant spin $\frac{1}{2}$ and 1 quantum spin systems on the complete graph" [89]. We present the paper essentially unchanged, with references to our use of representation-theoretic tools from the previous sections given as appropriate.

In this paper, we study a certain two-parameter family of quantum spin systems on the complete graph which generalises the spin $S=\frac{1}{2}$ quantum Heisenberg model, and which in particular has a certain invariance under the action of the orthogonal group $O(\theta) = O(2S+1) = O_{2S+1}(\mathbb{C})$, where $\theta = 2S+1$, and S is the spin of the model. It is equivalent in spin $S=\frac{1}{2}$ to the XXZ model, and in spin $S=1$ to the bilinear-biquadratic Heisenberg model, defined above. The work is motivated by a paper of Björnberg, whose model is $GL(\theta)$-invariant. In spin $S=\frac{1}{2}$ and $S=1$ we give an explicit formula for the free energy for all values of the two parameters, and for spin $S>1$ for when one of the parameters is non-negative. This allows us to draw phase diagrams, and determine critical temperatures. For spins $S=\frac{1}{2}$ and $S=1$, we give a magnetisation, the left and right derivatives of the free energy as the strength parameter of a certain magnetisation term in the Hamiltonian similar to (5.11) tends to zero. We also give a formula for a certain total spin observable in the style of a paper of Björnberg, Fröhlich and Ueltschi [14]. We also give a certain heuristic argument which points towards the structure of the Gibbs states in the models we study, analogously to an argument in [14].

The key technical tool used in this paper (and the next, in Chapter 6) is Schur-Weyl duality. Fundamentally, the partition function of the model is the trace of the action $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ (see 3.12) of some element of the Brauer algebra $\mathbb{B}_{n,\theta}$ on tensor space $V^{\otimes n}$. We can therefore use Schur-Weyl duality to write the partition function in terms of the irreducible characters of the symmetric group and the Brauer algebra, which is instrumental in being able to then take limits.

Let us present a more detailed introduction to this paper. Quantum spin systems and their phase transitions have been studied widely. Mermin and Wagner showed that no model with continuous symmetry has a phase transition in dimensions 1 and 2. This was done in particular for the quantum Heisenberg model [74] and the classical model [73]. Dyson, Lieb and Simon [33] showed a transition for a large class of models on $\mathbb{Z}^d$, $d \geq 3$, including the quantum Heisenberg antiferromagnet with spin $S \geq 1$. A phase transition on $\mathbb{Z}^d$, $d \geq 3$ for the ferromagnet remains unproved. Tóth [97] and Aizenman-Nachtergale [1] showed that the spin $S = \frac{1}{2}$ Heisenberg ferro- and anti-ferromagnet (respectively) have probabilistic representations as weighted interchange processes. Other spin systems have been studied with probabilistic representations, and interchange processes have been studied widely in their own right; see, for example, [5], [48], [61], [91].

The free energy of the spin $S = \frac{1}{2}$ Heisenberg ferromagnet on the complete graph was determined by Tóth [96] and Penrose [82]. This was extended by Björnberg [13] to a class of spin $S \in \frac{1}{2}\mathbb{N}$ models, with Hamiltonian equal to the sum of transposition operators. The model's probabilistic representation is that of the interchange process, where Tóth's weighting of $2^{\#\text{cycles}}$ is replaced by $(2S+1)^{\#\text{cycles}}$.

Motivated by [13], we give in Theorem 5.2.1 the free energy, on the complete graph and in spins $S = \frac{1}{2}$ and 1, of a model with Hamiltonian (5.20) given by linear combinations of the sum of transposition operators, and the sum of certain projection operators. For spins $S > 1$ we can apply a similar strategy to give in Theorem 5.2.2 the free energy in the case that one of the parameters of the Hamiltonian is non-negative. In spin $S = \frac{1}{2}$ the model is equivalent to the Heisenberg XXZ model (Hamiltonian (5.25)). In spin $S = 1$ it is equivalent to the bilinear-biquadratic Heisenberg model (Hamiltonian (5.29)), which is also known as the most general $SU(2)$-invariant spin $S = 1$ model (here $SU(2)$-invariance means invariance under the action of $SU(2)$ generated by the spin-operators). We give a full phase diagram in the two parameters of the Hamiltonian in the $S = \frac{1}{2}$ and 1 cases, and half of the diagram for $S > 1$ (the region where we have the free energy), giving the points of phase transitions in finite temperature, and ground state behaviour. These phase diagrams differ notably in shape from those on $\mathbb{Z}^d$, since the complete graph is not bipartite. Indeed, no phase transition is observed for the spin $S = \frac{1}{2}$ Heisenberg antiferromagnet, in contrast to $\mathbb{Z}^d$, [33], and the expected phase diagram for the spin $S = 1$ model in Z^d differs from ours on the complete graph - see Ueltschi's work [101] and [102]. In spins $S = \frac{1}{2}$ and 1 we give in Theorems 5.2.3 and 5.2.4 respectively expressions for a magnetisation and a total spin observable. These are motivated by corresponding results of [13] and [14] respectively.

The Hamiltonian in [13] is $GL(\theta)$-invariant, which allows it to be studied using the representation theory of the symmetric group (here and for the rest of the paper, by G-invariance, we mean that G acts on tensor space by $G \ni g \mapsto g^{\otimes n}$, and the Hamiltonian in question commutes with this action). Björnberg's key technical step is to express the partition function of the model in terms of the irreducible characters of the symmetric group. Our Hamiltonian is only $O(\theta)$-invariant, which requires us to look for more tools, as the symmetric group is not sufficient. (In fact, any $O(\theta)$-invariant pair-interaction Hamiltonian must be of the form (5.20)). The key representation-theoretic step in finding the free energy is to express the partition function in terms of the irreducible characters of both the symmetric group and the Brauer algebra. Indeed, the Brauer algebra was introduced by Brauer [19], as the algebra of invariants of the action of the orthogonal group on tensor space. A key technical step in our proofs is solving the problem of finding when the Brauer algebra - symmetric group branching coefficients are non-zero; we have a general solution for this problem in spins $S = \frac{1}{2}, 1$ in Propositions 5.7.6 and 5.7.8. For higher spins more work is needed to answer the problem fully, and handle the remaining parts of the phase diagram.

This paper is a continuation of several papers which analyse quantum spin systems and their interchange processes using representation theory (including [13]). Alon and Kozma [3] estimate the number of cycles of length k in the unweighted interchange process, on any graph. Berestycki and Kozma [9] give an exact formula for the same on the complete graph, and study the phase transitions present. In [4] Alon and Kozma give a formula for the magnetisation of the $2^{\#cycles}$ weighted process (equivalent to the spin $S = \frac{1}{2}$ ferromagnet) on any graph, which simplifies greatly in the mean-field.

The model we study was introduced by Ueltschi [101], generalising Tóth [97] and Aizenmann-Nachtergaele [1]. Ueltschi showed, for certain values of the parameters, equivalence with a weighted interchange process with "reversals". For these parameters, the model and interchange process have been studied on $\mathbb{Z}^d$ [16], [25], trees [10], [17], [49], graphs of bounded degree [76], and the complete graph [15], [14], the latter of which computes many observables. Our methods allow us to deal with all values of the parameters, not just those for which the probabilistic representation holds. The implications of our results for this interchange process seem to be limited to the following. In [15], the authors show that the transition time is independent of the parameter giving the ratio of "crosses" and "reversals"; our results indicate the same is most probably true for the weighted process.

In Section 5.2.1, we describe our model and precisely state our results. In Section 5.3 we give an introduction to the Brauer algebra. In Section 5.4 we prove our main result, Theorem 5.2.1, modulo the key ingredients Propositions 5.7.6 and 5.7.8 which are proved in Section 5.7. In Section 5.5 we give the free energy in higher spins, and prove our magnetisation and total spin results. In Section 5.6 we prove certain results on the analyticity of the free energy, which follow from Theorem 5.2.1, and give calculations which back up our interpretation of the phase diagrams.

5.2.1 Models and results

Let $S^{(1)}, S^{(2)}, S^{(3)}$ denote the usual spin-operators, satisfying the relations:

$$[S^{(1)}, S^{(2)}] = iS^{(3)}, \ [S^{(2)}, S^{(3)}] = iS^{(1)}, \ [S^{(3)}, S^{(1)}] = iS^{(2)},$$
$$(S^{(1)})^2 + (S^{(2)})^2 + (S^{(3)})^2 = S(S+1)\mathbf{id},$$

with $i = \sqrt{-1}$. For each $S \in \frac{1}{2}\mathbb{N}$ we use the standard spin S representation, with $S^{(j)}$, $j = 1, 2, 3$, Hermitian matrices acting on $V = \mathbb{C}^\theta$, $\theta = 2S + 1$. We will broadly adopt the bra-ket $\langle \cdot | \cdot \rangle$ statistical mechanical notation for vectors and operators on V and tensor products of V. We fix a non-degenerate, symmetric, bilinear form on V, such that an orthonormal basis of V with respect to this form is given by the eigenvectors $|a\rangle$ of $S^{(3)}$, with eigenvalues $a \in \{-S, \ldots, S\}$. Note we define the orthogonal group $O(\theta)$ as the group preserving this form, as in Chapter 2.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the complete graph on n vertices. We number the vertices $1, \ldots, n$.

At each vertex $1 \le i \le n$, we fix a copy $V_i = V$, and let the space $\mathbb{V} = \otimes_{1\le i\le n} V_i = V^{\otimes n}$. Now an orthonormal basis of $\mathbb{V}$ is given by vectors $|\mathbf{a}\rangle = \otimes_{1\le i\le n}|a_i\rangle$, where each $a_i \in \{-S, \dots, S\}$. If A is an operator acting on V, we define $A_i = A \otimes \mathbf{id}_{\mathcal{V}\setminus\{i\}}$ acting on $\mathbb{V}$ (i.e. $A_i(v_1 \otimes \dots \otimes v_i \otimes \dots \otimes v_n) = v_1 \otimes \dots \otimes Av_i \otimes \dots \otimes v_n$ for all $v_i \in V_i$, $1 \le i \le n$, extending linearly).

We define $T_{i,j}$ to be the transposition operator, and $Q_{i,j}$ to be a certain projection operator, first on $V_i \otimes V_j$:

$$\begin{aligned} T_{i,j}|a_i, a_j\rangle &= |a_j, a_i\rangle, \\ \langle a_i, a_j|Q_{i,j}|b_i, b_j\rangle &= \delta_{a_i,a_j}\delta_{b_i,b_j}, \end{aligned} \tag{5.19}$$

for basis vectors a_i, a_j, b_i, b_j of V_i, V_j as appropriate. We then identify $T_{i,j}$ with $T_{i,j} \otimes \mathbf{id}_{\mathcal{V}\setminus\{i,j\}}$, and $Q_{i,j}$ similarly. Note these are just (3.12). Let our Hamiltonian be defined as:

$$H = H(n, \theta, L_1, L_2) = -\sum_{i,j} (L_1 T_{i,j} + L_2 Q_{i,j}), \tag{5.20}$$

where $L_1, L_2 \in \mathbb{R}$, and the sum is over all pairs of vertices $1 \le i < j \le n$. We define the partition function as

$$Z_{n,\theta}(L_1, L_2) = \mathrm{Tr}\Big[e^{-\frac{1}{n}H(n,\theta,L_1,L_2)}\Big].$$

Note that usually we would write $e^{-\frac{\beta}{n}H}$, for inverse temperature β, but without loss of generality this β can be incorporated into L_1 and L_2. One could think of β as being expressed by the norm of the vector $(L_1, L_2) \in \mathbb{R}^2$. The factor $\frac{1}{n}$ compensates for the fact that on the complete graphs there are order n^2 interactions (as opposed to $\mathbb{Z}^d$, where the number of interactions is proportional to the volume).

We have the following results. The first and main result, Theorem 5.2.1, gives the free energy when the spin $S = \frac{1}{2}$ or 1, that is, $\theta = 2, 3$. Theorem 5.2.2 gives the free energy for all $\theta \ge 2$, but only for $L_2 \ge 0$; its proof is very similar to that of 5.2.1. Theorems 5.2.3 and 5.2.4 give formulae for a certain magnetisation and a certain total spin, respectively. In Theorems 5.2.5, 5.2.7 and 5.2.10 we analyse the free energies of Theorems 5.2.1 and 5.2.2 and discuss the phase diagrams that they produce.

Theorem 5.2.1. *For $\theta = 2, 3$, the free energy of the model with Hamiltonian given by (5.20) is:*

$$\lim_{n\to\infty} \frac{1}{n} \log Z_{n,\theta}(L_1, L_2) = \max_{(x,y)\in\Delta^*_\theta} \left[\frac{1}{2}\left((L_1 + L_2)\sum_{i=1}^{\theta} x_i^2 - L_2 y_1^2\right) - \sum_{i=1}^{\theta} x_i \log(x_i)\right], \tag{5.21}$$

where

$$\begin{aligned} \Delta^*_2 &= \{(x,y) = (x_1, x_2, y_1) \in [0,1]^3 \mid x_1 \ge x_2,\ x_1 + x_2 = 1, \quad 0 \le y_1 \le x_1 - x_2\}, \\ \Delta^*_3 &= \{(x,y) = (x_1, x_2, x_3, y_1) \in [0,1]^4 \mid x_1 \ge x_2 \ge x_3,\ x_1 + x_2 + x_3 = 1,\ 0 \le y_1 \le x_1 - x_3\}. \end{aligned} \tag{5.22}$$

From hereon in, we label the function being maximised by:

$$\phi = \phi_{\theta,L_1,L_2}(x,y) = \frac{1}{2}\left[(L_1 + L_2)\sum_{i=1}^{\theta} x_i^2 - L_2 y_1^2\right] - \sum_{i=1}^{\theta} x_i \log(x_i). \tag{5.23}$$

Our result for higher spins covers only the range $L_2 \geq 0$. As noted earlier in the introduction, this restriction is due to our only having a partial solution to determining when the Brauer algebra - symmetric group branching coefficients are non-zero, when the multiplicative parameter of the Brauer algebra $\theta = 2S+1$ is greater than 3 (see Section 5.3 for the definition of the multiplicative parameter, and (5.37) for the branching coefficients).

Theorem 5.2.2. *Let $\theta \geq 2$, and assume $L_2 \geq 0$. Then the free energy of the model with Hamiltonian (5.20) is:*

$$\lim_{n\to\infty} \frac{1}{n} \log Z_{n,\theta}(L_1, L_2) = \max_{x\in\Delta_\theta}\left[\frac{L_1+L_2}{2}\sum_{i=1}^{\theta} x_i^2 - \sum_{i=1}^{\theta} x_i \log(x_i)\right],$$

where $\Delta_\theta = \{x \in [0,1]^\theta \mid x_i \geq x_{i+1} \geq 0,\ \sum_{i=1}^{\theta} x_i = 1\}$.

Notice that for $\theta = 2, 3$ this theorem is consistent with Theorem 5.2.1, since in (5.21), for $L_2 \geq 0$, we must set $y_i = 0$ for all i. Note that when $L_2 = 0$, Theorems 5.2.1 and 5.2.2 recover Björnberg's result (Theorem 1.1 from [13]), with our L_2 equal to the β from that paper. Results equivalent to our following results are obtained (for all θ) for the case $L_2 = 0$ in [14]. Note in particular that there the symmetry is different - the model is $GL(\theta)$-invariant rather than the $O(\theta)$-invariance of our general $L_2 \neq 0$ model. The paper [14] also discusses the Gibbs states of the $L_2 = 0$ model, which are expected to be indexed by $\mathbb{CP}^{\theta-1}$, different from $L_2 \neq 0$ (see discussion below).

We can give two additional results, both for $\theta = 2, 3$. The first gives the free energy of the model when we add a certain magnetisation term with a real strength parameter h, and its left and right derivatives at $h = 0$. Let us modify the Hamiltonian (5.20):

$$H_h = H_h(n, \theta, L_1, L_2, W) = -\sum_{i,j}(L_1 T_{i,j} + L_2 Q_{i,j}) - h\sum_i W_i, \tag{5.24}$$

where h is real, and W is a $\theta \times \theta$ skew-symmetric matrix (i.e. $W^{\mathsf{T}} = -W$), with eigenvalues $1, -1$ for $\theta = 2$, and $1, 0, -1$ for $\theta = 3$. In this theorem and the next, the limitation of W being skew-symmetric is a technical one arising from the methodology. Note that $2S_i^{(2)}$ when $\theta = 2$, and $S_i^{(2)}$ when $\theta = 3$ is skew-symmetric with the appropriate eigenvalues. In our interpretation of phase diagrams, we will think of this magnetisation term as that in the $S^{(2)}$ direction. This theorem relates to Theorem 5.2.1 as Theorem 4.1 from [13] does to Theorem 1.1 from that paper.

Theorem 5.2.3. *Let $\theta = 2, 3$, and let $Z_{n,\theta}(L_1, L_2, h) = \mathrm{Tr}\left[e^{-\frac{1}{n}H_h}\right]$. The free energy of the model with Hamiltonian H_h (5.24) is given by:*

$$\Phi = \Phi_\theta(L_1, L_2, h) = \lim_{n\to\infty}\frac{1}{n}\log Z_{n,\theta}(L_1, L_2, h) = \max_{(x,y)\in\Delta_\theta^*}\left[\phi_{\theta,L_1,L_2}(x,y) + |h|y_1\right].$$

Further, the left and right derivatives of this free energy with respect to h, at $h = 0$, are given by:

$$\left.\frac{\partial\Phi}{\partial h}\right|_{h\searrow 0} = y_1^{\uparrow}, \qquad \left.\frac{\partial\Phi}{\partial h}\right|_{h\nearrow 0} = y_1^{\downarrow},$$

where $(x^\uparrow, y^\uparrow)$ is the maximiser of ϕ which maximises y_1, and $(x^\downarrow, y^\downarrow)$ the one which minimises y_1.

Note that $y_1^\uparrow$, $y_1^\downarrow$ depend on L_1, L_2, and we will show that both are zero when L_1, L_2 are small.

We now return to the model (5.20) with no magnetisation term. We give a total spin observable $\langle e^{\frac{h}{n}\sum_i W_i}\rangle$, for skew-symmetric matrices W. This theorem is an equivalent of Theorems 2.1-2.3 from [14], and its proof follows similar lines of reasoning to that of Theorem 2.3 from that paper, aided by the Brauer algebra technology that we develop in this paper.

Theorem 5.2.4. *Let $\theta = 2, 3$, $h \in \mathrm{R}$, and W skew-symmetric with eigenvalues $1, -1$ for $\theta = 2$, and $1, 0, -1$ for $\theta = 3$. Assume that the function ϕ_{θ, L_1, L_2} has a unique maximiser $(x^*, y^*) \in \Delta_\theta^*$. Then with H the Hamiltonian from (5.20), for $L_2 \neq 0$,*

$$\langle e^{\frac{h}{n}\sum_i W_i}\rangle := \lim_{n\to\infty} \frac{\mathrm{Tr}\left[e^{\frac{h}{n}\sum_i W_i} e^{-\frac{1}{n}H}\right]}{Z_{n,\theta}(L_1, L_2)} = \begin{cases} \cosh(hy_1^*), & \text{if } \theta = 2, \\ \frac{\sinh(hy_1^*)}{hy_1^*}, & \text{if } \theta = 3. \end{cases}$$

The quantity $\cosh(hy_1^*)$ is related to Ising spin-flip symmetry, see below in Proposition 5.2.6 and the discussion thereafter; and the quantity $\frac{\sinh(hy_1^*)}{hy_1^*}$ is related to $SU(2)$ (or $O(3)$) symmetry, see in Proposition 5.2.9 and the discussion thereafter, and in [14].

We now state our results in terms of two well known models, the spin $S = \frac{1}{2}$ Heisenberg XXZ model, and the spin $S = 1$ bilinear-biquadratic Heisenberg model.

5.2.2 Phase diagram for spin $S = \frac{1}{2}$

Let $S = \frac{1}{2}$, so $\theta = 2$. We consider the Hamiltonian of the XXZ model, which will be equivalent to (5.20). Let

$$H' = -\left(\sum_{i,j} K_1 S_i^{(1)} S_j^{(1)} + K_2 S_i^{(2)} S_j^{(2)} + K_1 S_i^{(3)} S_j^{(3)}\right), \tag{5.25}$$

with $K_1, K_2 \in \mathbb{R}$. Our result Theorem 5.2.1 leads us to the following theorem, which will give information about the phase diagram of this model. See Figures 5.1a and 5.1b.

Theorem 5.2.5. *The free energy of the model with Hamiltonian (5.25) is analytic everywhere in the (K_1, K_2) plane, except the half-lines $K_1 = 4, K_2 \leq 4$ and $K_2 = 4, K_1 \leq 4$, where it is differentiable, but not twice-differentiable, and the half-line $K_1 = K_2 \geq 4$, where it is not differentiable.*

Note that the free energy is trivially continuous everywhere (it is concave or convex, and in our case, the maximum of a smooth function).

Let us also formalise what we will prove about the magnetisation and finite volume ground states, which will aid our discussion of the phase diagram below.

Proposition 5.2.6. *Consider the spin $S = \frac{1}{2}$ Heisenberg XXZ model (5.25).*

1. *The magnetisation $y_1^\uparrow$ of Theorem 5.2.3 is positive if and only if $K_2 > 4$, $K_2 \geq K_1$, and is zero elsewhere.*

2. (a) *For $K_2 > 0$, $K_2 > K_1$, the finite volume ground states are spanned by the two product states $\otimes_{1\leq j\leq n}(|\frac{1}{2}\rangle \pm i|-\frac{1}{2}\rangle)$, (where $i = \sqrt{-1}$);*

(b) For $K_1 > 0$, $K_1 > K_2$, the finite (even) volume ground state is the vector (5.28).

Note that the two vectors $|\frac{1}{2}\rangle \pm i|-\frac{1}{2}\rangle$ are the eigenvectors of $S^{(2)}$.

Theorem 5.2.5 splits the plane into three regions of analyticity, which we identify as three phases of the model. We label the region $K_1 \le 4, K_2 \le 4$ disordered (illustrated in block pink in Figure 5.1b); the maximiser of the function ϕ in (5.21) is constant in this region, and it maximises the entropy term (the logarithms) of ϕ.

We label the region $K_2 > 4, K_2 > K_1$ the Ising phase (illustrated in dotted yellow in Figure 5.1b). It includes the half-line $K_1 = 0, K_2 \ge 4$, where the model is the supercritical classical Ising model, and further we will show the free energy in this region is independent of K_1 (it is perhaps slightly surprising that K_2 dominates to such a complete extent). There are two finite volume ground states in this region, the product states $\otimes_{1\le j\le n}(|\frac{1}{2}\rangle \pm i|-\frac{1}{2}\rangle)$. Further, for small values of h, adding $-h\sum_i S_i^{(2)}$ to the Hamiltonian as in (5.24) forces a unique ground state, $\otimes_{1\le j\le n}(|\frac{1}{2}\rangle + i|-\frac{1}{2}\rangle)$ when $h > 0$ and $\otimes_{1\le j\le n}(|\frac{1}{2}\rangle - i|-\frac{1}{2}\rangle)$ when $h < 0$. The magnetisation $y_1^{\uparrow}$ in the $S^{(2)}$ direction from Theorem 5.2.3 is positive.

The authors of [14] give a heuristic argument that points towards an expected structure of the set of extremal Gibbs states Ψ_β at inverse temperature β for several models on $\mathbb{Z}^d$, $d \ge 3$. The extremal Gibbs states in infinite volume are not well-defined on the complete graph, so the working is by analogy. Specifically, their heuristics indicate two expected equalities: first, that

$$\lim_{\Lambda_n \to \mathbb{Z}^d} \langle e^{\frac{h}{|\Lambda|}\sum_i S_i^{(2)}} \rangle_\Lambda = \int_{\Psi_\beta} e^{h\langle S_0^{(2)}\rangle_\psi} d\mu(\psi), \tag{5.26}$$

where $d\mu$ is the measure on Gibbs states corresponding to the symmetric state, $S_0^{(2)}$ is the spin operator at the lattice site 0, and the left hand side is the limit of successively larger boxes $\Lambda \in \mathbb{Z}^d$; and second that

$$\lim_{n\to\infty} \langle e^{\frac{h}{n}\sum_i S_i^{(2)}} \rangle_{\mathcal{G}} = \lim_{\Lambda_n \to \mathbb{Z}^d} \langle e^{\frac{h}{|\Lambda_n|}\sum_i S_i^{(2)}} \rangle_{\Lambda_n}, \tag{5.27}$$

where the left hand term is the observable on the complete graph. The left hand side of (5.27) is computed rigorously on the complete graph, and then, with the expected structure of Ψ_β inserted, the right hand side of (5.26) is rigorously computed, and the two are shown to be the same. This working is not a proof either of the expected equalities (5.26), (5.27) or of the expected structure of Ψ_β, but it points towards all three statements holding true. We can take the same approach for our models in several of the phases, with one small difference. We expect that the equality (5.27) holds for the complete graph models that we study in this paper on the left hand side, and on the right hand side models on other non-bipartite graphs, for example the triangular lattice, or the models on $\mathbb{Z}^d$, $d \ge 3$, with nearest and next-to-nearest neighbour interactions. For the Ising phase here, we can argue that the extremal Gibbs states in the Ising phase are expected to be indexed by $\{\pm e_2\}$, where e_2 is the second basis vector in $\mathbb{R}^3$. Indeed, with magnetisation $y_1^* = \langle S_0^{(2)}\rangle_{e_2}$ and $\Psi_\beta = \{\pm e_2\}$, the right hand side of this equality is $\cosh(hy_1^*)$; the left hand side is the same by Theorem 5.2.4.

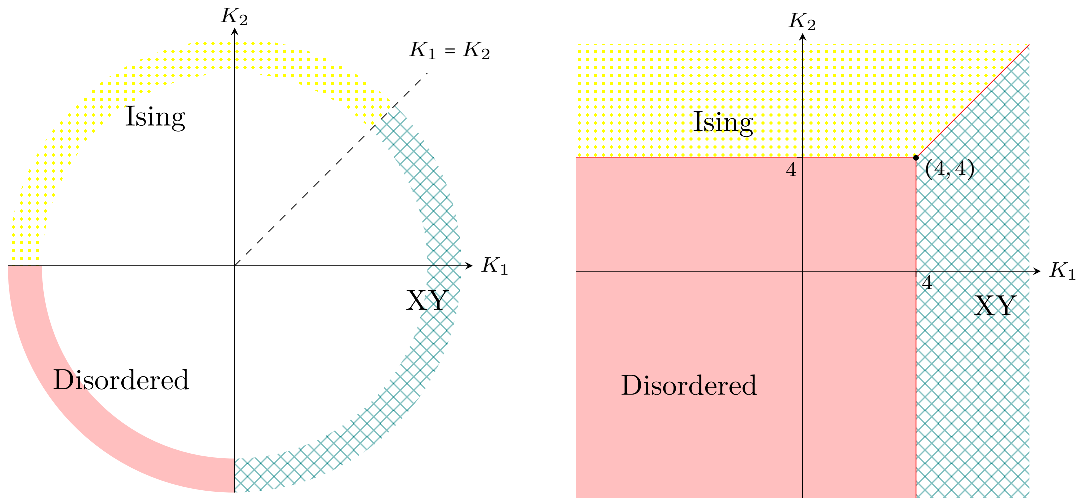
We label the region $K_1 > 4, K_1 > K_2$ the *XY* phase (illustrated in hatched blue in Figure 5.1b). We expect the $S_i^{(1)}$ and $S_i^{(3)}$ terms to dominate, and the extremal Gibbs states to be labelled by $\vec{a} \in \mathbb{S}^1$ in the $1-3$ directions. The magnetisation $y_1^{\uparrow}$ in the $S_i^{(2)}$ direction (from Theorem 5.2.3) is zero in this region, which is consistent with this picture. Similarly to the Ising phase, with $\Psi_\beta = \mathbb{S}^1$, the right hand side of (5.26) is 1, as is the left hand side by Theorem 5.2.4, so again we are encouraged in our labelling of the extremal states, and of (5.26). Equivalent calculations in the $S^{(1)}$ direction are done in [14]. Interestingly, the ground state in finite (even) volume is the vector

$$\sum_{\underline{m},\underline{m}'} \bigotimes_{i=1}^{n/2} \sum_{a=\frac{-1}{2},\frac{1}{2}} |a_{m_i}, a_{m_i'}\rangle, \tag{5.28}$$

where the sum is over all possible pairings $(\underline{m}, \underline{m}')$ of the vertices of $\mathcal{V}$ (that is, $(\underline{m}, \underline{m}') = ((m_1, \dots, m_{\frac{n}{2}}), (m_1', \dots, m_{\frac{n}{2}}'))$, with $\underline{m} \cup \underline{m}' = \mathcal{V}$, $\underline{m} \cap \underline{m}' = \varnothing$).

Note that the line $K_1 = K_2 \geq 0$ is the ferromagnetic Heisenberg model, and the extremal Gibbs states are expected to be labelled by $\vec{a} \in \mathbb{S}^2$. Here we can prove that the magnetisation $y_1^{\uparrow} > 0$ iff $K_1 = K_2 > 4$. The heuristics of (5.26) are given in Theorem 2.1 of [14].

The transitions from the disordered phase to each of the Ising and XY phases are second order, and the transition from Ising to XY is first order. The ground state phase diagram is illustrated in Figure 5.1a, and the finite temperature phase diagram is illustrated in Figure 5.1b.



(a) Ground state phase diagram

(b) Finite temperature phase diagram

Figure 5.1: On the left, the ground state phase diagram for the Spin $\frac{1}{2}$ Heisenberg XXZ model with Hamiltonian (5.25). The line $K_1 = K_2 \geq 0$ gives the Heisenberg ferromagnet. On the right, the phases at finite temperature, where varying temperature is given by varying the modulus $\|(K_1, K_2)\|$. Transitions between phases (points of non-analyticity of the free energy) shown in red lines.

5.2.3 Phase diagram for spin $S = 1$

In spin $S = 1$, we consider the bilinear-biquadratic Heisenberg model:

$$H'' = -\left(\sum_{i,j} J_1(\boldsymbol{S}_i \cdot \boldsymbol{S}_j) + J_2(\boldsymbol{S}_i \cdot \boldsymbol{S}_j)^2\right), \tag{5.29}$$

with $J_1, J_2 \in \mathbb{R}$. Our Theorem 5.2.1 leads us to the following theorem, which will give information about the phase diagram of this model. See Figures 5.2a and 5.2b. We can rigorously analyse the free energy in the $J_2 > J_1$ half of the phase diagram; on the other half we have partial results and numerical simulations to support Remark 5.2.8.

Theorem 5.2.7. *Within the region $J_2 > J_1$ of the (J_1, J_2) plane, the free energy of the model with Hamiltonian (5.29) is analytic everywhere, except the half-line $J_2 = \log 16, J_1 \leq \log 16$, where it is continuous, but not differentiable.*

Remark 5.2.8. We strongly suspect that Theorem 5.2.7 extends to the following: that the free energy of the model with Hamiltonian (5.29) is analytic everywhere in the (J_1, J_2) plane, apart from the half-lines $J_2 = \log 16, J_1 \leq \log 16$ and $J_1 = J_2 \geq \log 16$, where it is continuous, but not differentiable, and a curve (that we label $\mathcal{C}$) made up of the half-line $J_2 = 2J_1 - 3 \leq 3/2$ and a curve connecting the points $(\frac{9}{4}, \frac{3}{2})$ and $(\log 16, \log 16)$, which (as a function of J_1) is convex, with gradient in $[2, 3]$. It is unclear whether it is analytic on the half-line $J_1 = 0,\ J_2 \leq -3$.

Let us make clear what we will prove towards Remark 5.2.8 and the following discussion of the phase diagram of the model.

Proposition 5.2.9. *Consider the bilinear-biquadratic Heisenberg model with Hamiltonian (5.29).*

1. *The region $\mathcal{A}$ of the $J_1 - J_2$ plane where the point $(x, y) = ((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), (0,0,0))$ is a maximiser of ϕ_{3,J_1,J_2} (5.63) (the equivalent of ϕ_{3,L_1,L_2} (5.23) when we change variables appropriately) is closed and convex, and its boundary is the half-line $J_1 \leq J_2 = \log 16$, and a curve $\mathcal{C}$ as described in Remark 5.2.8.*

2. *The magnetisation term $y_1^{\uparrow}$ of Theorem 5.2.3 is zero in the region $\mathcal{A}$ and the region $J_2 \geq \log 16,\ J_2 > J_1$, and is positive in the region $J_1 \geq J_2$, strictly to the right of the curve $\mathcal{C}$.*

3. *(a) For $J_2 > 0,\ J_2 > J_1$, the finite (even) volume ground state is the vector (5.30);*

   (b) For $J_1 > 0,\ J_1 > J_2$, the finite volume ground states are those vectors invariant under S_n, and killed by all $P_{i,j}$ (5.61), which include the product states $\bigotimes_{1 \leq i \leq n} |a\rangle$, where $a_0^2 - a_1 a_{-1} = 0$;

   (c) For $\frac{1}{2} J_2 < J_1 < 0$, the finite volume ground states are those vectors are spanned by the vectors (5.31).

Let us now discuss the phase diagram. Theorem 5.2.7, Remark 5.2.8, and Proposition 5.2.9 divide the (J_1, J_2) plane into four regions, which we label as phases of the model. We label the region $\mathcal{A}$ (illustrated in block pink in Figure 5.2b) the disordered phase. The boundary of this region is made up of the half-line $J_2 = \log 16, J_1 \leq \log 16$ and the curve $\mathcal{C}$. The maximiser of the function ϕ from (5.21) is constant in this region, and maximises the entropy term.

We label the region of phase space to the right of the red line in Figure 5.2b, within the region $J_1 > J_2,\ J_1 > 0$, ferromagnetic (illustrated in dotted yellow in Figure 5.2b) (in fact, for large $\|(J_1, J_2)\|$, this region is that which is expected to be ferromagnetic in $\mathbb{Z}^3$, see [101]). The finite volume ground states include the product states $\otimes_{1\le i\le n}|a\rangle$, where $a_0^2 - a_1 a_{-1} = 0$, (eg. the ferromagnetic $|1\rangle$, $|-1\rangle$, as well as $|1\rangle + |0\rangle + |-1\rangle$). The magnetisation $y_1^{\uparrow}$ in the $S^{(2)}$ direction (Theorem 5.2.3) is positive in this phase. We expect that the extremal Gibbs states are indexed by $\vec{a} \in \mathbb{S}^2$, in which case (with $\langle S_0^{(2)}\rangle_{e_2} = y_1^*$) the right hand side of (5.26) equals $\frac{\sinh(hy_1^*)}{hy_1^*}$. Numerical simulations suggest that the maximiser y_1^* of ϕ is unique in this phase, so the left hand side of (5.26) should be the same, by Theorem 5.2.4. This encourages our expectation that the extremal states are indeed $\vec{a} \in \mathbb{S}^2$, and that (5.26) holds true. This extends the same analysis of the $J_2 = 0$ case given in Theorem 2.1 of [14].

We label the region of phase space $J_2 > \log 16, J_2 > J_1$ the nematic phase (illustrated in hatched blue in Figure 5.2b); we expect the $(\boldsymbol{S}_i \cdot \boldsymbol{S}_j)^2$ term to dominate, and the extremal Gibbs states to be given by $a \in \mathbb{RP}^2$. The magnetisation in the $S^{(2)}$ direction (Theorem 5.2.3), $y_1^{\uparrow}$, is zero in this phase, which matches the heuristics; we would expect to get something non-zero, for example, by replacing $S^{(2)}$ with its square. Interestingly, the finite (even) volume ground state in this phase is the vector

$$\sum_{\underline{m},\underline{m}'} \bigotimes_{i=1}^{n/2} \sum_{a=-1}^{1} |a_{m_i}, (-1)^a(-a)_{m_i'}\rangle, \tag{5.30}$$

where the sum is over all possible pairings $(\underline{m}, \underline{m}')$. This is the sum over all possible products of singlet states.

The fourth phase (illustrated in checkerboard orange in Figure 5.2b) occupies the region $\frac{1}{2}J_2 + \frac{3}{2} \le J_1 \le 0$. This phase is somewhat more mysterious. While the magnetisation $y_1^{\uparrow}$ is positive in this phase, the finite volume ground states are complicated. They depend on the ratio $\alpha = \frac{J_2}{J_1+J_2} \in [\frac{2}{3}, 1]$, and are spanned by vectors of the form

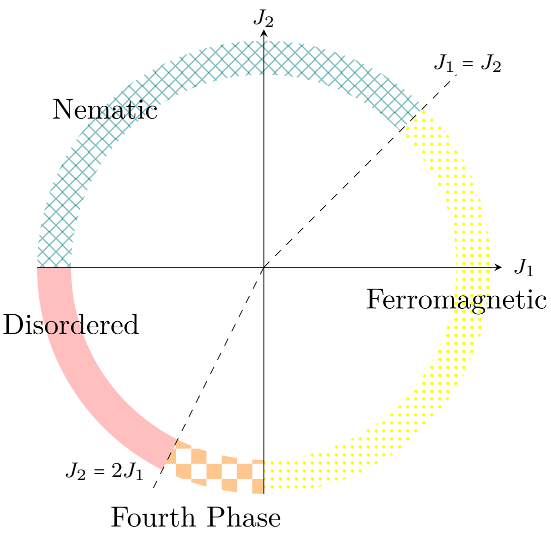
$$z_\tau \left[\left(\bigotimes_{i\in\mathcal{V}\setminus(\underline{m}\cup\underline{m}')} |a_i'\rangle \right) \otimes \left(\bigotimes_{i=1}^{(1-\alpha')n/2} \sum_{a=-1}^{1} |a_{m_i}, (-1)^a(-a)_{m_i'}\rangle \right) \right], \tag{5.31}$$

where α' is a fraction with denominator n close to α, $\underline{m}, \underline{m}'$ is a pairing of $(1-\alpha')n$ of the vertices, a' satisfies $(a_0')^2 - a_1' a_{-1}' = 0$, and z_τ is a Young symmetriser corresponding to the partition $(\alpha' n, (1-\alpha')n)$. The vector being symmetrised can be thought of as a proportion α of the volume being taken up by a ferromagnetic ground state, and $1-\alpha$ being taken up by a product of singlet states. However, it is difficult to interpret the vector once the Young symmetriser is applied.

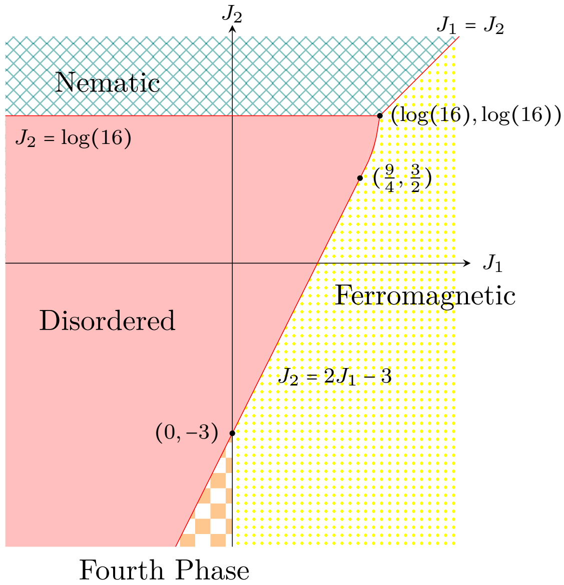
The transition from the disordered phase to the nematic phase is first order; we have not been able to prove similar statements for the other transitions. The ground state phase diagram is illustrated in Figure 5.2a, and the finite temperature phase diagram is illustrated in Figure 5.2b.

5.2.4 Higher spins

Recall that we only have the free energy of the model with Hamiltonian 5.20 for spins $S > 1$ in the region L_2. We can describe this half of the phase diagram in the (L_1, L_2)



(a) Ground state phase diagram



(b) Finite temperature phase diagram

Figure 5.2: On the left, the ground state phase diagram for the Spin 1 bilinear-biquadratic Heisenberg model with Hamiltonian (5.29). On the right, the phases at finite temperature, where varying temperature is given by varying the modulus $\|(J_1, J_2)\|$. Transitions between phases (points of non-analyticity of the free energy) shown in red lines (proved in the region $J_2 \geq J_1$, expected as shown for the rest of the plane).

plane for all spins as follows. Let

$$\beta_c = \beta_c(\theta) = \begin{cases} 2, & \text{for } \theta = 2 \\ 2\left(\frac{\theta-1}{\theta-2}\right)\log(\theta-1), & \text{for } \theta \geq 3. \end{cases} \tag{5.32}$$

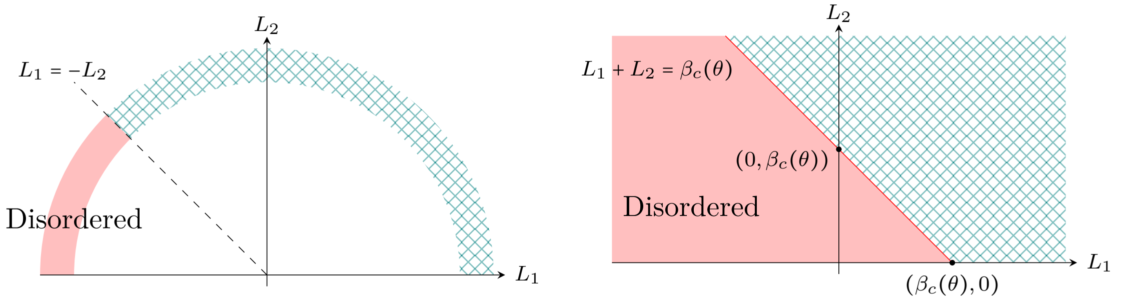
Theorem 5.2.10. *Let $S \geq \frac{1}{2}$ (so $\theta \geq 2$), and let $\beta_c = \beta_c(\theta)$ from (5.32). Within the region $L_2 > 0$ of the (L_1, L_2) plane, the free energy of the model with Hamiltonian (5.20) is analytic everywhere, except the half-line $L_1 + L_2 = \beta_c$, where for spin $S \geq 1$ it is continuous, but not differentiable, and for spin $S = \frac{1}{2}$ it is differentiable, but not twice-differentiable.*

We note that this theorem is a generalisation of Theorem 5.2.7 to spins $S \geq 1$, and indeed it implies Theorem 5.2.7. This can be seen by a unitary transformation of the spin $S = 1$ Hamiltonian (5.29) which we describe in Section 5.6.

The $L_2 \geq 0$ part of the phase diagram can be split into three phases. The disordered phase occupies the region $L_1 + L_2 < \beta_c$. The maximiser of ϕ_{θ,L_1,L_2} (5.23) maximises the entropy term (the logarithms) in this region. The region $L_1 + L_2 > \beta_c$, $L_2 > 0$ is a second phase. The finite (even) volume ground states include the vector

$$\sum_{\underline{m},\underline{m}'} \bigotimes_{i=1}^{n/2} \sum_{a=-S,S} |a_{m_i}, a_{m_i'}\rangle, \tag{5.33}$$

where the sum is over all possible pairings $(\underline{m}, \underline{m}')$ of the vertices of $\mathcal{V}$. The line $L_2 = 0$, $L_1 > 0$ is the quantum interchange model of [13]. The finite volume ground states are any vector which is invariant under the action of S_n.



(a) Ground state phase diagram

(b) Finite temperature phase diagram

Figure 5.3: On the left, the ground state phase diagram for the spin S model with Hamiltonian (5.20), in the region $L_2 \geq 0$. On the right, the phases for $L_2 \geq 0$ at finite temperature, where varying temperature is given by varying the modulus $\|(L_1, L_2)\|$. Transitions between phases (points of non-analyticity of the free energy) shown in red lines.

Remark 5.2.11 (A remark on the Interchange process with reversals)**.** As noted in the introduction, for certain values of the parameters, the model with Hamiltonian (5.20) has a probabilistic representation as an interchange process with reversals, re-weighed by $\theta^{\#\text{ loops}}$; see [101]. To be precise, let $L_1 = 1 - L_2 = u \in [0, 1)$, and introduce a temperature parameter β, that is, let $Z_{n,\theta}(u, \beta) := \mathrm{Tr}\big[e^{-\beta H(n,\theta,u,1-u)}\big]$. Then the corresponding interchange process is that described in Section 2A of [101], with β translating to time in the interchange process. It is natural to ask what our results imply, if anything, about this process; we have one remark to make on this topic. In [15], the authors consider the unweighted process with reversals, and prove that above a critical time β, the rescaled loop lengths converge to a Poisson-Dirichlet distribution, as n, the number of particles, tends to infinity. In particular, the critical time (and indeed the limiting distribution) are independent of the parameter u. Our result Theorem 5.2.10 indicates that a similar result might hold for the re-weighed process, since the transition in the spin model occurs at $\beta = \beta_c$ (5.32), independent of u.

For completeness, we make the following final remark. In [14], the authors obtain expressions for total spin observables of the form of Theorem 5.2.4, and note they are equal to certain observables of the corresponding interchange process, which are characteristic functions of the lengths of loops. Then they check that the limits of these observables, evaluated under the Poisson-Dirichlet distribution, are the same as the expressions obtained for the total spin observables. This supports the hypothesis that the rescaled loop lengths in the reweighed process are, in the limit, distributed is distributed according a Poisson Dirichlet distribution. It is tempting to try to play the same game here; however, we are unfortunately not able to with our specific total spin observables. Our total spin observable is trivial in the region where the probabilistic representation holds; we can give more details in the following.

In Theorem 2.3 of [14], the authors consider a total spin expression of the form of Theorem 5.2.4 for the case $u = 1$, the "Quantum Interchange Model", and for the matrix W replaced with any $\theta \times \theta$ matrix, with eigenvalues $h_1, \ldots, h_\theta$. That model has a probabilistic representation as the Interchange process (without reversals) with configurations re-weighed by $\theta^{\#\text{ loops}}$, described in Section 3.3 of that paper. A configuration of that process at time β is given by a configuration of certain loops; we label the lengths of the loops l_i, and the number of loops $l(\sigma)$. The total spin in finite volume is shown to be equal to the expectation of the observable $\prod_{i\geq 1}\frac{1}{\theta^{l(\sigma)}}(e^{h_1 l_i/n},\ldots,e^{h_\theta l_i/n})$. Now using Theorem 4.6 from [84], one can obtain the same expression for our total spin in Theorem 5.2.4, except the length of a loop, which before was the number of vertices at time $\beta = 0$ it visits, is replaced by the modulus of its winding number. The winding number definition comes from the algebraic equivalent in the Brauer algebra of the length of a cycle in the symmetric group, in that it defines conjugacy classes in the Brauer algebra (see Section 5.3 and Theorem 3.1 of [84]). (In the case of the interchange process without reversals, (and equivalently in the symmetric group) the length is the same as the modulus of the winding number, so there is no issue; this is not the case when reversals are introduced). Hence this observable does not tend to a function of the rescaled loop lengths, so cannot be compared with the Poisson-Dirichlet distribution.

5.3 The Brauer algebra

We essentially prove Theorem 5.2.1 by identifying the eigenspaces of the Hamiltonian, their dimensions, and their corresponding eigenvalues. We first observe that the Hamiltonian is actually the action of an element of the Brauer algebra on $\mathbb{V}$. Schur-Weyl duality gives us information on the irreducible invariant subspaces of this action, which leads us to the eigenspaces of the Hamiltonian.

In this subsection we will recall the definitions from Chapter 2 that will be of specific use to this paper. We will recall the definition of the Brauer algebra, and how its irreducible representations are enumerated, along with those of the symmetric group and the general linear and orthogonal groups.

Let $\theta \in \mathbb{C}$. The Brauer algebra $\mathbb{B}_{n,\theta}$ is the (formal) complex span of the set of pairings of $2n$ vertices. We think of pairings as graphs, which we will call diagrams, with each vertex having degree exactly 1. We arrange the vertices in two horizontal rows, labelling the upper row (the northern vertices) $1^+, 2^+, \ldots, n^+$, and the lower (southern) $1^-, \ldots, n^-$. We call an edge connecting two northern vertices (or two southern) a bar.

Multiplication of two diagrams is given by concatenation. If b_1, b_2 are two diagrams, we align the northern vertices of b_1 with the southern of b_2, and the result is obtained by removing these middle vertices (which produces a new diagram), and then multiplying the result by $\theta^{l(b_1,b_2)}$, where $l(b_1,b_2)$ is the number of loops in the concatenation. See Figure 5.4. This defines $\mathbb{B}_{n,\theta}$ as an algebra.

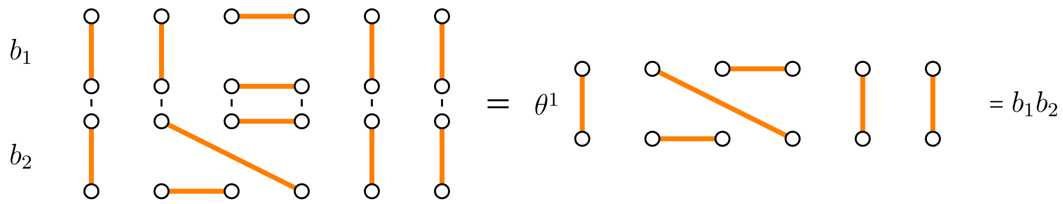


Figure 5.4: Two diagrams b_1 and b_2 (left), and their product (right). The concatenation contains one loop, so we multiply the concatenation (with middle vertices removed) by θ^1.

We call the set of diagrams B_n. Note that diagrams with no bars are exactly permutations, where $\sigma \in S_n$ is represented by the diagram where x^- is connected to $\sigma(x)^+$, so

$S_n \subset B_n$. Moreover the multiplication defined above reduces to multiplication in S_n, so $\mathbb{C}S_n$ is a subalgebra of $\mathbb{B}_{n,\theta}$. We write id for the identity - its diagram has all its edges vertical. We denote the transposition S_n swapping x and y by (i,j), and we write $(\overline{i,j})$ for the diagram with x^+ connected to y^+, and x^- connected to y^-, and all other edges vertical. See Figure 5.5. Note that just as the transpositions (i,j) generate the symmetric group, the Brauer algebra is generated by the transpositions and the elements $(\overline{i,j})$.

Let us note that the diagrams and multiplication depicted in Figures 5.4 and 5.5 mirror the paths in the interchange process with reversals (see [1], [101] for definitions). In a similar way to the interchange process without reversals being thought of as a continuous time random walk on the symmetric group, this shows that the process with reversals can be thought of as a random walk on the basis B_n of the Brauer algebra. See, for example, Figure 1 from [101].

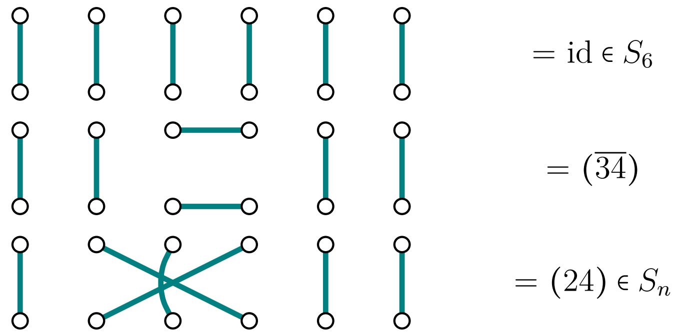


Figure 5.5: The identity element, the element $(\overline{34})$, and the transposition $(24) \in S_6$, all lying in B_6.

Let us turn to representations. A vector $\rho = (\rho_1, \ldots, \rho_t) \in \mathbb{Z}^t$ is a partition of n (we write $\rho \vdash n$) if $\rho_i \geq \rho_{i+1} \geq 0$ for all i, and $\sum_{i=1}^t \rho_i = n$. Recall that the irreducible representations (and characters) of $\mathbb{C}S_n$ are indexed by partitions of n. The Young diagram of $\rho \vdash n$ is the diagram of boxes of ρ with ρ_j boxes in the j^{th} row. When it is unambiguous, will denote the Young diagram of ρ simply by ρ. See Figure 5.6 for an illustration of the Young diagrams of the partitions $(5,5,3,1)$, $(4,1,1)$ respectively. We label by $\mathrm{ct}(\rho)$ the sum of contents of the boxes of the Young diagram of ρ, where the content of a box in row i and column j is given by $j - i$. For a partition ρ, $\rho^\intercal$ is the partition with Young diagram obtained by transposing the diagram of ρ (so $\rho_i^\intercal$ is the length of the i^{th} column of ρ). For $\rho \vdash n$ a partition, let $\psi_\rho^{S_n}$ be the irreducible representation corresponding to ρ, $\chi_\rho^{S_n}$ its character, and $d_\rho^{S_n}$ its dimension. The irreducible representations of the Brauer algebra

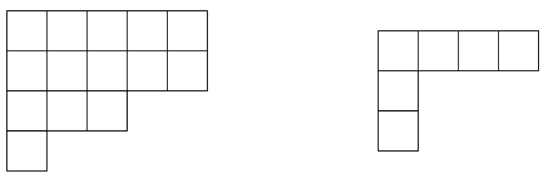


Figure 5.6: The Young diagrams of the partitions $(5,5,3,1)$ and $(4,1,1)$.

$\mathbb{B}_{n,\theta}$ are indexed by partitions $\lambda \vdash n - 2k$, $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$ (see Chapter 2 or, for example, [84] or [24]); let us denote them by $\psi_\lambda^{\mathbb{B}_{n,\theta}}$, and their characters by $\chi_\lambda^{\mathbb{B}_{n,\theta}}$, and dimensions by $d_\lambda^{\mathbb{B}_{n,\theta}}$.

We also recall that irreducible representations of the orthogonal group $O(\theta)$ are given by partitions λ of any size such that $\lambda_1^\intercal + \lambda_2^\intercal \le \theta$ (see Theorem 1.2 from [84]). The irreducible polynomial representations of $GL(\theta)$ are given by partitions ρ of any size with at most θ parts. Lastly, we note that the irreducible representations of the special orthogonal group $SO(\theta)$ are indexed by partitions of any size with at most $r = \lfloor \theta/2 \rfloor$ parts, with the exception that when $\theta = 2r$ even, the r^{th} part can be negative. For each of these three groups G, for a given partition π we denote the irreducible corresponding to π, and its character and dimension, by ψ_π^G, χ_π^G and d_π^G respectively.

In the following section we prove Theorem 5.2.1, and in the section after, Theorems 5.2.2, 5.2.3, 5.2.4, whose proofs are all based on that of 5.2.1. In Section 5.6 we prove Theorems 5.2.5 and 5.2.7, which follow from Theorem 5.2.1. In Section 5.7 we prove the two Propositions 5.7.6 and 5.7.8 which are key technical ingredients for the proof of Theorem 5.2.1.

5.4 Proof of Theorem 5.2.1

In this section we prove our main Theorem 5.2.1, modulo Propositions 5.7.6 and 5.7.8, whose proofs are postponed to Section 5.7. As noted above, our method is to identify the eigenspaces of our Hamiltonian, their dimensions and associated eigenvalues. We start by viewing the Hamiltonian (5.20) as the action of an element of the Brauer algebra $\mathbb{B}_{n,\theta}$ on $\mathbb{V}$.

Let $\theta \ge 2$. The Brauer algebra acts on $\mathbb{V} = (\mathbb{C}^\theta)^{\otimes \mathcal{V}} = V^{\otimes n}$ by $\mathfrak{p}^{\mathbb{B}_{n,\theta}}(\overline{i,j}) = Q_{i,j}$, and $\mathfrak{p}^{\mathbb{B}_{n,\theta}}(i,j) = T_{i,j}$, where recall $T_{i,j}$, $Q_{i,j}$ are given by (5.19). We therefore have $H = \mathfrak{p}^{\mathbb{B}_{n,\theta}}(\overline{H})$, where

$$\begin{aligned}\overline{H} &= -\sum_{i,j}\left(L_1(i,j) + L_2(\overline{i,j})\right)\\ &= -(L_1 + L_2)\sum_{i,j}(i,j) + L_2\sum_{i,j}\left((i,j) - (\overline{i,j})\right).\end{aligned}$$

Now $\overline{H}$ is a linear combination of two elements in $\mathbb{B}_{n,\theta}$: the sum of all transpositions, which is central in $\mathbb{C}S_n$, and the sum of all transpositions minus all elements $(\overline{i,j})$, which is a central element in $\mathbb{B}_{n,\theta}$. A central element of an algebra acts as a scalar on the irreducible representations of that algebra. Indeed, (from Lemma 2.1.10) for all $\rho \vdash n$,

$$\psi_\rho^{S_n}\left(\sum_{i,j}(i,j)\right) = \mathrm{ct}(\rho)\mathbf{id}, \tag{5.34}$$

and (see (2.10)) for all $\lambda \vdash n-2k$, $0 \le k \le \lfloor n/2 \rfloor$,

$$\psi_\lambda^{\mathbb{B}_{n,\theta}}\left(\sum_{i,j}((i,j) - (\overline{i,j}))\right) = (\mathrm{ct}(\lambda) + k(1-\theta))\mathbf{id}. \tag{5.35}$$

Finding the eigenspaces of the Hamiltonian requires two steps. First we find the irreducible invariant subspaces $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ of the action $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$, on each of which the element in (5.35) acts as a scalar. The element in (5.34) does not act as a scalar on these spaces $\psi_\lambda^{\mathbb{B}_{n,\theta}}$. Hence, the second step will be to further decompose these subspaces into smaller spaces (irreducibles $\psi_\rho^{S_n}$), on each of which the element in (5.34) does act as a scalar. These smaller spaces are therefore the eigenspaces of the Hamiltonian.

The first step, the decomposition of $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$, is given by a classical theorem called Schur-Weyl duality, which we now describe. The orthogonal group also has a natural action on $\mathbb{V}$; for $g \in O(\theta)$, $v_i \in \mathbb{C}_i^\theta$ for each $1 \le i \le n$, we have $g(v_1 \otimes \cdots \otimes v_n) = gv_1 \otimes \cdots \otimes gv_n$. Recall from Theorem 3.0.3 that Schur-Weyl duality states that the actions of the two algebras $\mathbb{B}_{n,\theta}$ and $\mathbb{C}O(\theta)$ on $\mathbb{V}$ centralise each other, and $\mathbb{V}$ can be viewed as a module of the tensor product $\mathbb{B}_{n,\theta} \otimes \mathbb{C}O(\theta)$, which decomposes as:

$$\mathbb{V} = \bigoplus_{\substack{\lambda \vdash n-2k \\ \lambda_1^\intercal + \lambda_2^\intercal \le \theta}} \psi_\lambda^{O(\theta)} \boxtimes \psi_\lambda^{\mathbb{B}_{n,\theta}}. \tag{5.36}$$

(Note the square tensor symbol denotes a representation of the tensor product of two algebras, as opposed to the circle tensor which denotes a representation of a single group or algebra). Hence the action $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ decomposes into irreducibles $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ (such that $\lambda_1^\intercal + \lambda_2^\intercal \le \theta$), each with multiplicity $d_\lambda^{O(\theta)}$. Note that a similar theorem (the original version of Schur-Weyl duality) holds for the general linear and symmetric groups (see equation (5.71) in Section 5.7). Here we only note that those representations of S_n which appear in $\mathbb{V}$ are all those with at most θ parts.

For the second step, we need to restrict $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ to the symmetric group and decompose into irreducibles. For $\rho \vdash n$ and $\lambda \vdash n-2k$, $0 \le k \le \lfloor n/2 \rfloor$, recall from (2.40)

$$\operatorname{res}_{S_n}^{\mathbb{B}_{n,\theta}} [\psi_\lambda^{\mathbb{B}_{n,\theta}}] = \bigoplus_{\rho \vdash n} (\psi_\rho^{S_n})^{\oplus b_{\lambda,\rho}^{n,\theta}}, \tag{5.37}$$

where res denotes the restriction of a representation. The coefficients $b_{\lambda,\rho}^{n,\theta}$ are the Brauer algebra - symmetric group branching coefficients. The eigenspaces of the Hamiltonian are therefore indexed by pairs (λ, ρ), each appearing with multiplicity $d_\lambda^{O(\theta)} b_{\lambda,\rho}^{n,\theta}$; their dimensions are $d_\rho^{S_n}$, and their corresponding eigenvalues are $-(L_1+L_2)\mathrm{ct}(\rho) + L_2[\mathrm{ct}(\lambda) + k(1-\theta)]$. Taking exponentials and traces, we see that

$$\begin{aligned} Z_{n,\theta}(L_1, L_2) &= \operatorname{Tr}\left[\mathfrak{p}^{\mathbb{B}_{n,\theta}}(e^{-\frac{1}{n}\overline{H}})\right] \\ &= \sum_{\substack{\lambda \vdash n-2k \\ \lambda_1^\intercal + \lambda_2^\intercal \le \theta}} \sum_{\rho \vdash n} d_\lambda^{O(\theta)} b_{\lambda,\rho}^{n,\theta} d_\rho^{S_n} \exp\left[\frac{1}{n}[(L_1+L_2)\mathrm{ct}(\rho) - L_2(\mathrm{ct}(\lambda) + k(1-\theta))]\right]. \end{aligned} \tag{5.38}$$

Now we need to take the limit of $\frac{1}{n}\log Z_{n,\theta}(L_1, L_2)$, which will essentially behave like $\frac{1}{n}\log$ of the largest term in the sum above. As $n \to \infty$, the behaviour of $\mathrm{ct}(\lambda)$, $\mathrm{ct}(\rho)$, and $\frac{1}{n}\log d_\rho^{S_n}$ are given by Björnberg [13]. We will show that $\frac{1}{n}\log d_\lambda^{O(\theta)} \to 0$. It remains to analyse the branching coefficients $b_{\lambda,\rho}^{n,\theta}$. In particular, since we are interested in the largest term in the sum above, and the sum is really only over those pairs (λ, ρ) for which $b_{\lambda,\rho}^{n,\theta} > 0$, it is crucial that we have good knowledge of when these coefficients are non-zero. Obtaining this knowledge is the main technical difficulty of this paper.

Let us introduce some notation. Define $\Lambda_n(\theta)$ to be the set of pairs (λ, ρ) of partitions with at most θ parts, $\lambda \vdash n-2k$ for some $0 \le k \le \lfloor \frac{n}{2} \rfloor$, with $\lambda_1^\top + \lambda_2^\top \le \theta$, $\rho \vdash n$. Let $P_n(\theta)$ be the set of $(\lambda, \rho) \in \Lambda_n(\theta)$ with the extra condition that $b_{\lambda,\rho}^{n,\theta} > 0$. Let $\frac{1}{n}P_n(\theta)$ be the set of pairs $(\frac{\lambda}{n}, \frac{\rho}{n})$, for $(\lambda,\rho) \in P_n(\theta)$.

For $\theta = 2,3$, we give a detailed description of $P_n(\theta)$ in Propositions 5.7.6 and 5.7.8, proved in Section 5.7. Essentially, (i.e. apart from a few edge cases which behave well), $(\lambda,\rho) \in \Lambda_n(2)$ lies in $P_n(2)$ iff $0 \le \lambda_1 \le \rho_1 - \rho_2$, and $(\lambda,\rho) \in \Lambda_n(3)$ lies in $P_n(3)$ iff $0 \le \lambda_1 \le \rho_1 - \rho_3$. As noted earlier, we do not know as much detail when $\theta > 3$ - we use what we do know to prove Theorem 5.2.3 in Section 5.5.

We will need to take the limit of the sequence $\frac{1}{n}P_n(\theta)$; let us make clear what we mean by this. Let $\Delta_\theta \subset \mathbb{R}^{2\theta}$ be the set of pairs $(x,y) \in ([0,1]^\theta)^2$ such that $\sum_{i=1}^\theta x_i = 1, x_i \ge x_{i+1}$ for all i, $\sum_{i=1}^\theta y_i \in [0,1], y_i \ge y_{i+1}$ for all i, and $y_i = 0$ for all $i > \lfloor \frac{\theta}{2} \rfloor$. Equip $\mathbb{R}^{2\theta}$ and subsets thereof with $||\cdot||$ the ∞-norm, $||z|| = \max_{i=1}^{2\theta}|z_i|$, and consider the Hausdorff distance $d_H(\cdot,\cdot)$ on sets in $\mathbb{R}^{2\theta}$:

$$d_H(U,W) = \inf\{\epsilon > 0 \mid U \subseteq W^\epsilon \text{ and } W \subseteq U^\epsilon\},$$

where $U^\epsilon = \{x \in \mathbb{R}^{2\theta} \mid ||x-u|| < \epsilon \text{ for some } u \in U\}$. Then Propositions 5.7.6 and 5.7.8 show that $\frac{1}{n}P_n(\theta) \to \Delta_\theta^*$ for $\theta = 2,3$ in this distance, where, recall,

$$\begin{aligned}
\Delta_2^* &= \{(x,y) \in ([0,1]^2)^2 \mid x_1 \ge x_2,\ x_1 + x_2 = 1,\ y_2 = 0,\ \ 0 \le y_1 \le x_1 - x_2\},\\
\Delta_3^* &= \{(x,y) \in ([0,1]^3)^2 \mid x_1 \ge x_2 \ge x_3,\ x_1 + x_2 + x_3 = 1,\ y_2, y_3 = 0,\ 0 \le y_1 \le x_1 - x_3\}.
\end{aligned} \tag{5.39}$$

The rest of the proof follows very similarly to Section 3 of Björnberg [13]. As in that paper, we prove a slightly more general convergence result, and then apply it to our setting.

Let Δ be any compact subset of $\mathbb{R}^t$, $t \in \mathbb{N}_{>0}$, and let $P_n \subset \Delta$ be a sequence of finite sets with $P_n \to \Delta$ in the Hausdorff distance, and $\frac{1}{n}\log|P_n| \to 0$, as $n \to \infty$. Let $\phi : \Delta \to \mathbb{R}$ continuous, and let $\phi_n : P_n \to \mathbb{R}$ such that $\phi_n \to \phi$ in the sense that there exists $\delta_n \to 0$ such that

$$|\phi_n(p_n) - \phi(p_n)| \le \delta_n, \tag{5.40}$$

uniformly in $p_n \in P_n$.

Lemma 5.4.1. *Given the assumptions above, we have that*

$$\lim_{n\to\infty} \frac{1}{n} \log\left(\sum_{p_n \in P_n} \exp\left[n\phi_n(p_n)\right] \right) = \max_{x\in\Delta}\phi(x).$$

Proof. Let us first prove an upper bound. We have that

$$\begin{aligned}
\frac{1}{n}\log\left(\sum_{p_n\in P_n}\exp[n\phi_n(p_n)]\right) &\le \frac{1}{n}\log\left(|P_n|\max_{p_n\in P_n}\{\exp[n\phi_n(p_n)\}]\right)\\
&= \max_{p_n\in P_n}[\phi_n(p_n)] + o(1)\\
&\le \max_{p_n\in P_n}[\phi(p_n)] + \delta_n + o(1)\\
&\le \max_{x\in\Delta}[\phi(x)] + \delta_n + o(1),
\end{aligned}$$

where in the second to last inequality we use that ϕ_n tends to ϕ (5.40), and in the last we use simply that $P_n \subset \Delta$. Hence we have $\limsup_{n\to\infty} \frac{1}{n}\log\left(\sum_{p_n\in P_n}\exp\left[n\phi_n(p_n)\right]\right) \le$

$\max_{x \in \Delta} \phi(x)$. For the lower bound, we have:

$$\frac{1}{n} \log \left(\sum_{(p_n) \in P_n} \exp[n\phi_n(p_n)] \right) \geq \frac{1}{n} \log \left(\max_{p_n \in P_n} \{\exp[n\phi_n(p_n)]\} \right)$$
$$= \max_{p_n \in P_n} [\phi_n(p_n)]$$
$$\geq \max_{p_n \in P_n} [\phi(p_n)] + \delta_n.$$

Now it suffices to prove that $\lim_{n\to\infty} \max_{p_n \in P_n} [\phi(p_n)] = \max_{x\in\Delta}[\phi(x)]$, which follows from convergence in the Hausdorff distance. Indeed, since Δ is compact, the maximum $\max_{x\in\Delta} \phi(x)$ is attained, say, at x^*. Then there exists a sequence of points $p_n \in P_n$ with $p_n \to x^*$. Now $\phi(p_n) \leq \max_{p_n \in P_n}[\phi(p_n)] \leq \max_{x\in\Delta}[\phi(x)] = \phi(x^*)$, again in the last inequality using the fact that $P_n \subset \Delta$, which gives the desired limit by continuity of ϕ. To conclude, $\liminf_{n\to\infty} \frac{1}{n} \log \left(\sum_{p_n \in P_n} \exp[n\phi_n(p_n)] \right) \geq \max_{x\in\Delta} \phi(x)$, which completes the proof. ∎

Now we set $\Delta = \Delta_\theta^*$, $P_n = \frac{1}{n} P_n(\theta)$, $\phi = \phi_{\theta,L_1,L_2}$ defined below (5.22), and $\phi_n = \phi_{n,\theta,L_1,L_2}$, where

$$\phi_{n,\theta,L_1,L_2}(\lambda,\rho) = \frac{1}{n}\log\left(d_\lambda^{O(\theta)}\right) + \frac{1}{n}\log\left(b_{\lambda,\rho}^{n,\theta}\right) + \frac{1}{n}\log\left(d_\rho^{S_n}\right) + \frac{1}{n^2}\left((L_1+L_2)\mathrm{ct}(\rho) - L_2(\mathrm{ct}(\lambda) + k(1-\theta))\right).$$

Now using Lemma 5.4.1, in order to prove Theorem 5.2.1, we note that $\frac{1}{n}P_n(\theta) \to \Delta_\theta^*$ in the Hausdorff distance by Propositions 5.7.6 and 5.7.8, and it remains to prove $\phi_{n,\theta,L_1,L_2} \to \phi_{\theta,L_1,L_2}$ in the sense of (5.40). Noting that $\frac{1}{n^2}(k(1-\theta)) \to 0$, the final two of the four terms in ϕ_{n,θ,L_1,L_2} give the desired limit; this is proved in Theorem 3.5 of [13], the salient points of which are that as $\frac{\rho}{n} \to x$, $\frac{1}{n}\log\left(d_\rho^{S_n}\right) \to -\sum_{i=1}^{\theta} x_i \log(x_i)$, and $\mathrm{ct}(\rho) \to \frac{1}{2}\sum_{i=1}^{\theta} x_i^2$. So it remains to prove only that $\frac{1}{n}\log\left(d_\lambda^{O(\theta)}\right) + \frac{1}{n}\log\left(b_{\lambda,\rho}^{n,\theta}\right)$ tends to zero as $n \to \infty$, uniformly in (λ,ρ). The second of these terms tends to zero by Corollaries 5.7.7 and 5.7.9 in Section 5.7. To show the first tends to zero, we note that Weyl's formula gives the dimension $d_\lambda^{SO(\theta)}$ of the irreducible representation of $SO(\theta)$ corresponding to $\lambda = (\lambda_1, \ldots, \lambda_r)$, where $r = \lfloor \theta/2 \rfloor$ (see, for example, Section 7 of [44]). For θ odd, and $\pi_i = n - i + \frac{1}{2}$,

$$d_\lambda^{SO(\theta)} = \prod_{1\leq i<j\leq r} \frac{(\lambda_i+\pi_i)^2 - (\lambda_j+\pi_j)^2}{\pi_i^2 - \pi_j^2} \prod_{1\leq i\leq r} \frac{\lambda_i + \pi_i}{\pi_i},$$

and for $\pi_i = n - i$, θ even, we have

$$d_\lambda^{SO(\theta)} = \prod_{1\leq i<j\leq r} \frac{(\lambda_i+\pi_i)^2 - (\lambda_j+\pi_j)^2}{\pi_i^2 - \pi_j^2}.$$

It's straightforward to see that these dimensions are bounded above by $(2n)^{6r}$. Finally, for $\lambda_1^\intercal \leq r$ recall from (2.23) that

$$\mathrm{res}_{SO(\theta)}^{O(\theta)} \chi_\lambda^{O(\theta)} = \mathrm{res}_{SO(\theta)}^{O(\theta)} \chi_{\lambda'}^{O(\theta)} = \chi_\lambda^{SO(\theta)}, \tag{5.41}$$

where λ' is identical to λ, except its first column is replaced by $\theta - \lambda_1^\intercal$, except in the case when $\theta = 2r$ even, and $\lambda_r > 0$, in which case recall from (2.24) that

$$\mathrm{res}^{O(\theta)}_{SO(\theta)}\chi^{O(\theta)}_{\lambda} = \chi^{SO(\theta)}_{\lambda} + \chi^{SO(\theta)}_{\lambda^\circ}, \tag{5.42}$$

where λ° is the same as λ except with λ_r replaced with $-\lambda_r$. As a consequence, the dimensions $d^{O(\theta)}_{\lambda}$ are bounded above by $2(2n)^{6r}$. This completes the proof of Theorem 5.2.1.

5.5 Proofs of Theorems 5.2.2, 5.2.3 and 5.2.4

Proof of Theorem 5.2.2. As noted above, the main technical difficulty in this paper is finding a detailed description for $P_n(\theta)$. For general θ, all of the working from the proof of Theorem 5.2.1 in Section 5.4 holds, apart from the fact that we do not know what the set $P_n(\theta)$ looks like for $\theta > 3$. For $L_2 \geq 0$, it turns out that enough information in contained in a theorem of Okada [81], which computes the coefficients $b^{n,\theta}_{\lambda,\rho}$ in certain special cases. Note that in [81], the coefficients are described in terms of the general linear and orthogonal groups - in Lemma 5.7.1 we show that this formulation is equivalent to ours.

Remark 5.5.1. Okada's result says: if $\lambda = (1^j)$, $j = 0, \ldots, \theta$, then $b^{n,\theta}_{\lambda,\rho} = 1$ if ρ has exactly j odd parts, and is zero otherwise (part (2) of Theorem 5.4 of [81]).

Now assume $L_2 \geq 0$. Recall the decomposition of $Z_{n,\theta}(L_1, L_2)$ from (5.38):

$$Z_{n,\theta}(L_1, L_2) = \sum_{(\lambda,\rho)\in P_n(\theta)} d^{O(\theta)}_{\lambda} b^{n,\theta}_{\lambda,\rho} d^{S_n}_{\rho} \exp\left[\frac{1}{n}\left[(L_1 + L_2)\mathrm{ct}(\rho) - L_2(\mathrm{ct}(\lambda) + k(1-\theta))\right]\right]. \tag{5.43}$$

Since the sum behaves like its maximal term, and $L_2 \geq 0$, it is clear that we would like to minimise $\mathrm{ct}(\lambda)$. Remark 5.5.1 allows us to do this, since the partitions (1^j) have $\mathrm{ct}((1^j))$ essentially zero.

Let us make this precise. Given $\rho \vdash n$ with $\rho^{\mathsf{T}}_1 \leq \theta$, let $j(\rho)$ be the number of parts of ρ of odd length. Then by Remark 5.5.1, the pair $((1^{j(\rho)}), \rho)$ lies in $P_n(\theta)$. Now take any pair $(\lambda, \rho) \in P_n(\theta)$. It is straightforward to show that $\mathrm{ct}(\lambda) \geq \mathrm{ct}((1^{j(\rho)})) - \theta^3$. Indeed, $\mathrm{ct}((1^{j(\rho)})) \leq 0$, and since λ has at most θ parts, it has at most θ^2 boxes with negative content, and those contents must be at least $-\theta$. Substituting into (5.43) gives

$$Z_{n,\theta}(L_1, L_2) \leq \sum_{(\lambda,\rho)\in P_n(\theta)} d^{O(\theta)}_{\lambda} b^{n,\theta}_{\lambda,\rho} d^{S_n}_{\rho} \exp\left[\frac{1}{n}\left[(L_1 + L_2)\mathrm{ct}(\rho) - L_2(\mathrm{ct}((1^{j(\rho)})) - \theta^3 + k(1-\theta))\right]\right]. \tag{5.44}$$

The lower bound is trivial, simply take the term $((1^{j(\rho)}), \rho)$ from the sum to achieve

$$Z_{n,\theta}(L_1, L_2) \geq \sum_{\substack{\rho\vdash n,\\ \rho^{\mathsf{T}}_1 \leq \theta}} d^{O(\theta)}_{(1^{j(\rho)})} b^{n,\theta}_{(1^{j(\rho)}),\rho} d^{S_n}_{\rho} \exp\left[\frac{1}{n}\left[(L_1 + L_2)\mathrm{ct}(\rho) - L_2(\mathrm{ct}((1^{j(\rho)})) + k(1-\theta))\right]\right]. \tag{5.45}$$

Now we can apply Lemma 5.4.1 to see the result, recalling that $\frac{1}{n}\log|P_n(\theta)|$, $\frac{1}{n}\log d_\lambda^{O(\theta)}$, $\frac{1}{n}\log b_{\lambda,\rho}^{n,\theta}$ and $\frac{1}{n^2}(\mathrm{ct}((1^{j(\rho)})) - \theta^3 + k(1-\theta))$ all tend to zero as $n \to \infty$. ■

Proof of Theorem 5.2.3. Let $\theta = 2,3$, and let W be a skew-symmetric $\theta \times \theta$ matrix with eigenvalues 1, -1 for $\theta = 2$, and $1, 0, -1$ for $\theta = 3$. Consider the model with Hamiltonian H_h given in (5.24), and let $Z_{n,\theta}(L_1, L_2, h) = \mathrm{Tr}\left[e^{-\frac{1}{n}H}\right]$. The same working as in Section 5.4, taking traces in (5.36), gives us:

$$Z_{n,\theta}(L_1,L_2,h) = \sum_{(\lambda,\rho)\in P_n(\theta)} \chi_\lambda^{O(\theta)}(e^{hW}) b_{\lambda,\rho}^{n,\theta} d_\rho^{S_n} \exp\left[\frac{1}{n}\left[(L_1+L_2)\mathrm{ct}(\rho) - L_2(\mathrm{ct}(\lambda) + k(1-\theta))\right]\right]. \tag{5.46}$$

Now by Lemma 5.4.1, to prove the free energy part of the theorem, it suffices to prove that as $\lambda/n \to y$ (as $n \to \infty$), we have

$$\frac{1}{n}\log \chi_\lambda^{O(\theta)}(e^{hW}) \to |h|y_1. \tag{5.47}$$

We prove a more general lemma, one which holds for all θ.

Lemma 5.5.2. *Let $\theta \geq 2$, let $\lambda \vdash n-2k$, $0 \leq k \leq \lfloor n/2 \rfloor$ with $\lambda_1^\intercal + \lambda_2^\intercal \leq \theta$, and let $\chi_\lambda^{O(\theta)}$ denote the irreducible representation of $O(\theta)$ indexed by λ. Let W be any $\theta\times\theta$ skew-symmetric matrix with real eigenvalues $w_1 \geq \cdots \geq w_\theta$ (note $w_i = -w_{r+1-i}$ for each $i = 1,\dots,\theta$). Let $r = \lfloor \theta/2 \rfloor$ (note $w_1 \geq \cdots \geq w_r \geq 0$). Then as $n\to\infty$ and $\lambda/n \to y$, we have*

$$\frac{1}{n}\log \chi_\lambda^{O(\theta)}(e^{hW}) \to |h|\sum_{i=1}^r w_i y_i.$$

Proof. Notice that $e^{hW} \in SO(\theta)$. Assume $\lambda_1^\intercal \leq \theta/2$, and recall from (5.41) that in all cases except θ even and $\lambda_1^\intercal = \theta/2$, we have that $\mathrm{res}_{SO(\theta)}^{O(\theta)}\chi_\lambda^{O(\theta)} = \mathrm{res}_{SO(\theta)}^{O(\theta)}\chi_{\lambda'}^{O(\theta)} = \chi_\lambda^{SO(\theta)}$, where the latter is the irreducible representation of $SO(\theta)$ with highest weight λ, and where λ' is identical to λ, except its first column is replaced by $\theta - \lambda_1^\intercal$. In this case, a formula due to King (see Theorem 2.5 of [94]) gives

$$\chi_\lambda^{SO(\theta)}(e^{\beta h w_1},\dots,e^{\beta h w_r}) = \sum_{\mathbb{T}} 2^{m(\mathbb{T})} e^{\beta h \sum_{i=1}^r w_i(m_i - m_{\bar{i}})}, \tag{5.48}$$

where the sum is over semistandard Young tableaux of shape λ filled with indices $1 < \bar{1} < 2 < \bar{2} < \cdots < r < \bar{r} < \infty$, such that:

1. The entries of row i are all at least i,
2. If i and $\bar{i}$ appear consecutively in a row, then there is an i in the box directly above the $\bar{i}$.

Here $m(\mathbb{T})$ is the number of occurrences of i directly above $\bar{i}$ in the first column of the tableau $\mathbb{T}$, with $\bar{i}$ in row i, and m_i is the number of times i appears in the tableau, $m_{\bar{i}}$ similar. We recall also that a semistandard Young tableau of shape λ is a Young diagram of shape λ with each box filled with one of a set of indices, such that along rows the indices are non-decreasing, and down columns they strictly increase. Let $h > 0$. The exponent in (5.48) is maximised by the tableau with every box in row i containing i. Indeed, taking any tableau $\mathbb{T}$, changing a single box in row i to contain i changes the exponent by either $h(w_i - w_j)$ (if the box contains $j \geq i$), $h(w_i + w_j)$ (if the box contains $\bar{j} \geq i$) or hw_i (if the box contains ∞). These are all non-negative, since we ordered $w_1 \geq \cdots \geq w_r \geq 0$. In a very similar way, if $h < 0$, the exponent is maximised by the tableau with $\bar{i}$ as each entry of row i. In either case, the maximum exponent is $|h| \sum_{i=1}^{s} w_i \lambda_i$. Now we have

$$e^{|h| \sum_{i=1}^{s} w_i \lambda_i} \leq \chi_{\lambda}^{SO(\theta)}(e^{hw_1}, \ldots, e^{hw_s}) \leq e^{|h| \sum_{i=1}^{s} w_i \lambda_i} \sum_{\mathbb{T}} 2^{m(\mathbb{T})}, \tag{5.49}$$

and noticing that $2^{m(\mathbb{T})}$ is bounded, and the number of $\mathbb{T}$, which is the dimension of the irreducible representation, satisfies $\frac{1}{n} \log d_{\lambda}^{SO(\theta)} \to 0$, we have the key claim. ∎

This proves (5.47), and therefore proves the free energy part of the theorem. It remains to prove the second part. Again we prove a more general lemma. Let $r = \lfloor \frac{\theta}{2} \rfloor$, and let $\Delta_{\theta}^{\bullet}$ be the set of pairs $(x, y)^2 \in ([0,1]^{\theta})^2$, with $x_i \geq x_{i+1} \geq 0$, $\sum_{i=1}^{\theta} x_i = 1$, $y_i \geq y_{i+1} \geq 0$, $y_i = 0, \text{for} i > r$, $\sum_{i=1}^{\theta} y_i \in [0,1]$.

Lemma 5.5.3. *Let h, L_1, L_2 be real, and let $w_1 \geq \cdots \geq w_r \geq 0$, where $r = \lfloor \theta/2 \rfloor$. Define a function Φ as*

$$\Phi = \Phi(L_1, L_2, h) = \max_{(x,y) \in \Delta_{\theta}^{\dagger}}(x, y) \left[\phi_{\theta, L_1, L_2} + |h| \sum_{i=1}^{r} w_i y_i \right],$$

where $\Delta_{\theta}^{\dagger}$ is some compact subset of $\Delta_{\theta}^{\bullet}$. Then

$$\left. \frac{\partial \Phi}{\partial h} \right|_{h \downarrow 0} = \sum_{i=1}^{s} w_i y_i^{\uparrow}, \qquad \left. \frac{\partial \Phi}{\partial h} \right|_{h \uparrow 0} = \sum_{i=1}^{s} w_i y_i^{\downarrow}, \tag{5.50}$$

where $(x^{\uparrow}, y^{\uparrow})$ is the maximiser of ϕ in $\Delta_{\theta}^{\dagger}$ which maximises the inner product $\sum_{i=1}^{s} w_i y_i$, and $(x^{\downarrow}, y^{\downarrow})$ the one which minimises the inner product.

Proof. The proof follows the proof of Theorem 4.1 from [13] very closely. We prove the case of the right derivative - the left derivative is almost identical. Note that $(x^{\uparrow}, y^{\uparrow})$ (resp. $(x^{\downarrow}, y^{\downarrow})$) may not be unique, but this does not matter for the proof; from hereon in by $(x^{\uparrow}, y^{\uparrow})$ we mean one such maximiser. We have

$$\frac{\Phi(L_1, L_2, h) - \Phi(L_1, L_2, 0)}{h} = \max_{(x,y) \in \Delta_{\theta}^{\dagger}} \left[\sum_{i=1}^{s} y_i w_i + \frac{\phi(x, y) - \phi(x^{\uparrow}, y^{\uparrow})}{h} \right]. \tag{5.51}$$

We denote the function being maximised on the right hand side by $f(x, y; h)$. Clearly its maximum is bounded below by $\sum_{i=1}^{s} y_i^{\uparrow} w_i$. For fixed $h > 0$, let $(x(h), y(h))$ maximise $f(x, y; h)$ (such a maximiser exists as x, y lie in compact sets, and f is continuous). It suffices to show that as $h \searrow 0$, $(x(h), y(h)) \to (x^{\uparrow}, y^{\uparrow})$. Certainly $(x(h), y(h))$ must tend to a maximiser of $\phi(x, y)$; if it did not, then by continuity, $\phi(x, y) - \phi(x^{\uparrow}, y^{\uparrow})$ would stay bounded away from zero (below some negative number), and the right hand side of (5.51) would tend to $-\infty$. This contradicts the lower bound we noted above. To conclude, $(x(h), y(h))$ must tend to $(x^{\uparrow}, y^{\uparrow})$ (and not a different maximiser), since the sum $\sum_{i=1}^{r} = y_i w_i$ defining $y^{\uparrow}$ appears in $f(x, y; h)$. ∎

This concludes the proof of Theorem 5.2.3. ∎

Proof of Theorem 5.2.4. We have, taking traces in (5.36),

$$\begin{aligned}\langle e^{(h/n)W}\rangle_{n,\theta} &= \frac{\sum_\lambda \chi_\lambda^{O(\theta)}(e^{(h/n)W})\chi_\lambda^{\mathbb{B}_{n,\theta}}(e^{(h/n)W})}{\sum_\lambda d_\lambda^{O(\theta)}\chi_\lambda^{\mathbb{B}_{n,\theta}}(e^{(h/n)W})} \\ &= \frac{\sum_\lambda d_\lambda^{O(\theta)}\chi_\lambda^{\mathbb{B}_{n,\theta}}(e^{(h/n)W})\frac{\chi_\lambda^{O(\theta)}(e^{(h/n)W})}{d_\lambda^{O(\theta)}}}{\sum_\lambda d_\lambda^{O(\theta)}\chi_\lambda^{\mathbb{B}_{n,\theta}}(e^{(h/n)W})}.\end{aligned} \tag{5.52}$$

Using Lemma B.1 from [14], it suffices to show that for $\theta = 3$, $\chi_\lambda^{O(\theta)}(e^{(h/n)W})/d_\lambda^{O(\theta)} \to sinh(hy_1)/hy_1$ as $\lambda/n \to y$, while for $\theta = 2$, the limit is $cosh(hy_1)$. Let $\theta = 3$. Using the determinental formula for the character of the orthogonal group [84], and the Weyl dimension formula [44], we have

$$\frac{\chi_\lambda^{O(3)}(e^{(h/n)W})}{d_\lambda^{O(3)}} = \frac{e^{(h/n)(\lambda_1+1/2)} - e^{-(h/n)(\lambda_1+1/2)}}{e^{h/2n} - e^{-h/2n}} \cdot \frac{1/2}{\lambda_1 + 1/2},$$

which, on expanding the exponentials in the denominator, clearly tends to the desired limit. The $\theta = 2$ case is simpler. The dimension $d_\lambda^{O(2)} = 2$ for all λ except $\lambda = \varnothing$ or (1^2) (the trivial and determinant representations), which are both one-dimensional. In the latter two cases, $\chi_\lambda^{O(2)}(e^{(h/n)W})/d_\lambda^{O(2)} = 1$, since $e^{(h/n)W} \in SO(2)$, and in the former case, $\chi_\lambda^{O(2)}(e^{(h/n)W})/d_\lambda^{O(2)} = (e^{h\lambda_1/n} + e^{-h\lambda_1/n})/2$, which has the desired limit. ∎

5.6 Phase diagrams

In this section we prove Theorems 5.2.5, 5.2.7 and 5.2.10 and justify their descriptions of the phase diagrams for their respective systems. We begin with proving Theorem 5.2.10, and then show that Theorems 5.2.5 and 5.2.7 can be reduced to 5.2.10.

5.6.1 Higher spins; Proof of Theorem 5.2.10

Proof of Theorem 5.2.10. Recall the result Theorem 5.2.2, which gives the free energy of the model with Hamiltonian 5.20 in the region $L_2 \geq 0$:

$$\lim_{n\to\infty} \frac{1}{n} \log Z_{n,\theta}(L_1, L_2) = \max_{x\in\Delta_\theta}\left[\frac{L_1 + L_2}{2}\sum_{i=1}^{\theta} x_i^2 - \sum_{i=1}^{\theta} x_i \log(x_i)\right],$$

where $\Delta_\theta = \{x \in [0,1]^\theta \mid x_i \geq x_{i+1} \geq 0,\ \sum_{i=1}^\theta x_i = 1\}$. Let us label the function being maximised ϕ^{int}. This function ϕ^{int} is that from Theorem 1.1 of [13]. In that paper and Lemma C.1 of [14], it is proved that the maximisers of ϕ^{int} are always of the form $(x, \frac{1-x}{\theta-1}, \ldots, \frac{1-x}{\theta-1})$, that for $L_1 + L_2 \neq \beta_c$ the maximiser is unique, and that at $L_1 + L_2 = \beta_c$, there are exactly two maximisers, at $x = \frac{1}{\theta}$ and $x = 1 - \frac{1}{\theta}$ (which become a single unique maximiser when $\theta = 2$). Here β_c is given by (5.32). Moreover, it suffices to work with the modified function $\phi^{\text{int}}(x_1, \ldots, x_\theta) - \phi^{\text{int}}(\frac{1}{\theta}, \ldots, \frac{1}{\theta}) = \phi^{\text{int}}(x_1, \ldots, x_\theta) - \frac{L_1+L_2}{2\theta} - \log\theta$, since we are subtracting a smooth function of $L_1 + L_2$ independent of the variables x_i. Combining the above facts, we can consider a function of one variable; let $\phi:[0,1]\to\mathbb{R}$ as

$$\phi(x)=\phi_{\theta,\beta}(x)=\frac{\beta}{2}\left(x^2+(\theta-1)\left(\frac{1-x}{\theta-1}\right)-\frac{1}{\theta}\right)-x\log x-(1-x)\log\left(\frac{1-x}{\theta-1}\right)-\log\theta,$$

and let $\Phi(\beta)=\Phi_\theta(\beta)=\max_{x\in[\frac{1}{\theta},1]}\phi(x)$. To prove Theorem 5.2.10, it suffices to prove that $\Phi_\theta(\beta)$ is smooth for $\beta\neq\beta_c$, and is differentiable but not twice-differentiable at β_c for $\theta\geq 3$, and is continuous but not differentiable at β_c for $\theta=2$.

By [13], for all θ, $\Phi(\beta)=0$ for all $\beta\leq\beta_c$. Let us denote by $x^*=x^*(\beta)$ the unique maximiser of ϕ for all $\beta>\beta_c$. For $\beta>\beta_c$, we can obtain a formula for β in terms of the maximiser x^*, indeed, setting $\frac{\partial\phi}{\partial x}=0$ gives

$$\beta=\frac{\theta-1}{\theta x^*-1}\log\left(\frac{x^*(\theta-1)}{1-x^*}\right); \tag{5.53}$$

this function is smooth and increasing for $x^*\in(1-\frac{1}{\theta},1)$, tends to β_c as x^* tends to $1-\frac{1}{\theta}$ and to $+\infty$ as x^* tends to 1. By the inverse function theorem, x^* is a smooth function of β in the region (β_c,∞). Hence $\Phi_\theta(\beta)=\phi(x^*(\beta))$ is a smooth function on the interval (β_c,∞). We now turn to the behaviour at β_c.

Let $\theta=2$. Recall that $\beta_c=2$. To show Φ is differentiable at $\beta=\beta_c$, we need to show that its right derivative exists, and equals 0. By expanding the Taylor series of the logarithms, calculations yield that for $x\in(0,1)$,

$$\phi(x)=\frac{\beta}{4}(2x-1)^2+\sum_{i=1}^{\infty}\left(\frac{1}{2i}-\frac{1}{2i-1}\right)(2x-1)^{2i}.$$

Noting that $\left(\frac{1}{2i}-\frac{1}{2i-1}\right)<0$, we have

$$\begin{aligned}\lim_{\beta\to2^+}\frac{\Phi(\beta)-\Phi(2)}{\beta-2}&=\lim_{\beta\to2^+}\frac{\phi(x^*(\beta))}{\beta-2}\\&\leq\lim_{\beta\to2^+}\frac{1}{4}(2x^*(\beta)-1)^2,\end{aligned}$$

which is zero, as $x^*(\beta)\to\frac{1}{2}$ as $\beta\searrow 2$. The limit is also at least zero, since $\Phi(\beta)-\Phi(2)\geq0$ for all $\beta>2$ straightforwardly. Hence the right derivative of Φ is 0 and Φ is differentiable at $\beta=\beta_c$.

To show Φ is not twice-differentiable at $\beta_c=2$, we show that for $\beta\in(2,2+\epsilon_1)$, $\Phi(\beta)>f(\beta)$ for some smooth function f, with $f(2)=0$ and a strictly positive right derivative at 2. We have that

$$\Phi'(\beta)\geq\frac{\partial\phi}{\partial\beta}\Big|_{x^*(\beta)}=g(x^*(\beta))\geq g(x_0(\beta)),$$

where $g(x)=\frac{1}{4}(2x-1)^2$ and $x_0(\beta)$ is some function satisfying $\frac{1}{2}\leq x_0(\beta)\leq x^*(\beta)$, the last inequality coming from the monotonicity of g on the interval $(\frac{1}{2},1)$. We claim that $x_0(\beta)=(\frac{1}{10}(\beta-2))^{\frac{1}{2}}+\frac{1}{2}$ is such a function. We then define $f(\beta):=g(x_0(\beta))=\frac{1}{10}(\beta-2)$; clearly f satisfies the required conditions. It remains to prove the claim that $\frac{1}{2}\leq x_0(\beta)\leq x^*(\beta)$. Consider the inverse function of x_0, $\beta_0(x)=10(x-\frac{1}{2})^2+2$. By calculating that the first derivatives of β_0 and $\beta(x^*)$ (5.53) are both zero at $\frac{1}{2}$, and their second derivatives satisfy

$\beta_0''(\frac{1}{2}) > \beta''(\frac{1}{2}) > 0$, we see there exists an interval $(\frac{1}{2}, \frac{1}{2} + \epsilon_0)$ on which $\beta_0(x) \geq \beta(x)$. Now since both functions are also strictly increasing on this interval, there exists an interval $(2, 2 + \epsilon_1)$ on which their inverses have the reverse inequality, which is what we wanted to prove.

Now let $\theta \geq 3$. The function Φ is clearly continuous at $\beta = \beta_c$. Similarly to the $\theta = 2$ case, to show Φ is not differentiable at $\beta = \beta_c$, we show that on an interval $(\beta_c, \beta_c + \epsilon_2)$, $\Phi(\beta) \geq f(\beta)$, where f is some smooth function with $f(\beta_c) = 0$, whose right derivative at β_c is strictly positive. We have that

$$\Phi(\beta) = \phi(x^*(\beta)) \geq \phi\left(1 - \frac{1}{r}\right) =: f(\beta),$$

the inequality arising since x^* is defined to be the maximiser of ϕ. Lengthy calculations yield $\frac{\partial f}{\partial \beta} = \frac{(r-2)^2}{2r(r-1)}$, which is clearly positive for $r \geq 3$. This concludes the proof of Theorem 5.2.10. ■

Before finishing this subsection, let us prove that for $L_1 + L_1 > 0$, $L_2 > 0$, the finite (even) volume ground state is the vector given in (5.33). In order to give explicit finite volume ground states, we will need a concrete realisation of the eigenspaces of the Hamiltonian. This comes from our working in the proof of Theorem 3.0.3 and Proposition 3.0.11. Recall for $\underline{m}, \underline{m}'$ a pairing of $2k$ vertices in $\mathcal{V}$, let $Q_{\underline{m},\underline{m}'} = \prod_{i=1}^{k} Q_{m_i, m_i'}$, and recall $[Q_{\underline{m},\underline{m}'} \cdot \mathbb{V}]^0$ is the set of vectors in $Q_{\underline{m},\underline{m}'} \cdot \mathbb{V}$ which are killed by any $Q_{i,j}$ with $i, j \in \mathcal{V} \smallsetminus (\underline{m} \cup \underline{m}')$. Recall that by our working in Section 5.4, the eigenspaces of the Hamiltonian (5.20) (for any $\theta \geq 2$) are indexed by pairs (λ, ρ), partitions of $n - 2k$ $(0 \leq k \leq \lfloor \frac{n}{2} \rfloor)$ and n respectively, with $\lambda_1^\top + \lambda_2^\top \leq \theta$, $\rho_1^\top \leq \theta$, and $(\lambda, \rho) \in P_n(\theta)$. Let τ and π be standard tableaux (see Chapter 2) with shapes λ, ρ, and entries from $\mathcal{V} \smallsetminus (\underline{m} \cup \underline{m}')$ and $\mathcal{V}$ respectively. Following the proof of Proposition 3.0.11, the eigenspace itself can be realised as the span of the sets

$$z_\tau z_\pi \cdot \left[Q_{\underline{m},\underline{m}'} \cdot \mathbb{V}\right]^0, \tag{5.54}$$

where $z_\tau \in \mathbb{C}S_{|\mathcal{V} \smallsetminus (\underline{m} \cup \underline{m}')|}$ is a Young symmetriser (see (2.4)) for the partition λ acting on $\bigotimes_{i \in \mathcal{V} \smallsetminus (\underline{m} \cup \underline{m}')} V_i$, and $z_\rho \in \mathbb{C}S_n$ is a Young symmetriser for the partition ρ acting on all of $\mathbb{V}$. While the formula (5.54) is complicated for general pairs (λ, ρ), we will see that for some explicit pairs it simplifies greatly. By our working in Section 5.4, the dimension of the eigenspace is $d_\lambda^{O(\theta)} b_{\lambda,\rho}^{n,\theta} d_\rho^{S_n}$.

Let $\theta \geq 2$, $L_1 + L_1 > 0$, $L_2 > 0$. By our working in Section 5.4, the eigenvalues of the Hamiltonian are

$$-\left[(L_1 + L_2)\mathrm{ct}(\rho) - L_2(\mathrm{ct}(\lambda) + k(1 - \theta))\right], \tag{5.55}$$

indexed by pairs $(\lambda, \rho) \in P_n(\theta)$, where $\lambda \vdash n - 2k$. While (as noted in Section 5.5) we do not know the structure of $P_n(\theta)$ for $\theta > 3$, calculations yield that for n even, the eigenvalue is minimised in $\Lambda_n(\theta)$ at the pair $(\varnothing, (n))$, and by Remark 5.5.1, $(\varnothing, (n)) \in P_n(\theta)$. Now the dimension of the associated eigenspace is $d_\varnothing^{O(\theta)} b_{\varnothing,(n)}^{n,2} d_{(\theta)}^{S_n} = 1$, and using (5.54) it is straightforward to check it is spanned by (5.33).

5.6.2 Spin $\frac{1}{2}$, $\theta = 2$

Let us now prove Theorem 5.2.5, and justify the description of the phase diagram of the spin $S = \frac{1}{2}$ Heisenberg XXZ model illustrated in Figures 5.1a and 5.1b. Let $S = \frac{1}{2}$, so $\theta = 2$. Recall the Hamiltonian of the Heisenberg XXZ model is given by

$$H' = -\left(\sum_{i,j} K_1 S_i^{(1)} S_j^{(1)} + K_2 S_i^{(2)} S_j^{(2)} + K_1 S_i^{(3)} S_j^{(3)}\right). \tag{5.56}$$

We use the two identities $4(S_i^{(1)}S_j^{(1)} + S_i^{(2)}S_j^{(2)} + S_i^{(3)}S_j^{(3)}) + \mathrm{id} = 2T_{i,j}$ and $4(S_i^{(1)}S_j^{(1)} - S_i^{(2)}S_j^{(2)} + S_i^{(3)}S_j^{(3)}) + \mathrm{id} = 2Q_{i,j}$ (see, for example, Section 7 of [101]), which show that the Hamiltonian H' is, up to addition of a constant,

$$H = H(n, K_1, K_2) = -\frac{1}{4}\left(\sum_{i,j}(K_1 + K_2)T_{i,j} + (K_1 - K_2)Q_{i,j}\right). \tag{5.57}$$

Note that the line $K_1 = K_2 > 0$ gives the spin $S = \frac{1}{2}$ Heisenberg ferromagnet, and the line $K_1 = K_2 < 0$ gives the antiferromagnet. Let $Z_n(K_1, K_2) = \mathrm{Tr}\left[e^{-\frac{1}{n}H}\right]$, where H is from (5.57). Setting $L_1 = \frac{1}{4}(K_1 + K_2)$, $L_2 = \frac{1}{4}(K_1 - K_2)$ from Theorem 5.2.1 we have that the free energy of the system with Hamiltonian H given by (5.57) is

$$\lim_{n\to\infty} \frac{1}{n} \log Z_n(K_1, K_2) = \max_{(x,y)\in\Delta_2^*} \phi_{2,K_1,K_2}(x,y),$$

where we have

$$\phi_{2,K_1,K_2}((x_1,x_2),(y_1,0)) = \frac{1}{8}\left(2K_1(x_1^2 + x_2^2) + (K_2 - K_1)y_1^2\right) - \sum_{i=1}^{2} x_i \log(x_i),$$

and $\Delta_2^* = \{(x,y) \in ([0,1]^2)^2 \mid x_1 \geq x_2,\ x_1 + x_2 = 1,\ y_2 = 0,\quad 0 \leq y_1 \leq x_1 - x_2\}$.

Proof of Theorem 5.2.5. We analyse this free energy by considering different regions of the (K_1, K_2) plane. If $K_1 \geq K_2$ we can set $y_1 = 0$. This is the region covered by Theorem 5.2.2, and the free energy is exactly that of Theorem 1.1 from [13], with β from that paper replaced with $\frac{K_1}{2}$. The result of Theorem 5.2.10 shows that in this region, the free energy is smooth apart from at the line $K_1 = 4$, where it is differentiable, but not twice-differentiable.

Note that if we insert the condition $x_2 = 1 - x_1$ into $\phi_{2,K_1,K_2}(x_1)$ in this region $K_1 \geq K_2$ (with $y_1 = 0$), we can rewrite it as $\phi_{2,K_1,K_2}(x_1) = K_1(2x_1 - 1)^2 - x_1\log(x_1) - (1 - x_1)\log(1 - x_1)$. Now consider the region $K_2 \geq K_1$. We have to set $y_1 = x_1 - x_2$ in order to maximise ϕ. Rearranging, and inserting $x_2 = 1 - x_1$ now gives almost the same function as above, but with K_1 replaced with K_2:

$$\phi_{2,K_1,K_2}(x_1) = K_2(2x_1 - 1)^2 + K_1 - x_1\log(x_1) - (1 - x_1)\log(1 - x_1). \tag{5.58}$$

The extra term K_1 does not affect the location of the maximiser. So, in the region $K_2 \geq K_1$, the free energy is (up to the addition of K_1) that of Theorem 5.2.2 and Theorem 1.1 from [13]. So by our proof of Theorem 5.2.10, it is smooth everywhere in the region $K_2 \geq K_1$ apart from the line $K_2 = 4$, where it is differentiable but not twice-differentiable.

It remains to join the two regions $K_1 \geq K_2$ and $K_2 \geq K_1$ together. Clearly the free energy is continuous on the whole plane. The above working shows that in the region $K_1 \leq 4$, $K_2 \leq 4$, the maximiser of ϕ_{2,K_1,K_2} is at $((\frac{1}{2},\frac{1}{2}),(0,0))$, so the free energy is smooth in this region. To conclude, let us consider the free energy on a line $K_1 = C \geq 4$ as it crosses the half-line $K_1 = K_2 \geq 4$. For $K_2 \leq 4$ on this line it is constant by our working above. If we denote the free energy by $\Phi(K_1,K_2)$, then using (5.58), for $K_2 > 4$,

$$\frac{\partial \Phi}{\partial K_2} = \frac{\partial(\phi_{2,K_1,K_2}(x^*))}{\partial K_2} \geq \frac{\partial \phi_{2,K_1,K_2}}{\partial K_2}\Big|_{x^*(K_2)} = (2x_1^*(K_2)-1)^2 > 0,$$

the last inequality coming from our working in the proof of Theorem 5.2.10. Hence the free energy is not differentiable on the half-line $K_1 = K_2 \geq 4$, which completes the proof of Theorem 5.2.5. ∎

Proof of Proposition 5.2.6. Let us now comment on the phase diagram that Theorem 5.2.5 indicates, and in the process prove Proposition 5.2.6. We label the region $K_1 \leq 4$, $K_2 \leq 4$ (the region where $(x^*,y^*) = ((\frac{1}{2},\frac{1}{2}),(0,0))$ maximises ϕ_{2,K_1,K_2}) the disordered phase, since it maximises the entropy term (the logarithms) in ϕ_{2,K_1,K_2}. It is illustrated as the solid pink region in Figure 5.1b. The maximiser $y_1^* = (0,0)$ gives the magnetisation of Theorem 5.2.3 $y_1^{\uparrow} = 0$.

We label the region $K_2 > K_1$, $K_2 > 4$ the Ising phase, illustrated as the dotted yellow region in Figure 5.1b. Proposition 5.7.8 and our working to prove Theorem 5.2.5 show that the maximiser of ϕ_{2,K_1,K_2} is unique in the Ising phase, and of the form $(x^*,y^*) = ((x_1^*,x_2^*),(x_1^*-x_2^*,0))$, with $x_1^* > x_2^*$. Then the magnetisation $y_1^{\uparrow}$ of Theorem 5.2.3 is strictly positive.

As $\|(K_1,K_2)\| \to \infty$, the maximiser of ϕ_{2,K_1,K_2} tends to $((1,0),(1,0))$. Recall that from our working in Section 5.4 and Proposition 5.7.8, the eigenvalues of the Hamiltonian (5.57) are given by

$$-\left[2K_1 \mathrm{ct}(\rho) - (K_1-K_2)(\mathrm{ct}(\lambda) + k(1-\theta))\right], \tag{5.59}$$

where (λ,ρ) are partitions of $n-2k$ $(0 \leq k \leq \lfloor \frac{n}{2} \rfloor)$ and n respectively, with $\lambda_1^{\mathsf{T}} + \lambda_2^{\mathsf{T}} \leq 2$, $\rho_1^{\mathsf{T}} \leq 2$, and $\lambda_1 \leq \rho_1 - \rho_2$. It is not hard to see that for $K_2 > K_1$, $K_2 > 0$, the finite volume ground states are the eigenspace corresponding to the pair $(\lambda,\rho) = ((n),(n))$. Using (5.54), this is the space of vectors invariant under the action of S_n, and killed by any $Q_{i,j}$, and has dimension $d_{(n)}^{O(2)} b_{(n),(n)}^{n,2} d_{(n)}^{S_n} = 2$. A dimension count shows that it is therefore spanned by the two product states $\otimes_{1\leq j\leq n}(|\frac{1}{2}\rangle \pm i|-\frac{1}{2}\rangle)$ (where i here is $\sqrt{-1}$). Further, if we consider the Hamiltonian with a magnetisation term $-h\sum_{1\leq i\leq n} S_i^{(2)}$ added, since these product states are eigenvectors of the magnetisation term, for h small and positive $\otimes_{1\leq j\leq n}(|\frac{1}{2}\rangle + i|-\frac{1}{2}\rangle)$ must be the unique ground state, and vice-versa for h small and negative.

We label the region $K_1 > K_2$, $K_1 > 4$ the XY region, illustrated as the hatched blue region in Figure 5.1b. By our working in the proof of Theorem 5.2.5, the maximiser of ϕ_{2,K_1,K_2} is unique and of the form $((x_1^*,x_2^*),(0,0))$, so the magnetisation $y_1^{\uparrow}$ from Theorem 5.2.3 is zero. As $\|(K_1,K_2)\| \to \infty$, the maximiser of ϕ_{2,K_1,K_2} tends to $((1,0),(0,0))$, and as we have already shown in arbitrary spins, the finite volume ground states are given by the eigenspace corresponding to the pair $(\lambda,\rho)=(\varnothing,(n))$. This space is one-dimensional, since $d_{\varnothing}^{O(2)}=b_{\varnothing,(n)}^{n,2}=d_{(n)}^{S_n}=1$, and using (5.54), is spanned by the vector (5.28).

On the half-line $K_1=K_2>4$ (the supercritical isotropic Heisenberg model), the y term in ϕ_{2,K_1,K_2} disappears, so if (x^*,y^*) is a maximiser of ϕ_{2,K_1,K_2}, then (x^*,y) is too, so long as $(x^*,y)\in\Delta_2^*$. Hence $y_1^{\uparrow}=x_1^*-x_2^*>0$ by the proof of Theorem 5.2.10, or [13]. This concludes the proof of Proposition 5.2.6. ∎

From Theorem 5.2.5, the transition from the disordered to either of the other two phases is second order, and from XY to Ising is first order. By our working above, the transitions from the Ising to the other phases can also be observed in the quantities from Theorems 5.2.3 and 5.2.4, since $y_1^{\uparrow}=y_1^*>0$ in the Ising phase, and is zero in the other phases. This transition in $y_1^{\uparrow}=y_1^*$ is continuous in the Ising-disordered transition, and discontinuous in the Ising-XY transition.

5.6.3 Spin 1; $\theta=3$

Proof of Theorem 5.2.7. Let $S=1$, and recall the Hamiltonian of the bilinear-biquadratic Heisenberg model:

$$H''=-\left(\sum_{i,j}J_1(\boldsymbol{S}_i\cdot\boldsymbol{S}_j)+J_2(\boldsymbol{S}_i\cdot\boldsymbol{S}_j)^2\right). \tag{5.60}$$

where $J_1,J_2\in\mathbb{R}$ and $\boldsymbol{S}_i\cdot\boldsymbol{S}_j=\sum_{k=1}^3 S_i^{(k)}S_j^{(k)}$. Let $P_{i,j}$ be (a scalar multiple of) the spin-singlet operator, given by

$$\langle a_i,a_j|P_{i,j}|b_i,b_j\rangle=(-1)^{a_i-b_i}\delta_{a_i,-a_j}\delta_{b_i,-b_j}. \tag{5.61}$$

Note that the line $J_2=0$ gives the Heisenberg ferromagnet $(J_1>0)$, and antiferromagnet $(J_1<0)$. We use the relations $\boldsymbol{S}_i\cdot\boldsymbol{S}_j=T_{i,j}-P_{i,j}$ and $(\boldsymbol{S}_i\cdot\boldsymbol{S}_j)^2=P_{i,j}+\mathrm{id}$ (see Lemma 7.1 from [101]) to show that Hamiltonian (5.29) is, up to addition of a constant,

$$H(n,J_1,J_2)=-\left(\sum_{i,j}J_1T_{i,j}+(J_2-J_1)P_{i,j}\mathrm{id}\right). \tag{5.62}$$

Let $Z_n(J_1,J_2)=\mathrm{Tr}\left[e^{-\frac{1}{n}H}\right]$, where H is given by 5.62. Ueltschi (Theorem 3.2 of [101]) shows that for θ odd, this partition function is the same as when $P_{i,j}$ is replaced with $Q_{i,j}$. For completeness, we show that this equality can be derived from an isomorphism of representations (Lemma 5.9.1). Now, setting $L_1=J_1$, $L_2=J_2-J_1$, Theorem 5.2.1 shows that the free energy of the model with Hamiltonian (5.62) is

$$\lim_{n\to\infty}\frac{1}{n}\log Z_n(J_1,J_2)=\max_{(x,y)\in\Delta_3^*}\phi_{3,J_1,J_2}(x,y),$$

where we have

$$\phi_{3,J_1,J_2}((x_1,x_2,x_3),(y_1,0,0))=\frac{1}{2}\left(J_2\sum_{i=1}^3 x_i^2+(J_1-J_2)y_1^2\right)-\sum_{i=1}^3 x_i\log(x_i). \tag{5.63}$$

The proof of Theorem 5.2.7 now follows from Theorem 5.2.10 using the change of variables above.

■

Proof of Proposition 5.2.9. In the rest of this section we provide the proof of Proposition 5.2.9, which backs up Remark 5.2.8 and our description of the phases of the bilinear-biquadratic Heisenberg model, illustrated in Figure 5.2b.

Let $\phi = \phi_{3,J_1,J_2}$. We define the disordered phase to be the set $\mathcal{A}$ of values of (J_1, J_2) such that $((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), (0, 0, 0))$ is a maximiser of ϕ; this maximises the entropy term (the logarithms) of ϕ.

Let us prove Proposition 5.2.9 first in the region $J_2 > J_1$. Here, we set $y_1 = 0$, so ϕ reduces to ϕ^{int}, and the disordered phase is the region $J_2 \leq \log 16$, by [13]. We label the region $J_2 > J_1$, $J_2 > \log 16$ the nematic phase. It is illustrated as the hatched blue region in Figure 5.2b. As noted above, we must set $y_1 = 0$, so the magnetisation $y_1^{\uparrow}$ in Theorem 5.2.3 is zero in this phase. We can say that the transition from disordered to nematic is first order, by Theorem 5.2.7. Lastly, let us show that for $J_2 > J_1$, $J_2 > 0$, the finite (even) volume ground state is the vector (5.30). By our working in Section 5.4 and Proposition 5.7.8, the eigenvalues of the Hamiltonian (5.62) are given by

$$-\left[J_2 \mathrm{ct}(\rho) + (J_1 - J_2)(\mathrm{ct}(\lambda) + k(1-\theta))\right], \tag{5.64}$$

where (λ, ρ) are partitions of $n - 2k$ $(0 \leq k \leq \lfloor \frac{n}{2} \rfloor)$ and n, respectively, with $\lambda_1^{\mathsf{T}} + \lambda_2^{\mathsf{T}} \leq 3$, $\rho_1^{\mathsf{T}} \leq 3$, and $\lambda_1 \leq \rho_1 - \rho_3$. For $J_2 > J_1$, $J_2 > 0$, as we have already shown in arbitrary spins, this is minimised by the pair $(\lambda, \rho) = (\varnothing, (n))$, the corresponding eigenspace has dimension $d_{\varnothing}^{O(3)} b_{\varnothing,(n)}^{n,2} d_{(n)}^{S_n} = 1$, and using (5.54), the unique ground state of the transformed Hamiltonian (i.e. (5.62) with $P_{i,j}$ replaced with $Q_{i,j}$) is the vector given by the sum over all $\frac{n}{2}$-fold tensor products of the vector $\sum_{a=-1}^{1} |a, a\rangle$. Transforming this back to the original Hamiltonian, we have the sum over all possible tensor products of singlet states, which is precisely (5.30).

We can now turn to proving Proposition 5.2.9 in the region $J_2 \leq J_1$; this region is more complicated. The function ϕ does not reduce to ϕ^{int}. We must let $y_1 = x_1 - x_3$. Setting $x_3 = 1 - x_1 - x_2$, we rewrite ϕ as a function of x_1 and x_2:

$$\begin{aligned}\phi = \phi_{3,J_1,J_2}(x_1, x_2) = \frac{1}{2}&\left(J_2(-2x_1^2 + x_2^2 - 2x_1x_2 + 2x_1) + J_1(2x_1 + x_2 - 1)^2\right) \\ &- x_1 \log(x_1) - x_2 \log(x_2) - (1 - x_1 - x_2)\log(1 - x_1 - x_2).\end{aligned} \tag{5.65}$$

Note we are analysing this function in the region R defined by $x_1 \geq x_2$, $1 - x_2 \geq x_1 \geq 1 - 2x_2$ (see Figure 5.7).

In this region $J_1 \geq J_2$, the boundary of the disordered phase $\mathcal{A}$ is difficult to identify - recall we will show it is a curve $\mathcal{C}$ made up of the half-line $J_2 = 2J_1 - 3 \leq \frac{3}{2}$ and a curve connecting the points $(\frac{9}{4}, \frac{3}{2})$ and $(\log 16, \log 16)$. Outside of the disordered phase $\mathcal{A}$ (within the region $J_1 \geq J_2$), we can show that $y_1^{\uparrow}$ from Theorem 5.2.3 is strictly positive. Indeed, if (x^*, y^*) is a maximiser of ϕ_{3,J_1,J_2}, then $y_1^* = x_1^* - x_3^*$, meaning the only point with $y_1^* = 0$ is $((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), (0, 0, 0))$, and the claim follows from the definition of the disordered phase $\mathcal{A}$. Numerical simulations suggest that $y_1^{\uparrow}$ is a unique maximiser everywhere in $J_1 \geq J_2$ except for the curve between $(\frac{9}{4}, \frac{3}{2})$ and $(\log 16, \log 16)$ which is part of the curve C, so the y_1^*

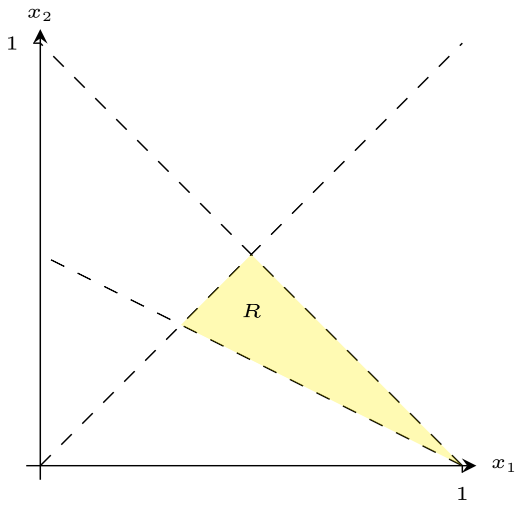


Figure 5.7: The region R.

form Theorem 5.2.4 would exist and be positive; we have not been able to prove this.

Let us consider the ground state behaviour outside the disordered phase. As $\|(J_1, J_2)\| \to \infty$, the logarithm terms in ϕ will become negligible. Let ϕ_0 be ϕ with the logarithm terms removed. We maximise ϕ_0 in the region R. Setting $x_3 = 1 - x_1 - x_2$, we have

$$\begin{aligned}\frac{\partial \phi_0}{\partial x_1} &= (2J_1 - J_2)(2x_1 + x_2 - 1);\\ \frac{\partial \phi_0}{\partial x_2} &= J_1(2x_1 + x_2 - 1) + J_2(x_2 - x_1).\end{aligned} \tag{5.66}$$

Now since $2J_1 - J_2 > 0$, (and as we take our limit we are beyond the conjectured boundary $\mathcal{C}$), so the maximum of ϕ_0 must lie on the boundary line $x_1 + x_2 = 1$ of R. Note this implies $x_3 = 0$, and so $y_1 = x_1$. Substituting $x_2 = 1 - x_1$, and rearranging, we have the quadratic

$$\phi_0(x_1) = \frac{J_1 + J_2}{2}\left(\left(x_1 - \frac{J_2}{J_1 + J_2}\right)^2 + \frac{J_1 J_2}{(J_1 + J_2)^2}\right), \tag{5.67}$$

where recall we are concerned with the region $x_1 \in [\frac{1}{2}, 1]$.

Calculations yield that in the region $J_2 < J_1$, $J_1 \geq 0$, this quadratic has maximum at $x_1 = 1$. So the maximiser of ϕ_{3,J_1,J_2} in this region as $\|(J_1, J_2)\| \to \infty$ tends to $((1,0,0),(1,0,0))$. Indeed the finite volume ground states (of the transformed Hamiltonian (5.62)) are given by the eigenspace corresponding to the pair $(\lambda, \rho) = ((n),(n))$. This space has dimension $d^{O(3)}_{(n)} b^{n,2}_{(n),(n)} d^{S_n}_{(n)} = 2n + 1$, and using (5.54), is the set of vectors invariant under S_n (equivalently, invariant under any $T_{i,j}$), which are killed by any $Q_{i,j}$. The corresponding eigenspace of the original Hamiltonian is the set invariant under S_n and killed by any $P_{i,j}$. Straightforward analysis of $P_{i,j}$ shows that the product states $\bigotimes_{1 \leq i \leq n} |a\rangle$ with $a_0^2 - a_1 a_{-1} = 0$ lie in this set (although they do not span it), which include the ferromagnetic $|a\rangle = |1\rangle$ and $|-1\rangle$ as well as $|1\rangle + |0\rangle + |-1\rangle$. We label this region, $J_2 < J_1$, $J_1 \geq 0$ and to the right of the curve $\mathcal{C}$, ferromagnetic. It is illustrated as the dotted yellow region in Figure 5.2b.

Now consider the region $2J_1 - J_2 > 0$, and $J_1 < 0$, illustrated by the checkerboard orange region in Figure 5.2b. In this case, the quadratic (5.67) has maximum at $x_1 = \alpha := \frac{J_2}{J_1 + J_2}$, which lies in the range $[\frac{2}{3}, 1]$. Then the maximiser of ϕ_{3,J_1,J_2} is $((\alpha, 1-\alpha, 0), (\alpha, 0, 0))$. While we do not label this fourth phase, occupying the region $0 > J_1 > \frac{1}{2}(3 - J_2)$, this phase has some similarity with the ferromagnetic phase.

In finite volume, calculations from analysing (5.64) show that the set of ground states in the fourth phase is the eigenspace corresponding to a pair $(\lambda, \rho) = ((\alpha'), (\alpha', 1-\alpha'))$, where $\alpha' = \lambda_1/n$ is close to $\alpha = \frac{J_2}{J_1+J_2}$ (and tends to α as $n \to \infty$). Using (5.54), the eigenspace is spanned by vectors (5.31).

The rest of this section completes the proof of Proposition 5.2.9 by determining the boundary of the disordered phase $\mathcal{A}$ within the region $J_1 \geq J_2$. Recall we will show it is a curve $\mathcal{C}$ made up of the half-line $J_2 = 2J_1 - 3 \leq \frac{3}{2}$ and a curve connecting the points $(\frac{9}{4}, \frac{3}{2})$ and $(\log 16, \log 16)$. From here till the end of the section we work with ϕ given in (5.65). The partial derivatives of ϕ of first and second order are:

$$\begin{aligned}
\frac{\partial \phi}{\partial x_1} &= (2J_1 - J_2)(2x_1 + x_2 - 1) - \log(x_1) + \log(1 - x_1 - x_2);\\
\frac{\partial \phi}{\partial x_2} &= J_1(2x_1 + x_2 - 1) + J_2(x_2 - x_1) - \log(x_2) + \log(1 - x_1 - x_2);\\
\frac{\partial^2 \phi}{\partial x_1^2} &= 2(2J_1 - J_2) - \frac{1}{x_1} - \frac{1}{1 - x_1 - x_2};\\
\frac{\partial^2 \phi}{\partial x_1 \partial x_2} &= (2J_1 - J_2) - \frac{1}{1 - x_1 - x_2};\\
\frac{\partial^2 \phi}{\partial x_2^2} &= J_1 + J_2 - \frac{1}{x_1} - \frac{1}{1 - x_1 - x_2}.
\end{aligned} \tag{5.68}$$

Lemma 5.6.1. *The point $(x_1, x_2) = (\frac{1}{3}, \frac{1}{3})$ is always an inflection point of ϕ, and it is a local maximum point if $2J_1 - J_2 < 3$, and if $2J_1 - J_2 > 3$ it is not a local maximum point and does not maximise ϕ in R.*

Proof. Setting $(x_1, x_2) = (\frac{1}{3}, \frac{1}{3})$ in the above shows it is always an inflection point. If $\mathfrak{H}$ is the Hessian matrix of ϕ, then for any vector $(p, q) \in \mathbb{R}^2$, we have

$$(p, q)\mathfrak{H}(p, q)^{\mathsf{T}} = (p + q)^2(2J_1 - J_2 - 3) + p^2((2J_1 - J_2 - 3) + b^2(2J_2 - J_1 - 3)),$$

which is negative for all $2J_1 - J_2 < 3$ in our region $J_2 \leq J_1$, meaning $(\frac{1}{3}, \frac{1}{3})$ is a local maximum. Clearly for $2J_1 - J_2 > 3$, $\frac{\partial^2 \phi}{\partial x_1^2} > 0$, so $(\frac{1}{3}, \frac{1}{3})$ is not a local maximum, and it cannot maximise ϕ in the region R. ∎

Let $\mathcal{A}'$ be the region within the region $J_2 \leq J_1$ where $(\frac{1}{3}, \frac{1}{3})$ is a global maximum of ϕ in R (this is the region $\mathcal{A}$ intersected with $J_2 \leq J_1$). By the above, all of $\mathcal{A}'$ must lie within the region $2J_1 - J_2 \leq 3$, (or, not to the right of the line $J_2 = 2J_1 - 3$).

Lemma 5.6.2. *The set $\mathcal{A}'$ is convex.*

Proof. Let $J^{(1)}, J^{(2)}$ be two points in $\mathcal{A}'$. Let $J = sJ^{(1)} + (1-s)J^{(2)}$, $s \in [0, 1]$. Since ϕ is linear in J_1, J_2, we have that for any $(x_1, x_2) \in R$,

$$\begin{aligned}\phi_{2,J}(x_1,x_2) &= s\phi_{2,J^{(1)}}(x_1,x_2) + (1-s)\phi_{2,J^{(2)}}(x_1,x_2) \\ &\le s\phi_{2,J^{(1)}}\left(\frac{1}{3},\frac{1}{3}\right) + (1-s)\phi_{2,J^{(2)}}\left(\frac{1}{3},\frac{1}{3}\right) = \phi_{2,J}\left(\frac{1}{3},\frac{1}{3}\right).\end{aligned}$$

■

Lemma 5.6.3. *If a point $(J_1^{(0)}, J_2^{(0)})$ within the region $J_1 \ge J_2$ lies outside of $\mathcal{A}'$, then the point $(J_1^{(0)}, J_2^{(0)}) + v$, when it lies within the region $J_1 \ge J_2$, also lies outside of $\mathcal{A}'$, for any $v = \mu(1,2) + \nu(-1,-3)$, $\mu, \nu > 0$.*

Proof. We have that

$$\begin{aligned}\frac{\partial}{\partial J_1}\left(\phi(x_1,x_2) - \phi\left(\frac{1}{3},\frac{1}{3}\right)\right) &= \frac{1}{2}(2x_1 + x_2 - 1)^2 \\ \frac{\partial}{\partial J_2}\left(\phi(x_1,x_2) - \phi\left(\frac{1}{3},\frac{1}{3}\right)\right) &= \frac{1}{2}\left(1 - 2x_1^2 + x_2^2 - 2x_1x_2 + 2x_1 - \frac{4}{3}\right).\end{aligned}$$

Firstly, we consider:

$$\left(\frac{\partial}{\partial J_1} + 2\frac{\partial}{\partial J_2}\right)\left(\phi(x_1,x_2) - \phi\left(\frac{1}{3},\frac{1}{3}\right)\right) = \frac{3}{2}x_2^2 - x_2 + \frac{1}{6} = \frac{1}{6}(3x_2 - 1)^2,$$

which has single root and minimum at $x_2 = \frac{1}{3}$. Let $(J_1^{(0)}, J_2^{(0)})$ lie outside of $\mathcal{A}'$, so there exists some global maximiser $(x_1^*, x_2^*) \ne (\frac{1}{3}, \frac{1}{3})$ in R, $\phi(x_1^*, x_2^*) > \phi(\frac{1}{3}, \frac{1}{3})$. Now the above shows that moving $(J_1^{(0)}, J_2^{(0)})$ in the direction $(1,2)$ does not increase ϕ at $(\frac{1}{3}, \frac{1}{3})$ any faster than at any other point of R, so for all $\mu > 0$, $(J_1^{(0)}, J_2^{(0)}) + \mu(1,2)$ cannot lie in $\mathcal{A}'$.

Secondly,

$$\begin{aligned}\left(-\frac{\partial}{\partial J_1} - 3\frac{\partial}{\partial J_2}\right) &= -\frac{1}{2}\left(1 - 2x_1^2 + x_2^2 - 2x_1x_2 + 2x_1 - \frac{4}{3}\right) \\ &= \left(\left(x_1 - \frac{1}{3}\right) + \frac{1}{2}\left(x_2 - \frac{1}{3}\right)\right)^2 - \frac{9}{4}\left(x_2 - \frac{1}{3}\right)^2,\end{aligned}$$

which takes the value zero exactly on the lines $x_1 = x_2$ and $x_1 = 1 - 2x_2$, two of the boundary lines of R, and is positive in the rest of R. By the same argument as above, if $(J_1^{(0)}, J_2^{(0)}) \notin A$, then $(J_1^{(0)}, J_2^{(0)}) + \nu(-1,-3) \notin A$, for all $\nu > 0$. The lemma follows. ■

Lemma 5.6.4. *The region bounded by and including the line $J_1 = J_2$, $J_1 \le \log(16)$, the line $2J_1 - J_2 = 3, J_2 \le \frac{3}{2}$, and the straight line from the point $(\log(16), \log(16))$ to the point $(\frac{9}{4}, \frac{3}{2})$, lies within $\mathcal{A}'$.*

Proof. To begin with, note that on the line $J_1 = J_2$, we can use our results from the case $J_2 \ge J_1$. This means all $J_1 = J_2$, $J_1 \le \log(16)$ lie in $\mathcal{A}'$. Now the previous lemma implies that all $2J_1 - J_2 \le \log(16)$ lies in $\mathcal{A}'$, since if it were not true, we would be able to move from a point not in $\mathcal{A}'$ in the direction $(1,2)$ and arrive at a point in $\mathcal{A}'$. Now by the same logic, and the fact that $\mathcal{A}'$ is convex, it suffices to show that the point $(J_1, J_2) = (\frac{9}{4}, \frac{3}{2})$ lies in $\mathcal{A}'$. We show that at this point, there are no inflection points of ϕ besides $(\frac{1}{3}, \frac{1}{3})$, and $(\frac{1}{3}, \frac{1}{3})$ maximises ϕ on the boundary of R.

Set $(J_1, J_2) = (\frac{9}{4}, \frac{3}{2})$. Substituting $z = 1 - x_1 - x_2, w = x_1$ into (5.68) gives $\frac{\partial \phi}{\partial x_1} = 0$ if and only if $z = 1$ or

$$w = \frac{-\log(z)}{3(1-z)}. \tag{5.69}$$

Note that the region R is transformed into R', given by $\frac{1}{2}(\frac{1}{w} - 1) \geq z \geq \frac{1}{w} - 2$, $zw \geq 0$. The line $z = 1$ intersects R' at the single point $(w, z) = (\frac{1}{3}, 1)$, which corresponds to $(x_1, x_2) = (\frac{1}{3}, \frac{1}{3})$. Substituting (5.69) into (5.68) gives that, on the line where $\frac{\partial \phi}{\partial x_1} = 0$, the value of $\frac{\partial \phi}{\partial x_2}$ is:

$$\frac{\partial \phi}{\partial x_1}(z) = \frac{3}{2} + \frac{\log(z)(1+5z)}{4(1-z)} + \log\left(\frac{-z\log(z)}{3(1-z)+(1+z)\log(z)}\right).$$

Remark 5.6.5. Let r be the unique zero of $3(1-z) + (1+z)\log(z)$. This function $\frac{\partial \phi}{\partial x_1}(z)$ is positive in the range $(r, 1)$, except at $z = 1$, where it is zero. (It is not defined in $(0, r]$).

Hence either there are no points of inflection in R, or $(\frac{1}{3}, \frac{1}{3})$ is the only one. Proving Remark 5.6.5 by hand is difficult. However, a rigorous computer-assisted argument is available, which is due to Dave Platt. See Appendix 5.8.

It remains to analyse ϕ on the boundary of R. Substituting $x_1 = 1 - 2x_2$ into ϕ, we have

$$\phi(x_2) = \frac{15}{8} - \frac{21}{4}x_2 + \frac{63}{8}x_2^2 - (1-2x_2)\log(1-2x_2) - 2x_2\log(x_2),$$

which it is not hard to prove is maximised at $x_2 = \frac{1}{3}$ in the region $x_2 \in [0, \frac{1}{2}]$. Indeed, its first derivative $\frac{1}{4}(-21 + 63x_2 + 8\log(1-2x_2) - 8\log(x_2))$ is zero at $x_2 = 0$, and its second derivative $\frac{126x_2^2 - 62x_2 + 8}{8x^2 - 4x}$ is negative in the range. Substituting $x_1 = x_2$ into ϕ gives exactly the same function as above. As $x_1 + x_2 \to 1$, the first order derivatives of ϕ tend to $-\infty$. Hence ϕ on the boundary of R must be maximised at $(\frac{1}{3}, \frac{1}{3})$, so the same holds over all R, and so we can conclude that $(\frac{9}{4}, \frac{3}{2}) \in \mathcal{A}'$, which is what we wanted to prove; this completes the proof of Lemma 5.6.4.

∎

Combining the above lemmas give us the information we need about the boundary of $\mathcal{A}'$. Lemma 5.6.2 implies that its boundary exists, and adding Lemmas 5.6.3, 5.6.1 and 5.6.4 shows that its boundary is made up of the line $J_1 = J_2 \leq \log(16)$, and a curve $\mathcal{C}$ which (as a function of J_1) is a continuous, convex line, which is the line $2J_1 - J_2 = 3$ for $J_2 \leq \frac{3}{2}$, and that its gradient lies in $[2, 3]$. This curve must meet the line $J_1 = J_2$ at the point $(\log(16), \log(16))$. Indeed, recall $\mathcal{A}$ is the region of the whole plane where $((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), (0, 0, 0))$ maximises our original $\phi_{3,J_1,J_2}(x, y)$; we have $\mathcal{A}' = \mathcal{A} \cap \{J_2 \leq J_1\}$. Then the same proof as above can be employed to show that $\mathcal{A}$ is convex, which is what we need. This completes the description of the boundary of $\mathcal{A}$, which in turn completes the proof of Proposition 5.2.9. ∎

5.7 Branching Coefficients

As noted in Section 5.4, the aim of this section is to prove Propositions 5.7.6 and 5.7.8, which determine the sets $P_n(\theta)$ and the limits $\frac{1}{n}P_n(\theta)$, $\theta = 2, 3$. Recall $\Lambda_n(\theta)$ is the set of pairs of partitions (λ, ρ), $\lambda \vdash n - 2k$, $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$, $\rho \vdash n$, such that $\lambda_1^\top + \lambda_2^\top \leq \theta$ and $\rho_1^\top \leq \theta$. Recall that $b_{\lambda,\rho}^{n,\theta}$ is the coefficient of the irreducible ρ in the restriction of the irreducible λ from $\mathbb{B}_{n,\theta}$ to $\mathbb{C}S_n$. Then $P_n(\theta)$ is the set of $(\lambda, \rho) \in \Lambda_n(\theta)$ such that $b_{\lambda,\rho}^{n,\theta} > 0$. Most of the work in proving Propositions 5.7.6 and 5.7.8 is contained in three lemmas which we begin this section with. The first shows that the coefficients $b_{\lambda,\rho}^{n,\theta}$ are also the branching coefficients of the orthogonal and general linear groups. The second is a useful recurrence relation, and the third determines $b_{\lambda,\rho}^{n,\theta}$ for certain values of ρ, in terms of the Littlewood-Richardson coefficients.

5.7.1 Useful lemmas for all θ

Fix $\theta \geq 2$. First we rephrase the coefficients $b_{\lambda,\rho}^{n,\theta}$ in terms of the general linear and orthogonal groups, using Schur-Weyl duality. Recall that the irreducible polynomial representations of $GL(\theta)$ are indexed by ρ, partitions of any non-negative integer with at most θ parts. Similarly, those of $O(\theta)$ are indexed by λ, partitions of any non-negative integer whose first two columns sum to at most θ. Let $\rho \vdash n$, and let $g_{\lambda,\rho}^{n,\theta}$ denote the coefficient of $\psi_\lambda^{O(\theta)}$ in the restriction of $\psi_\rho^{GL(\theta)}$ from $GL(\theta)$ to $O(\theta)$.

Lemma 5.7.1. *The symmetric group-Brauer algebra and orthogonal group-general linear group branching coefficients are the same. That is, for all $(\lambda, \rho) \in \Lambda_n(\theta)$, we have that $g_{\lambda,\rho}^{n,\theta} = b_{\lambda,\rho}^{n,\theta}$.*

Proof. Recall that Schur-Weyl duality (5.36) states that as a module of $\mathbb{B}_{n,\theta} \otimes \mathbb{C}O_n(\mathbb{C})$,

$$\mathbb{V} = \bigoplus_{\substack{\lambda \vdash n-2k \\ \lambda_1^\top + \lambda_2^\top \leq \theta}} \psi_\lambda^{\mathbb{B}_{n,\theta}} \boxtimes \psi_\lambda^{O(\theta)}. \tag{5.70}$$

The equivalent statement for the symmetric and general linear groups says that as a module of $\mathbb{C}S_n \otimes \mathbb{C}GL_n(\mathbb{C})$,

$$\mathbb{V} = \bigoplus_{\substack{\rho \vdash n \\ \rho_1^\top \leq \theta}} \psi_\rho^{S_n} \boxtimes \psi_\rho^{GL(\theta)}. \tag{5.71}$$

Restricting each $\psi_\lambda^{\mathbb{B}_{n,\theta}}$ in the first equation to $\mathbb{C}S_n$, and each $\psi_\rho^{GL(\theta)}$ in the second to $O_n(\mathbb{C})$, we have, as a module of $\mathbb{C}S_n \otimes \mathbb{C}O_n(\mathbb{C})$,

$$\bigoplus_{(\lambda,\rho) \in \Lambda_n(\theta)} b_{\lambda,\rho}^{n,\theta} \psi_\rho^{S_n} \boxtimes \psi_\lambda^{O(\theta)} = \bigoplus_{(\lambda,\rho) \in \Lambda_n(\theta)} g_{\lambda,\rho}^{n,\theta} \psi_\rho^{S_n} \boxtimes \psi_\lambda^{O(\theta)}, \tag{5.72}$$

and hence the result. ∎

From hereon in we simply use $b_{\lambda,\rho}^{n,\theta}$ to denote either itself or $g_{\lambda,\rho}^{n,\theta}$. For λ a partition with $\lambda_1^\top + \lambda_2^\top \leq \theta$, recall λ' is the partition such that $(\lambda')_1^\top = \theta - \lambda_1^\top$, and $(\lambda')_j^\top = \lambda_j^\top$ for all $j > 1$. Note that $\lambda'' = \lambda$. We next prove a useful recurrence relation. Let

Lemma 5.7.2. *The symmetric group-Brauer algebra branching coefficients satisfy the following recurrence relation. Let $(\lambda, \rho) \in \Lambda_n(\theta)$, such that $\rho_\theta > 0$. Then $b^{n,\theta}_{\lambda,\rho} = b^{n-\theta,\theta}_{\lambda',\rho-\underline{1}}$, where $\underline{1}$ is the partition with all parts equal to* 1.

Remark 5.7.3. Note that as a consequence, if $(\lambda, \rho) \in \Lambda_n(\theta)$, $\rho_\theta > 0$, then

$$b^{n,\theta}_{\lambda,\rho} = \begin{cases} b^{n,\theta}_{\lambda,\rho-\underline{\rho_\theta}} & \rho_\theta \text{ even} \\ b^{n,\theta}_{\lambda',\rho-\underline{\rho_\theta}} & \rho_\theta \text{ odd,} \end{cases}$$

where $\underline{\rho_\theta}$ is the partition with all parts equal to ρ_θ.

Proof. We use the fact that $b^{n,\theta}_{\lambda,\rho}$ is a coefficient in the restriction of the irreducible $\psi^{GL(\theta)}_\rho$ of $GL(\theta)$ to $O(\theta)$. Recall the character orthogonality of the orthogonal group from (2.3). We have

$$b^{n,\theta}_{\lambda,\rho} = \int_{O(\theta)} \chi^{GL(\theta)}_\rho(g)\chi^{O(\theta)}_\lambda(g)\ dg, \tag{5.73}$$

where dg denotes the Haar measure on the orthogonal group. By the Pieri rule (2.35) (or, for example, the remarks after equation (1) of [93]), $\chi^{GL(\theta)}_\rho = \chi^{GL(\theta)}_{\rho-\underline{1}}\chi^{GL(\theta)}_{\underline{1}}$. Then we note that $\chi^{GL(\theta)}_{\underline{1}}$ is the determinant character of $GL(\theta)$ (or $O(\theta)$, when restricted), and that $\chi^{O(\theta)}_{\underline{1}}\chi^{O(\theta)}_\lambda = \chi^{O(\theta)}_{\lambda'}$ (see, for example, the remark after Proposition 2.6 in [81]). Substituting into (5.73) completes the proof. ∎

The last lemma in this subsection gives us control of the coefficients $b^{n,\theta}_{\lambda,\rho}$ for certain values of ρ. In order to prove it, we need to introduce some more representation theory of the Brauer algebra.

The Brauer algebra's semisimplicity is dependent on the parameter θ. When θ is a positive integer, $\mathbb{B}_{n,\theta}$ is semisimple if and only if $\theta \geq n-1$ and when $\theta \notin \mathbb{Z}$, it is always semisimple. See [103], [87]. The Brauer algebra has indecomposable representations, known as the cell modules (see [24]), indexed by partitions $\lambda \vdash n-2k$, $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$. When $\mathbb{B}_{n,\theta}$ is semisimple, these are exactly the irreducibles. Their characters are described by Ram [84]. Note that Ram's results on the cell characters are stated for when the algebra is semisimple, but they extend to the case when it is not.

When $\mathbb{B}_{n,\theta}$ is not semisimple, the cell modules are not necessarily irreducible (in fact they are not even necessarily semisimple). The irreducible representation corresponding to λ is then a quotient of the cell module corresponding to λ. Let us denote the character of the cell module corresponding to λ by $\gamma^{\mathbb{B}_{n,\theta}}_\lambda$.

The restrictions of representations of $\mathbb{C}S_n$ to $\mathbb{C}S_{n-1}$ and $\mathbb{B}_{n,\theta}$ to $\mathbb{B}_{n-1,\theta}$ are well studied. Let $\rho \vdash n$, $\lambda \vdash n-2k$, $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$. We have the following, in terms of characters, from (2.25) and (2.26) (or, for example, Sections 4 and 5 (and Figures 1 and 2) from [32], and Proposition 1.3 from [78]):

$$\begin{aligned} \operatorname{res}^{S_n}_{S_{n-1}}[\chi^{S_n}_\rho] &= \sum_{\overline{\rho}=\rho-\square} \chi^{S_{n-1}}_{\overline{\rho}}; \\ \operatorname{res}^{\mathbb{B}_{n,\theta}}_{\mathbb{B}_{n-1,\theta}}[\gamma^{\mathbb{B}_{n,\theta}}_\lambda] &= \sum_{\overline{\lambda}=\lambda\pm\square} \gamma^{\mathbb{B}_{n-1,\theta}}_{\overline{\lambda}}; \end{aligned} \tag{5.74}$$

and if $\theta \geq 2$ is an integer and λ further satisfies $\lambda_1^{\intercal} + \lambda_2^{\intercal} \leq \theta$,

$$\operatorname{res}^{\mathbb{B}_{n,\theta}}_{\mathbb{B}_{n-1,\theta}}[\chi^{\mathbb{B}_{n,\theta}}_{\lambda}] = \sum_{\substack{\overline{\lambda}=\lambda\pm\square \\ \overline{\lambda}_1^{\intercal}+\overline{\lambda}_2^{\intercal}\leq\theta}} \chi^{\mathbb{B}_{n-1,\theta}}_{\overline{\lambda}}, \tag{5.75}$$

where in the first equality the sum is over all $\overline{\rho} \vdash n-1$ whose Young diagram can be obtained from that of ρ by removing a box; in the second the sum is over $\overline{\lambda} \vdash n-1-2r$, $0 \leq r \leq \lfloor \frac{n-1}{2} \rfloor$, whose Young diagram can be obtained from that of λ by removing or adding a box; and in the third the sum is the same as the second, except we are restricted to those $\overline{\lambda}$ with $\overline{\lambda}_1^{\intercal} + \overline{\lambda}_2^{\intercal} \leq \theta$.

We now describe how cell modules of $\mathbb{B}_{n,\theta}$ decompose when restricted to $\mathbb{C}S_n$. We call a partition π even if all its parts π_i are even. Let $\lambda \vdash n-2k$, $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$. Then from (2.37):

$$\operatorname{res}^{\mathbb{B}_{n,\theta}}_{S_n}[\gamma^{\mathbb{B}_{n,\theta}}_{\lambda}] = \sum_{\rho\vdash n} \tilde{b}^{n,\theta}_{\lambda,\rho}\chi^{S_n}_{\rho} = \chi^{S_{n-2k}}_{\lambda} \times \sum_{\substack{\pi\vdash 2k \\ \pi \text{ even}}} \chi^{S_{2k}}_{\pi}, \tag{5.76}$$

or,

$$\tilde{b}^{n,\theta}_{\lambda,\rho} = \sum_{\substack{\pi\vdash 2k \\ \pi \text{ even}}} c^{\rho}_{\lambda,\pi}.$$

Let us make a few useful remarks.

Remark 5.7.4. 1. Since the irreducible representation of $\mathbb{B}_{n,\theta}$ corresponding to λ is a quotient of the cell module corresponding to λ, we have $b^{n,\theta}_{\lambda,\rho} \leq \tilde{b}^{n,\theta}_{\lambda,\rho}$ for all λ, ρ.

2. Since $c^{\rho}_{\lambda,\pi}$ is determined by π and the skew-diagram $\rho \smallsetminus \lambda$, we have that $\tilde{b}^{n,\theta}_{\lambda,\rho}$ is fully determined by the skew-diagram $\rho \smallsetminus \lambda$.

3. If $\lambda \nsubseteq \rho$ then $c^{\rho}_{\lambda,\pi} = 0$, so as a consequence, $\tilde{b}^{n,\theta}_{\lambda,\rho} = 0$ (and therefore $b^{n,\theta}_{\lambda,\rho} = 0$) if $\lambda \nsubseteq \rho$.

4. Combining the above with Remark 5.7.3, we have that $b^{n,\theta}_{\lambda,\rho} = 0$ in the following cases: if $\lambda_j > \rho_j - \rho_\theta$ for $j \leq \lfloor \theta/2 \rfloor$, or if $\rho_j = \rho_\theta$, $j > \lfloor \theta/2 \rfloor$ with either ρ_θ odd, $\lambda_j = 0$, or ρ_θ even, $\lambda_j = 1$.

Lemma 5.7.5. *Let $(\lambda, \rho) \in \Lambda_n(\theta)$, such that $\rho_1^{\intercal} + \rho_2^{\intercal} \leq \theta + 1$. Then $b^{n,\theta}_{\lambda,\rho} = \tilde{b}^{n,\theta}_{\lambda,\rho}$.*

Proof. We work by induction on n. The base case, $n = 1$, is straightforward, since $\mathbb{B}_{1,\theta} = \mathbb{C}S_1$. Assume the theorem is proved for $n-1, n-2, \ldots$. Since $\rho_1^{\intercal} + \rho_2^{\intercal} \leq \theta + 1$, in almost all cases $\rho = \pi + \square$, (meaning the Young diagram of ρ can be obtained from a valid Young diagram $\pi \vdash n-1$ by adding a box), with $\pi_1^{\intercal} + \pi_2^{\intercal} \leq \theta$; the exception is the case where θ is odd, $\rho_1^{\intercal} = \rho_2^{\intercal} = (\theta+1)/2$, and $\rho_{(\theta+1)/2} \geq 3$. We will deal with this exceptional case second, and the former case now. Let

$$\operatorname{res}^{\mathbb{B}_{n,\theta}}_{S_{n-1}}[\chi^{\mathbb{B}_{n,\theta}}_{\lambda}] = \sum_{\pi\vdash n-1} \alpha^{n,\theta}_{\lambda,\pi}\chi^{S_{n-1}}_{\pi}, \qquad \operatorname{res}^{\mathbb{B}_{n,\theta}}_{S_{n-1}}[\gamma^{\mathbb{B}_{n,\theta}}_{\lambda}] = \sum_{\pi\vdash n} \tilde{\alpha}^{n,\theta}_{\lambda,\pi}\chi^{S_{n-1}}_{\pi}.$$

Note that in a similar way to part 1 of Remark 5.7.4, $\alpha^{n,\theta}_{\lambda,\pi} \leq \tilde{\alpha}^{n,\theta}_{\lambda,\pi}$ for all λ, π. Now fix a $\pi \vdash n-1$ with $\rho = \pi + \square$, with $\pi_1^{\intercal} + \pi_2^{\intercal} \leq \theta$. We will exploit the fact that there are two ways to restrict from the Brauer algebra $\mathbb{B}_{n,\theta}$ to $\mathbb{C}S_{n-1}$; either by restricting first to $\mathbb{B}_{n-1,\theta}$, or first to $\mathbb{C}S_n$. Formulaically, using (5.74) and (5.75), this reads:

$$\alpha^{n,\theta}_{\lambda,\pi} = \sum_{\substack{\overline{\lambda}=\lambda\pm\square \\ \overline{\lambda}^{\mathsf{T}}_1+\overline{\lambda}^{\mathsf{T}}_2\le\theta}} b^{n-1,\theta}_{\overline{\lambda},\pi} = \sum_{\overline{\pi}=\pi+\square} b^{n,\theta}_{\lambda,\overline{\pi}}, \tag{5.77}$$

and

$$\tilde{\alpha}^{n,\theta}_{\lambda,\pi} = \sum_{\overline{\lambda}=\lambda\pm\square} \tilde{b}^{n-1,\theta}_{\overline{\lambda},\pi} = \sum_{\overline{\pi}=\pi+\square} \tilde{b}^{n,\theta}_{\lambda,\overline{\pi}}. \tag{5.78}$$

Since $\pi^{\mathsf{T}}_1 + \pi^{\mathsf{T}}_2 \le \theta$, by part 3 of Remark 5.7.4, each $\overline{\lambda}$ with $\tilde{b}^{n-1,\theta}_{\overline{\lambda},\pi} > 0$ must also have $\overline{\lambda}^{\mathsf{T}}_1 + \overline{\lambda}^{\mathsf{T}}_2 \le \theta$. Now the central sums in equations (5.77) and (5.78) are sums over the same set of partitions $\overline{\lambda}$. Now by the inductive assumption, each $b^{n-1,\theta}_{\overline{\lambda},\pi} = \tilde{b}^{n-1,\theta}_{\overline{\lambda},\pi}$, which gives $\alpha^{n,\theta}_{\lambda,\pi} = \tilde{\alpha}^{n,\theta}_{\lambda,\pi}$. Equating the right hand terms in the equations (5.77) and (5.78), and recalling that $b^{n,\theta}_{\lambda,\overline{\pi}} \le \tilde{b}^{n,\theta}_{\lambda,\overline{\pi}}$, we must have the equality $b^{n,\theta}_{\lambda,\overline{\pi}} = \tilde{b}^{n,\theta}_{\lambda,\overline{\pi}}$, for each $\overline{\pi} = \pi + \square$. Since $\rho = \pi + \square$, we are done.

It remains to prove the lemma for the special case where θ is odd, $\rho^{\mathsf{T}}_1 = \rho^{\mathsf{T}}_2 = (\theta+1)/2$, and $\rho_{(\theta+1)/2} \ge 3$. Here, we let $\rho = \pi + \square$, where the differing square lies on row $(\theta+1)/2$. Now $\pi^{\mathsf{T}}_1 = \pi^{\mathsf{T}}_2 = (\theta+1)/2$. The equations (5.77) and (5.78) still hold, but now there exists one possible summand of the central sum in (5.78) where $\overline{\lambda}^{\mathsf{T}}_1 + \overline{\lambda}^{\mathsf{T}}_2 > \theta$. This summand appears in the case when $\lambda^{\mathsf{T}}_1 = (\theta+1)/2$, $\lambda^{\mathsf{T}}_2 = (\theta-1)/2$, and the summand itself is $\overline{\lambda}$, obtained by adding a box in row $(\theta+1)/2$ (column 2). In other instances of λ, we use the same method as the first part of the proof.

Now, again employing the inductive assumption on the terms in the central sums of (5.77) and (5.78), we have that $\alpha^{n,\theta}_{\lambda,\pi} + \tilde{b}^{n-1,\theta}_{\overline{\lambda},\pi} = \tilde{\alpha}^{n,\theta}_{\lambda,\pi}$, where λ and $\overline{\lambda}$ are the specific partitions described above. Plugging this into the right hand sides of (5.77) and (5.78), we have

$$\tilde{b}^{n-1,\theta}_{\overline{\lambda},\pi} + \sum_{\overline{\pi}=\pi+\square} b^{n,\theta}_{\lambda,\overline{\pi}} = \sum_{\overline{\pi}=\pi+\square} \tilde{b}^{n,\theta}_{\lambda,\overline{\pi}}. \tag{5.79}$$

Let π^* be π with one box added in row $(\theta+1)/2+1$ (column 1). Note that π^* is the only $\overline{\pi} = \pi + \square$ satisfying $\overline{\pi}^{\mathsf{T}}_1 + \overline{\pi}^{\mathsf{T}}_2 > \theta + 1$. We will prove that $b^{n,\theta}_{\lambda,\pi^*} + \tilde{b}^{n-1,\theta}_{\overline{\lambda},\pi} = \tilde{b}^{n,\theta}_{\lambda,\pi^*}$. Then (5.79) becomes

$$\sum_{\substack{\overline{\pi}=\pi+\square \\ \overline{\pi}\ne\pi^*}} b^{n,\theta}_{\lambda,\overline{\pi}} = \sum_{\substack{\overline{\pi}=\pi+\square \\ \overline{\pi}\ne\pi^*}} \tilde{b}^{n,\theta}_{\lambda,\overline{\pi}},$$

and similar to the first part of the proof, recalling $b^{n,\theta}_{\lambda,\overline{\pi}} \le \tilde{b}^{n,\theta}_{\lambda,\overline{\pi}}$ gives $b^{n,\theta}_{\lambda,\overline{\pi}} = \tilde{b}^{n,\theta}_{\lambda,\overline{\pi}}$ for all $\overline{\pi} = \pi + \square$, with $\overline{\pi}^{\mathsf{T}}_1 + \overline{\pi}^{\mathsf{T}}_2 \le \theta + 1$. This covers ρ. So, it remains to prove $b^{n,\theta}_{\lambda,\pi^*} + \tilde{b}^{n-1,\theta}_{\overline{\lambda},\pi} = \tilde{b}^{n,\theta}_{\lambda,\pi^*}$.

Now, Okada [81] gives an explicit algorithm for calculating $b^{n,\theta}_{\lambda,\pi^*}$. Working through that algorithm, we find that $b^{n,\theta}_{\lambda,\pi^*} = \tilde{b}^{n,\theta}_{\lambda,\pi^*} - \tilde{b}^{n,\theta}_{\hat{\lambda},\pi^*}$, where $\hat{\lambda}$ is obtained from λ by adding two boxes, one in each of the first two columns. Now it is straightforward to see that $\tilde{b}^{n,\theta}_{\hat{\lambda},\pi^*} = \tilde{b}^{n-1,\theta}_{\overline{\lambda},\pi}$, since $\pi^* \smallsetminus \hat{\lambda}$ and $\pi \smallsetminus \overline{\lambda}$ are identical skew-diagrams (remark 5.7.4). This completes the proof. ∎

We can now determine the sets $P_n(\theta)$ for $\theta = 2, 3$.

5.7.2 Spin $\frac{1}{2}$; $\theta = 2$

Recall $\Lambda_n(\theta)$ is the set of pairs of partitions (λ, ρ), $\lambda \vdash n-2k$, $0 \le k \le \lfloor \frac{n}{2} \rfloor$, $\rho \vdash n$, such that $\lambda_1^{\top} + \lambda_2^{\top} \le \theta$ and $\rho_1^{\top} \le \theta$. Recall the set $P_n(\theta)$ is given by $(\lambda,\rho) \in \Lambda_n(\theta)$ such that $b^{n,\theta}_{\lambda,\rho} > 0$, where $b^{n,\theta}_{\lambda,\rho}$ is the coefficient of the irreducible $\chi^{S_n}_{\rho}$ in the restriction of $\chi^{\mathbb{B}_{n,\theta}}_{\lambda}$ from $\mathbb{B}_{n,\theta}$ to $\mathbb{C}S_n$.

Proposition 5.7.6. *For $\theta = 2$, the $\mathbb{C}S_n$-$\mathbb{B}_{n,2}$ branching coefficient $b^{n,2}_{\lambda,\rho}$ is strictly positive if and only if $\lambda_1 \le \rho_1 - \rho_2$, with the exceptions of $\lambda = \varnothing$ or $\lambda = (1,1)$, in which case both rows of ρ must be even or odd, respectively. Hence $\frac{1}{n}P_n(2) \to \Delta_2^*$ in the Hausdorff distance, where*

$$\Delta_2^* = \{(x,y) \in ([0,1]^2)^2 \mid x_1 \ge x_2,\ x_1 + x_2 = 1,\ y_2 = 0,\quad 0 \le y_1 \le x_1 - x_2\}.$$

Proof. We prove first that the irreducible representation $\psi^{\mathbb{B}_{n,2}}_{(n-2k)}$ of $\mathbb{B}_{n,2}$ restricts to the symmetric group as:

$$\mathrm{res}^{\mathbb{B}_{n,2}}_{S_n}[\chi^{\mathbb{B}_{n,2}}_{(n-2k)}] = \sum_{i=0}^{k} \chi^{S_n}_{(n-i,i)}. \tag{5.80}$$

Indeed, by Remark 5.7.3 (using $(n-2k)' = (n-2k)$) and Lemma 5.7.5, we have $b^{n,2}_{(n-2k),(n-i,i)} = b^{n-2i,2}_{(n-2k),(n-2i)} = \tilde{b}^{n-2i,2}_{(n-2k),(n-2i)} = \mathbb{1}\{0 \le i \le k\}$, the last equality coming from (5.76) and the definition of the Littlewood-Richardson coefficients. Combining (5.80) and Okada's result in Remark 5.5.1 gives the first part of the proof. The second is a straightforward application of the definition of the Hausdorff distance. ∎

The proof of Proposition 5.7.6 also implies the following corollary.

Corollary 5.7.7. We have that for $\theta = 2$, $\lim_{n\to\infty} \frac{1}{n} \log\Big(\max_{(\lambda,\rho)\in\Lambda_n(2)} b^{n,2}_{\lambda,\rho}\Big) = 0$.

5.7.3 Spin 1; $\theta = 3$

Proposition 5.7.8. *For $\theta = 3$, the $\mathbb{C}S_n$-$\mathbb{B}_{n,3}$ branching coefficient $b^{n,3}_{\lambda,\rho}$ is strictly positive if and only if $\lambda_1 \le \rho_1 - \rho_3$, with the following exceptions:*

1. *If $\lambda = (n-2k)$, and $\rho_2 = \rho_3$ odd, or $\rho_1 = \rho_2$ odd, then $b^{n,3}_{\lambda,\rho} = 0$;*
2. *If $\lambda = (n-2k-1,1)$, and $\rho_2 = \rho_3$ even, or $\rho_1 = \rho_2$ even, then $b^{n,3}_{\lambda,\rho} = 0$;*
3. *If $\lambda = (1^j)$, $j = 0,\dots,3$, then $b^{n,3}_{\lambda,\rho} > 0$ if and only if ρ has j odd parts.*

As a consequence, $\frac{1}{n}P_n(3) \to \Delta_3^*$ in the Hausdorff distance, where

$$\Delta_3^* = \{(x,y) \in ([0,1]^3)^2 \mid x_1 \ge x_2 \ge x_3,\ x_1 + x_2 + x_3 = 1,\ y_2 = y_3 = 0,\quad 0 \le y_1 \le x_1 - x_3\}.$$

Proof. From Remark 5.7.3, we see that if $\lambda_1 > \rho_1 - \rho_3$ then $b^{n,3}_{\lambda,\rho} = 0$. For the rest of the first part of the Proposition, let $\lambda = (n-2k)$ or $(n-2k-1,1)$, and let $\lambda_1 \le \rho_1 - \rho_3$. Then, using Remark 5.7.3 and Lemma 5.7.5, $b^{n,3}_{\lambda,(\rho_1,\rho_2,\rho_3)} = b^{n-3\rho_3,3}_{\lambda^*,(\rho_1-\rho_3,\rho_2-\rho_3)} = \tilde{b}^{n-3\rho_3,3}_{\lambda^*,(\rho_1-\rho_3,\rho_2-\rho_3)}$, where $\lambda^* = \lambda$ if ρ_3 even, $\lambda^* = \lambda'$ if ρ_3 odd. The cases where $\lambda^* \nleq (\rho_1-\rho_3, \rho_2-\rho_3)$ (which give $b^{n,3}_{\lambda,(\rho_1,\rho_2,\rho_3)} = 0$) are the cases: $\lambda = (n-2k)$, $\rho_2 = \rho_3$ odd, and $\lambda = (n-2k-1,1)$, $\rho_2 = \rho_3$ even.

It remains to determine when $\tilde{b}^{n,3}_{\lambda,(\rho_1,\rho_2)}$ is non-zero. We need to show that if $\lambda \le \rho$, it is non-zero unless $\lambda = (n-2k)$, and $\rho_1 = \rho_2$ odd, or $\lambda = (n-2k-1,1)$, and $\rho_1 = \rho_2$ even. Recall the coeffiecient (from (5.76)) is given by

$$\tilde{b}^{n,3}_{\lambda,(\rho_1,\rho_2)} = \sum_{\substack{\tau \vdash 2k \\ \tau \text{ even}}} c^{(\rho_1,\rho_2)}_{\lambda,\tau}. \tag{5.81}$$

Let us prove the $\lambda = (n-2k)$ case. By the Littlewood-Richardson rule (or its special case the Pieri rule - see Section I.9 of Macdonald [68]), $\tilde{b}^{n,3}_{(n-2k),(\rho_1,\rho_2)}$ is equal to $|A|$, where A is the set of even partitions $\tau \vdash 2k$, $\tau \le \rho$, such that $\rho \setminus \tau$ is a skew diagram with no two boxes the same column. Wlog $\tau = (2k-2m, 2m)$. For $\tau \in A$, we must have $0 \le 2m \le \rho_2$, and $\rho_2 \le 2k-2m \le \rho_1$. (Note that we certainly have $2k \ge \rho_2$, which follows from $n-2k \le \rho_1$). Now the only case where no such τ exists is when $\rho_1 = \rho_2$ odd, since in this case, any even τ must give $\rho \setminus \tau$ with two boxes in the last column. The case $\lambda = (n-2k-1,1)$ is obtained in a similar way.

The third special case $\lambda = (1^j)$ is given by Okada [81] - see Remark 5.5.1.

The final part of the theorem now follows by applying the first part, and the definition of the Hausdorff distance. ∎

We also have the following corollary.

Corollary 5.7.9. We have that for $\theta = 3$, $\lim_{n\to\infty} \frac{1}{n} \log\Big(\max_{(\lambda,\rho)\in\Lambda_n(3)} b^{n,3}_{\lambda,\rho}\Big) = 0$.

Proof. By Lemma 5.7.2 and 5.7.5, each non-zero $b^{n,3}_{\lambda,\rho}$ is equal to some $\tilde{b}^{m,3}_{\lambda',\rho'}$, where $m \le n$, $(\lambda',\rho') \in \Lambda_m(3)$, $\rho' \le \rho$. Now $\tilde{b}^{m,3}_{\lambda',\rho'}$ is (from (5.76)):

$$\tilde{b}^{m,3}_{\lambda',\rho'} = \sum_{\substack{\tau \vdash 2j \\ \tau \text{ even}}} c^{\rho'}_{\lambda',\tau}.$$

Since ${\rho'}^{\mathsf{T}}_1 \le 3$, the number of $\tau \vdash 2j$ with $\tau \le \rho'$ is bounded by n^3. Then the Littlewood-Richardson coefficient $c^{\rho'}_{\lambda',\tau}$ is bounded by n^2, since $\lambda'_2, \lambda'_3 \le 1$, and $\lambda_j = 0$ for $j \ge 4$. Hence $\tilde{b}^{m,3}_{\lambda',\rho'}$ is bounded by n^{5+2}, which gives the result. ∎

5.8 Numerical proof of Remark 5.6.5

Recall the function

$$w(z) := \frac{\partial \phi}{\partial x_1}(z) = \frac{3}{2} + \frac{\log(z)(1+5z)}{4(1-z)} + \log\left(\frac{-z\log(z)}{3(1-z)+(1+z)\log(z)}\right).$$

We need to prove that this function $w(z)$ is positive in the range $(r,1)$, where r is the unique root of $3(1-z)+(1+z)\log(z)$ in $(0,1)$. This proof is due to Dave Platt.

Away from r and 1, this can be done straightforwardly using ARB, a C library for rigorous real and complex arithmetic (see `https://arblib.org/index.html`). We split the interval into small pieces, and use the program to show positivity on each piece. This works on the interval $[81714053/2^{30}, 1013243800/2^{30}]$. Near r, the function is large, and we can show by hand that it is positive. Indeed, $\log(z)(1+5z)/(4(1-z))$ and $\log(-z\log(z))$ are both increasing on the interval $[r, 81714053/2^{30}]$, and their sum, plus $3/2$, is easily bounded on the interval by 1 in magnitude. Then $-\log(3(1-z)+(1-z)\log(z))$ is decreasing on the interval, and its value at $81714053/2^{30}$ is far larger than 1.

It remains to show that $w(z)$ is positive on $[1013243800/2^{30}, 1]$. The function's first three derivatives are zero at 1, and the fourth is positive at 1. We use the argument principle, and compute the integral $w'(z)/(2\pi i w(z))$ along a circle centre 1 and radius 1/16. There are no poles within this circle, and there are four zeros at 1, so computing the integral to be 4 implies there are no more zeros within the circle. We use a double exponential quadrature technique due to Pascal Molin. This approximates the integral to a sum with an explicit error term. We use Theorem 3.10 from [75], with $D = 1$, $h = 0.15$ and $n = 91$, which, using ARB, gives the sum to be $[4.00000 \pm 5.24e-6] + [\pm 5.10e-6] * I$. The integral must be an integer by the argument principle, and $D = 1$ means the explicit error term is at most e^{-1}, hence the integral must equal 4.

5.9 Equivalence of $Q_{i,j}$ and $P_{i,j}$

In this second appendix we study a second representation of $\mathbb{B}_{n,\theta}$, which we'll prove is isomorphic to the representation $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ (3.12), for θ odd, and not isomorphic for θ even. Recall

$$\mathfrak{p}^{\mathbb{B}_{n,\theta}}(\overline{i,j}) = Q_{i,j}, \qquad \mathfrak{p}^{\mathbb{B}_{n,\theta}}(i,j) = T_{i,j}. \tag{5.82}$$

This will give the equivalence, in spin $S = 1$, between our model with Hamiltonian (5.20), and the bilinear-biquadratic Heisenberg model with Hamiltonian (5.29); equality of their partition functions was proved by Ueltschi ([101], Theorem 3.2).

Recall $\langle a_i, a_j | P_{i,j} | b_i, b_j \rangle = (-1)^{a_i - b_i} \delta_{a_i, -a_j} \delta_{b_i, -b_j}$. Define $\tilde{\mathfrak{p}}^{\mathbb{B}_{n,\theta}} : \mathbb{B}_{n,\theta} \to \mathrm{End}(\mathbb{V})$, given by

$$\tilde{\mathfrak{p}}^{\mathbb{B}_{n,\theta}}(\overline{i,j}) = P_{i,j}, \qquad \tilde{\mathfrak{p}}^{\mathbb{B}_{n,\theta}}(i,j) = T_{i,j}. \tag{5.83}$$

Lemma 5.9.1. *For θ odd, and all n, the representations $\tilde{\mathfrak{p}}^{\mathbb{B}_{n,\theta}}$ and $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ of $\mathbb{B}_{n,\theta}$ are isomorphic via a unitary transformation, and for θ even, the two are not isomorphic.*

Proof. Since the elements (i,j) and $(\overline{i,j})$ generate the algebra $\mathbb{B}_{n,\theta}$, it suffices to find an invertible linear function $\psi_n : \mathbb{V} \to \mathbb{V}$ such that

$$\psi_n^{-1} T_{i,j} \psi_n = T_{i,j}, \quad \psi_n^{-1} Q_{i,j} \psi_n = P_{i,j}, \tag{5.84}$$

for all i,j. By the Schur-Weyl duality for the general linear and symmetric groups (5.71), the first condition holds if and only if $\psi_n = \psi^{\otimes n}$ for some $\psi \in GL(\theta)$. Then the second condition also holds if and only if $(\psi^{\otimes 2})^{-1} Q_{i,j} \psi^{\otimes 2} = P_{i,j}$ for all i,j, which holds if and only if:

$$\begin{aligned}
(-1)^{a_i - b_i} \delta_{a_i, -a_j} \delta_{b_i, -b_j} &= \sum_{r_i, r_j, s_i, s_j} \psi_{a_i, r_i} \psi_{a_j, r_j} \delta_{r_i, r_j} \delta_{s_i, s_j} (\psi^{-1})_{s_i, b_i} (\psi^{-1})_{s_j, b_j} \\
&= \sum_{r,s} \psi_{a_i, r} \psi_{a_j, r} (\psi^{-1})_{s, b_i} (\psi^{-1})_{s, b_j} \\
&= (\psi \psi^{\mathsf{T}})_{a_i, a_j} ((\psi^{-1})^{\mathsf{T}} (\psi^{-1}))_{b_i, b_j}.
\end{aligned}$$

Hence the two representations are isomorphic if and only if the two are isomorphic if and only if there exists an invertible $\theta \times \theta$ matrix ψ such that

$$\psi\psi^{\mathsf{T}} = \begin{bmatrix} & & & & (-1)^{S} \\ & & & (-1)^{S-1} & \\ & & \iddots & & \\ & (-1)^{1-S} & & & \\ (-1)^{-S} & & & & \end{bmatrix},$$

and

$$(\psi^{-1})^{\mathsf{T}}\psi^{-1} = \begin{bmatrix} & & & & (-1)^{S} \\ & & & (-1)^{S-1} & \\ & & \iddots & & \\ & (-1)^{1-S} & & & \\ (-1)^{-S} & & & & \end{bmatrix}^{\mathsf{T}}.$$

where recall $\theta = 2S + 1$. For θ odd the two matrices on the right hand sides above are the same, so it suffices to note that we can set the central entry in ψ to be 1, and the rest to be made up of nested invertible 2x2 matrices g_1, g_2, given by, for example,

$$g_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & i \\ -1 & -i \end{bmatrix}, \quad g_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & i \\ 1 & i \end{bmatrix},$$

since

$$g_1 g_1^{\mathsf{T}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad g_2 g_2^{\mathsf{T}} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

This shows that for θ odd, the representations $\tilde{\mathfrak{p}}^{\mathbb{B}_{n,\theta}}$ and $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ are indeed isomorphic, and since g_1, g_1 are unitary, so is ψ_n. For θ even, there are fractional powers of (-1) appearing, so we have to make a choice, say, of $(-1)^{\frac{1}{2}} = \pm 1$, and then the rest of the entries are determined by $(-1)^{a} = (-1)^{a-\frac{1}{2}}(-1)^{\frac{1}{2}}$. Whichever we choose though, $\psi\psi^{\mathsf{T}}$ will always be a symmetric matrix, and

$$\begin{bmatrix} & & & & (-1)^{S} \\ & & & (-1)^{S-1} & \\ & & \iddots & & \\ & (-1)^{1-S} & & & \\ (-1)^{-S} & & & & \end{bmatrix}$$

will always be anti-symmetric (and non-zero), so the two cannot be equal. This concludes the proof.

∎

Chapter 6

Quantum Spin Systems on the complete bipartite graph

In this section we present the results of the paper "Heisenberg models and Schur–Weyl duality" [12], which is joint work with Jakob Björnberg and Hjalmar Rosengren. As in the previous section, we present the paper essentially unchanged, with references to our use of representation-theoretic tools from the previous sections as appropriate.

6.1 Introduction and results

When Werner Heisenberg in 1928 introduced his famous model for ferromagnetism, he described it in terms of an *exchange interaction* between neighbouring valence electrons ("Austausch von Elektronen" [56]). In modern notation, for the spin $S = \frac{1}{2}$ system he was considering, this interaction can be written as $T_{i,j} = 2(\boldsymbol{S}_i \cdot \boldsymbol{S}_j) + \frac{1}{2}$, where $T_{i,j}$ acts on a pure tensor $v_i \otimes v_j$ in $\mathbb{C}^2 \otimes \mathbb{C}^2$ by transposing the factors, $T_{i,j}(v_i \otimes v_j) = v_j \otimes v_i$, and $\boldsymbol{S} = (S^{(1)}, S^{(2)}, S^{(3)})$ are spin $S = \frac{1}{2}$-matrices. Two natural generalisations to higher spin immediately suggest themselves: we can take the interaction to be the transposition $T_{i,j}$ acting on $\mathbb{C}^\theta \otimes \mathbb{C}^\theta$, or to be a (positive multiple of) $\boldsymbol{S}_i \cdot \boldsymbol{S}_j$, where the $\boldsymbol{S}$ are now spin-S-matrices and $\theta = 2S + 1$. For $S > \frac{1}{2}$, these choices are no longer equivalent; while both are natural generalisations, some authors usually reserve the name Heisenberg model for the model with interaction $\boldsymbol{S}_i \cdot \boldsymbol{S}_j$. The model with interaction $T_{i,j}$ has been called the interchange model and is one of the main topics of this paper.

The name interchange model can be traced back to works by Harris [55], Powers [83], and Tóth [97], and is motivated by a probabilistic representation of the model. Powers [83] was first to notice that the ferromagnetic (spin-$\frac{1}{2}$) Heisenberg model can be represented in terms of a random walk on permutations generated by transpositions. The latter random walk was constructed on infinite lattices by Harris [55]. Tóth [97] was first to use this representation to obtain an important result for the Heisenberg model: a bound on the free energy of the model on $\mathbb{Z}^3$ that was the best known for many years [21]. The underlying random walk on permutations has come to be known as the interchange *process* in the literature on mixing times of Markov chains [2]. The present paper does not use the probabilistic representation, however; indeed our methods apply also in cases where such a representation is not available.

For the antiferromagnetic spin $S = \frac{1}{2}$ Heisenberg model, Aizenman and Nachtergaele [1] discovered a similar probabilistic representation based on the identity $P_{i,j} = \frac{1}{2} - 2\boldsymbol{S}_i \cdot \boldsymbol{S}_j$ where $P_{i,j}$ is (twice) the projection onto the singlet subspace of $\mathbb{C}^2 \otimes \mathbb{C}^2$ (eigenspace for the total spin operator with eigenvalue 0). On a bipartite graph, such as the line $\mathbb{Z}$ considered by Aizenman and Nachtergaele, the Hamiltonian with interactions $P_{i,j}$ is unitarily equivalent to that with interactions $Q_{i,j}$ defined by

$$\langle e_{\alpha_1} \otimes e_{\alpha_2} | Q_{i,j} | e_{\alpha_3} \otimes e_{\alpha_4} \rangle = \delta_{\alpha_1,\alpha_2} \delta_{\alpha_3,\alpha_4}, \tag{6.1}$$

where the e_α are a basis for $\mathbb{C}^2$. The interaction $Q_{i,j}$ has a natural interpretation in terms of random loops, and plays a central role in the present work. The definition (6.1) generalises straightforwardly to higher spin.

If we take the underlying lattice to be the complete graph K_n, consisting of n vertices with an edge between each pair of distinct vertices, then the interchange model is a mean-field system with Hamiltonian

$$-\frac{1}{n} \sum_{1 \le i < j \le n} T_{i,j}, \quad \text{acting on } V^{\otimes n} = (\mathbb{C}^\theta)^{\otimes n}, \theta \ge 2. \tag{6.2}$$

This model was studied in the papers [13, 14], where the key step of the analysis was to note that the Hamiltonian (6.2) is a central element of the group algebra $\mathbb{C}[S_n]$ of the symmetric group, represented on the tensor space $V^{\otimes n}$. This means that the eigenspace decomposition for the Hamiltonian (6.2) coincides with the decomposition of $V^{\otimes n}$ into irreducible S_n-modules, which is well-studied. Ryan [89] implemented a similar approach for the model with Hamiltonian

$$-\frac{1}{n} \sum_{1 \le i < j \le n} (a\, T_{i,j} + b\, Q_{i,j}) \quad \text{acting on } V^{\otimes n}, \tag{6.3}$$

with $a, b \in \mathbb{R}$ and $\theta \ge 2$, which can similarly be diagonalised using the irreducible representations of the Brauer algebra (defined below).

The unifying principle behind this approach is a classical algebraic theory called Schur–Weyl duality. This term is used for specific instances of a general result in representation theory called the double centraliser theorem, which states the following (see Theorem 3.0.1) [34, Theorem 4.54]. Let $\mathbb{V}$ be a finite-dimensional vector space, and $\mathbb{A} \subseteq \mathrm{End}(\mathbb{V})$ a semi-simple algebra of linear mappings (endomorphisms) $\mathbb{V} \to \mathbb{V}$. Then the centraliser $\mathbb{B}$ of $\mathbb{A}$, i.e. the algebra of endomorphisms commuting with all elements of $\mathbb{A}$, is also semi-simple, and as a representation of $\mathbb{A} \otimes \mathbb{B}$ we have

$$\mathbb{V} = \bigoplus_i U_i \otimes V_i, \tag{6.4}$$

where the U_i (respectively V_i) are an exhaustive list of non-isomorphic irreducible representations of $\mathbb{A}$ (respectively $\mathbb{B}$). The most famous instances of this (and relevant in the present work) are obtained by letting $\mathbb{V} = V^{\otimes n}$. If we let $\mathbb{A}$ consist of all invertible endomorphisms of $\mathbb{C}^\theta$, acting diagonally on $\mathbb{V}$, then $\mathbb{B}$ is generated by the permutations of the tensor factors of $\mathbb{V}$: this gives the Schur–Weyl duality between the general linear group $GL(\theta)$ and the symmetric group S_n (see (6.54) for details) which facilitates the analysis of the interchange model (6.2). If instead we take $\mathbb{A}$ to consist of orthogonal matrices, then $\mathbb{B}$ is the Brauer-algebra used in the analysis of (6.3).

Let us note that this work follows a line of papers analysing the interchange process and Heisenberg model with algebraic methods (including the aforementioned [13], [14], [89]). Alon and Kozma [3] analysed the interchange process on a general graph, and estimated the number of k-cycles at a given time; Berestycki and Kozma [9] gave an exact formula for the same on the complete graph; Alon and Kozma [4] gave an exact formula for the magnetisation of the mean-field spin $S = \frac{1}{2}$ Heisenberg model.

In this work we carry the methods described above further, to inhomogeneous models on the complete graph where the coupling constants between different vertices take finitely many different values. The models for which our analysis goes the deepest are what we call *two-block* models, where coupling constants can take at most three values (one each for the interactions within each of the two blocks, and one for interactions between the two blocks). Our results on these models come in three parts: we first compute in Theorems 6.1.1 and 6.1.2 the free energy; in Propositions 6.1.3 to 6.1.6, we give results on phase transitions, and, for certain values of the parameters, we compute a critical temperature; finally in Theorems 6.1.7 and 6.1.8 we give a magnetisation and limits of certain correlation functions. We then give the free energy for what we call *multi-block* models in Theorem 6.1.9, where coupling constants can take finitely many values, and we allow certain many-body interactions. Finally, in Section 6.1.5, we give heuristics for descriptions of the extremal Gibbs states for some of the models we study, and comment on their phase diagrams.

6.1.1 Two-block models: Free energy

For $a, b, c \in \mathbb{R}$, and $1 \le m \le n$, we define the AB-interchange-model, or *AB-model* for short, through its Hamiltonian

$$H_n^{\mathrm{AB}} = -\frac{1}{n}\Big(a \sum_{1\le i<j\le m} T_{i,j} + b \sum_{m+1\le i<j\le n} T_{i,j} + c \sum_{1\le i\le m<j\le n} T_{i,j}\Big). \tag{6.5}$$

For $\beta > 0$, introduce the partition function $Z_n^{\mathrm{AB}}(\beta) = \mathrm{Tr}\big[e^{-\beta H_n^{\mathrm{AB}}}\big]$. We call this a *two-block* model since we may think of it as a spin system on a graph with vertex set $\{1, 2, \dots, n\}$ partitioned into the two blocks $A = \{1, \dots, m\}$ and $B = \{m+1, \dots, n\}$. The form of the Hamiltonian (6.5) means that spins at two vertices within A interact with coupling constant a, spins at two vertices within B interact with coupling constant b, and the spin at a vertex in A interacts with the spin at a vertex in B with coupling constant c. In the homogeneous case $a = b = c$ we obtain the interchange model on the complete graph (6.2), while if $a = b = 0$ and $c \ne 0$ then we obtain a model on the complete bipartite graph $K_{m,n-m}$.

We write

$$F(x_1, \dots, x_\theta; y_1, \dots, y_\theta) = \textstyle\sum_{i=1}^{\theta} f(x_i, y_i), \tag{6.6}$$

where $x_i, y_i \ge 0$ and

$$f(x, y) = -x \log x - y \log y + \tfrac{\beta}{2}\big(ax^2 + by^2 + 2cxy\big). \tag{6.7}$$

We have the following result about the free energy:

Theorem 6.1.1. *Let $a, b, c \in \mathbb{R}$ be fixed. If $n, m \to \infty$ such that $m/n \to \rho \in (0,1)$, then the free energy of the model* (6.5) *satisfies*

$$\Phi_\beta^{AB}(a,b,c) := \lim_{n\to\infty} \tfrac{1}{n} \log Z_n^{AB}(\beta) = \max F(x_1,\dots,x_\theta; y_1,\dots,y_\theta) \tag{6.8}$$

where the maximum is taken over $x_1,\dots,x_\theta, y_1,\dots,y_\theta \geq 0$ subject to $\sum_{i=1}^\theta x_i = 1 - \sum_{i=1}^\theta y_i = \rho$.

Note that if $(x_1,\dots,x_\theta; y_1,\dots,y_\theta)$ is a maximum point of F, and we order the x-entries so that

$$x_1 \geq x_2 \geq \cdots \geq x_\theta, \tag{6.9}$$

then for $c > 0$ we necessarily have $y_1 \geq \cdots \geq y_\theta$, while for $c < 0$ we necessarily have $y_1 \leq \cdots \leq y_\theta$. Indeed, the only term in F which is dependent on the relative order of the entries is the term $\sum_{i=1}^\theta x_i y_i$, which is indeed maximised when the orders are the same and minimised if they are reversed.

We next consider another two-block model but where the interaction "between" the blocks uses the operator Q defined in (6.1). We let

$$H_n^{\text{WB}} = -\frac{1}{n}\Big(a \sum_{1\leq i<j\leq m} T_{i,j} + b \sum_{m+1\leq i<j\leq n} T_{i,j} + c \sum_{1\leq i\leq m<j\leq n} Q_{i,j} \Big). \tag{6.10}$$

Also let $Z_n^{\text{WB}}(\beta) = \mathrm{Tr}\big[e^{-\beta H_n^{\text{WB}}}\big]$. Let us note here that for all $\theta \geq 2$, this model is unitarily equivalent to the same model with each $Q_{i,j}$ replaced with $P_{i,j}$, the latter being (θ times) the projection onto the singlet state:

$$\langle e_{\alpha_1} \otimes e_{\alpha_2} | P_{i,j} | e_{\alpha_3} \otimes e_{\alpha_4} \rangle = (-1)^{\alpha_1 - \alpha_3} \delta_{\alpha_1, -\alpha_2} \delta_{\alpha_3, -\alpha_4}. \tag{6.11}$$

(Here we index the basis e_α for $\mathbb{C}^\theta$ with $\alpha \in \{-S, -S+1, \dots, S\}$ where $S = (\theta-1)/2$.) Indeed, for the model with $a = b = 0$ and $c > 0$ the equivalence of partition functions was proved by Aizenman and Nachtergaele in [1]; we give an algebraic proof for general $a, b, c \in \mathbb{R}$ in Lemma 6.7.1. We use the notation WB for this model as its analysis is based on the representation theory of the walled Brauer algebra, see Section 6.2.2. Interestingly, this model has the exact same free energy as the two-block interchange model:

Theorem 6.1.2. *Let $a, b, c \in \mathbb{R}$ be fixed. If $n, m \to \infty$ such that $m/n \to \rho \in (0,1)$, then the free energy of the model* (6.5) *satisfies*

$$\Phi_\beta^{WB}(a,b,c) := \lim_{n\to\infty} \tfrac{1}{n} \log Z_n^{\text{WB}}(\beta) = \Phi_\beta^{AB}(a,b,c), \tag{6.12}$$

where $\Phi_\beta^{AB}(a,b,c)$ is given in Theorem 6.1.1.

In the case $\theta = 2$, Theorem 6.1.2 can be deduced from Theorem 6.1.1 in the following elementary manner. For $\theta = 2$ we have [101, Section 7.1]

$$T_{i,j} = 2(\boldsymbol{S}_i \cdot \boldsymbol{S}_j) + \tfrac{1}{2}, \qquad Q_{i,j} = 2(S_i^{(1)} S_j^{(1)} - S_i^{(2)} S_j^{(2)} + S_i^{(3)} S_j^{(3)}) + \tfrac{1}{2}. \tag{6.13}$$

Letting $W = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ we have that $W_j^{-1} T_{i,j} W_j = -Q_{i,j} + 1$, so conjugating $H_n^{\mathrm{AB}}(a,b,-c)$ with $\prod_{j=m+1}^{n} W_j$ gives $H_n^{\mathrm{WB}}(a,b,c) - cm(n-m)/n$. Thus $\Phi_\beta^{\mathrm{WB}}(a,b,c) = \Phi_\beta^{\mathrm{AB}}(a,b,-c) + c\rho(1-\rho)$. This is consistent with Theorem 6.1.2 since (indicating the dependence on c with a subscript) $F_c(x_1,x_2;y_1,y_2) - F_{-c}(x_1,x_2;y_2,y_1) = c(x_1+x_2)(y_1+y_2) = c\rho(1-\rho)$, meaning that by Theorem 6.1.1 we have $\Phi_\beta^{\mathrm{AB}}(a,b,-c) + c\rho(1-\rho) = \Phi_\beta^{\mathrm{AB}}(a,b,c)$. However, for general θ the rank of $T_{i,j}$ is $\theta(\theta+1)/2$ while the rank of $Q_{i,j}$ is 1, so when $\theta > 2$, conjugating $T_{i,j}$ cannot give a linear combination of $Q_{i,j}$ and the identity.

6.1.2 Two-block models: Phase transition and critical temperature

Next we discuss phase transitions as β is varied, via the maximiser of the function F. Essentially, when a transition is present, we expect the maximiser of F to be fixed (at ω_0 (6.16)) for small β, and then at some critical β_{c} to begin to move. This β_{c} then corresponds to a point of phase transition in the model. For $\beta = \beta_{\mathrm{c}}$ it can happen either that ω_0 is unique or that there are other maximum points. We will see that the phase-transition is also reflected in the behavior of observables (Theorem 6.1.7) and the magnetisation (Theorem 6.1.8).

In Proposition 6.1.3, we characterise completely the values of a,b,c for which there exists such a phase transition. When it exists, finding explicit formulae for β_{c} seems difficult in general; we can do it in two cases, firstly in Proposition 6.1.4 when $\theta = 2$ (that is, spin $S = \frac{1}{2}$), and secondly in Proposition 6.1.5 when $c \geq 0$, $\theta \geq 3$ and

$$(a-c)\rho = (b-c)(1-\rho). \tag{6.14}$$

In the latter case, we further prove in Proposition 6.1.6 that for $\beta_{\mathrm{c}} < \beta < \beta_{\mathrm{c}} + \varepsilon$ and $\varepsilon > 0$ small, the maximiser of F is unique.

In what follows, we write $\vec{x} = (x_1,\dots,x_\theta)$, $\vec{y} = (y_1,\dots,y_\theta)$, and

$$\Omega = \left\{(\vec{x};\vec{y}) : x_1,\dots,x_\theta,y_1,\dots,y_\theta \geq 0,\ \textstyle\sum_{i=1}^{\theta} x_i = 1 - \sum_{i=1}^{\theta} y_i = \rho\right\}. \tag{6.15}$$

Elements of Ω will typically be denoted $\omega = (\vec{x};\vec{y})$. We write

$$\omega_0 = \left(\tfrac{\rho}{\theta}, \tfrac{\rho}{\theta}, \dots, \tfrac{\rho}{\theta}; \tfrac{1-\rho}{\theta}, \tfrac{1-\rho}{\theta}, \dots, \tfrac{1-\rho}{\theta}\right) \in \partial\Omega, \tag{6.16}$$

and we write $Q(x,y) = \frac{1}{2}(ax^2 + by^2 + 2cxy)$ for the quadratic form appearing in the function $f(x,y)$.

Proposition 6.1.3. *If Q is negative semidefinite, that is,*

$$a \leq 0, \quad b \leq 0, \quad \text{and} \quad ab \geq c^2, \tag{6.17}$$

then F assumes it maximum value at ω_0 for all $\beta > 0$. Otherwise, there exists a number $\beta_{\mathrm{c}} > 0$ such that F assumes it maximum value at ω_0 if and only if $0 < \beta \leq \beta_{\mathrm{c}}$, and this maximum is unique if $0 < \beta < \beta_{\mathrm{c}}$.

Let us write $\beta_{\mathrm{c}}(\theta)$ to highlight the dependence on θ. The next proposition gives $\beta_{\mathrm{c}}(2)$ when it exists.

Proposition 6.1.4. *Let $\theta = 2$ and assume that Q is not negative semidefinite, so that β_c exists. Then*

$$\beta_c = \beta_c(2) := \begin{cases} \dfrac{\rho a + (1-\rho)b - \sqrt{(\rho a - (1-\rho)b)^2 + 4\rho(1-\rho)c^2}}{\rho(1-\rho)(ab - c^2)}, & ab \neq c^2, \\ \dfrac{2}{a\rho + b(1-\rho)}, & ab = c^2. \end{cases} \tag{6.18}$$

Moreover, for $\beta = \beta_c$, ω_0 is the unique maximum point.

In the homogeneous spin $S = \frac{1}{2}$ AB-model, i.e. $\theta = 2$ and $a = b = c = 1$, we recover the critical point $\beta_c = 2$ first identified by Tóth [96] and by Penrose [82]. In the bipartite case $a = b = 0$ we get the critical value $\beta_c = 2/\sqrt{c^2\rho(1-\rho)}$; this has, to the best of our knowledge, not appeared previously in the literature.

The next proposition gives $\beta_c(\theta)$, $\theta \geq 3$ in the special case that $c \geq 0$ and (6.14) holds.

Proposition 6.1.5. *Suppose that $(a-c)\rho = (b-c)(1-\rho)$ as in (6.14) and let t denote either side of that identity. Suppose also that $c \geq 0$, that Q is not negative semidefinite so that β_c exists, and that $\theta \geq 3$. Then*

$$\beta_c = \beta_c(\theta) := \frac{2(\theta-1)\log(\theta-1)}{(\theta-2)(c+t)}. \tag{6.19}$$

Moreover, if $\beta = \beta_c$ there are exactly two maximum points satisfying (6.9), namely ω_0 of (6.16) and $\omega_1 = (\vec{x};\vec{y})$ given by

$$x_1 = \tfrac{(\theta-1)\rho}{\theta}, \quad x_2 = \cdots = x_\theta = \tfrac{\rho}{\theta(\theta-1)}, \tag{6.20a}$$

$$y_1 = \tfrac{(\theta-1)(1-\rho)}{\theta}, \quad y_2 = \cdots = y_\theta = \tfrac{1-\rho}{\theta(\theta-1)}. \tag{6.20b}$$

For all $\theta \geq 2$ we expect the maximum point to be unique for all β (subject to (6.9)), except possibly at $\beta = \beta_c$. For $\beta > \beta_c$ we can prove this under the conditions in Proposition 6.1.5 and for β close to the critical point (see also Proposition 6.5.1 for another special case).

Proposition 6.1.6. *Under the assumptions of Proposition 6.1.5, there exists $\varepsilon > 0$ such that, if $\beta_c < \beta < \beta_c + \varepsilon$, there is a unique maximiser of F in Ω with entries ordered as in (6.9). Moreover as $\beta \searrow \beta_c$, this maximiser tends to ω_1 given in (6.20).*

6.1.3 Two-block models: Correlations and magnetisation

We next move on to results about correlations which extend [14, Theorem 2.3]. To state them, introduce the function

$$R(w_1, \dots, w_\theta; z_1, \dots, z_\theta) = \det\left[e^{w_i z_j}\right]_{i,j=1}^{\theta} \prod_{1 \leq i < j \leq \theta} \frac{j-i}{(w_i - w_j)(z_i - z_j)}. \tag{6.21}$$

For $\# \in \{\text{AB}, \text{WB}\}$, we write

$$\langle \mathcal{O} \rangle_{\beta,n}^{\#} = \frac{\mathrm{Tr}_{\mathbb{V}}\left[\mathcal{O} e^{-\beta H_n^{\#}}\right]}{Z_n^{\#}(\beta)} \tag{6.22}$$

for the usual equilibrium state expectation of a linear operator $\mathcal{O}$ on $\mathbb{V}$.

Theorem 6.1.7. *Let $a, b, c \in \mathbb{R}$ and $\beta > 0$ be such that F has a unique maximum point $\omega^\star = (\vec{x}^\star; \vec{y}^\star)$ satisfying (6.9). Let W be an $\theta \times \theta$ matrix with eigenvalues $w_1, \dots, w_\theta \in \mathbb{C}$. As $n, m \to \infty$ such that $m/n \to \rho \in (0,1)$, we have that*

$$\begin{aligned}
&\lim_{n\to\infty} \big\langle \exp\big\{\tfrac{1}{n}\textstyle\sum_{i=1}^n W_i\big\}\big\rangle_{\beta,n}^{AB} = R(w_1,\dots,w_\theta; z_1^\star,\dots,z_\theta^\star)\\
&\lim_{n\to\infty} \big\langle \exp\big\{\tfrac{1}{n}\big(\textstyle\sum_{i=1}^m W_i - \sum_{i=m+1}^n W_i^\top\big)\big\}\big\rangle_{\beta,n}^{WB} = R(w_1,\dots,w_\theta; z_1^\dagger,\dots,z_\theta^\dagger),
\end{aligned} \tag{6.23}$$

where the superscript $^\top$ denotes transpose, and

$$z_j^\star = x_j^\star + y_j^\star, \qquad z_j^\dagger = x_j^\star - y_j^\star. \tag{6.24}$$

As a concrete example, for $W = h\,\mathrm{diag}(0,1,2,\dots,\theta-1)$ we have

$$R(w_1,\dots,w_\theta; z_1,\dots,z_\theta) = \prod_{1\le i<j\le\theta} \frac{e^{hz_i} - e^{hz_j}}{h(z_i - z_j)}. \tag{6.25}$$

The phase-transition at β_c is reflected in the fact that $R \equiv 1$ when $\omega^\star = \omega_0$, while R is non-trivial if the entries of $\vec{z}$ are non-constant. The latter occurs e.g. in the AB-model for $\beta > \beta_c$.

For a second concrete example, let $c > 0$. We will prove in Proposition 6.5.1 that any maximiser $(\vec{x}^\star; \vec{y}^\star)$ of F satisfying (6.9) is then of the form

$$x_1^\star \ge x_2^\star = \cdots = x_\theta^\star, \qquad y_1^\star \ge y_2^\star = \cdots = y_\theta^\star, \tag{6.26}$$

in which case $z^\star$ (6.24) will be of the same form. Letting W be an arbitrary rank 1 projection, with eigenvalues $1, 0, \dots, 0$, and writing $u^\star = z_1^\star - z_2^\star$, we have

$$\lim_{n\to\infty} \big\langle \exp\big\{\tfrac{1}{n}\textstyle\sum_{i=1}^n W_i\big\}\big\rangle_{\beta,n}^{\mathrm{AB}} = \tfrac{(2S)!}{(hu^\star)^{2S}} e^{\frac{h}{2S+1}(1-u^\star)} \textstyle\sum_{j=2S}^\infty \tfrac{(hu^\star)^j}{j!}. \tag{6.27}$$

(The calculation of R is performed in [14, Section 6].)

Theorem 6.1.7 also shows that the AB- and WB-models are not equivalent, despite having the same free energy (for any anti-symmetric matrix W, the observables on the left in (6.23) are the same, while their limiting expectations are different). The result is also relevant for understanding extremal states, see Section 6.1.5.

Finally we have the following result about the (thermodynamic) magnetisation. Let W be an $\theta \times \theta$ matrix with real eigenvalues $w_1 \ge \cdots \ge w_\theta$, let $h \in \mathbb{R}$, and write

$$Z_n^{\mathrm{AB}}(\beta, h) = \mathrm{Tr}_{\mathbb{V}}[\exp\big(-\beta H_n^{\mathrm{AB}} + h \textstyle\sum_{1\le i\le n} W_i\big)], \tag{6.28}$$

and let

$$Z_n^{\mathrm{WB}}(\beta, h) = \mathrm{Tr}_{\mathbb{V}}[\exp\big(-\beta H_n^{\mathrm{WB}} + h\big(\textstyle\sum_{1\le i\le m} W_i - \sum_{m<i\le n} W_i^\top\big)\big)]. \tag{6.29}$$

In Theorem 6.2.4 we will obtain explicit expressions for the limits

$$\Phi_{\beta,h}^{\#}(a, b, c, \vec{w}) \coloneqq \lim_{n\to\infty} \tfrac{1}{n} \log Z_n^{\#}(\beta, h) \tag{6.30}$$

where $\# \in \{\textsc{ab}, \textsc{wb}\}$ (this turns out to depend on W only through its spectrum $\vec{w}$). The magnetisation is given by the left and right derivatives of this free energy with respect to h, at $h = 0$.

Theorem 6.1.8. *Let $a, b, c \in \mathbb{R}$ and $w_1 \geq \dots \geq w_\theta$ be fixed. Let $\Phi^{AB}(\beta, h) = \Phi^{AB}_{\beta,h}(a, b, c, \vec{w})$ and $\Phi^{WB}(\beta, h) = \Phi^{WB}_{\beta,h}(a, b, c, \vec{w})$, regarded as functions of β and h. Then*

$$\begin{aligned} \frac{\partial \Phi^{AB}}{\partial h}\Big|_{h\downarrow 0} &= \max_{(\vec{x}^\star;\vec{y}^\star)} \sum_{i=1}^{\theta} z_i^\star w_i, & \frac{\partial \Phi^{AB}}{\partial h}\Big|_{h\uparrow 0} &= \min_{(\vec{x}^\star;\vec{y}^\star)} \sum_{i=1}^{\theta} z_i^\star w_{\theta+1-i}, \\ \frac{\partial \Phi^{WB}}{\partial h}\Big|_{h\downarrow 0} &= \max_{(\vec{x}^\star;\vec{y}^\star)} \sum_{i=1}^{\theta} z_i^\star w_i, & \frac{\partial \Phi^{WB}}{\partial h}\Big|_{h\uparrow 0} &= \min_{(\vec{x}^\star;\vec{y}^\star)} \sum_{i=1}^{\theta} z_i^\star w_{\theta+1-i}, \end{aligned} \tag{6.31}$$

where the maxima and minima are over all maximisers $(\vec{x}^\star;\vec{y}^\star)$ of $F(\vec{x};\vec{y})$, with entries ordered decreasing, and $z_1^\star, \dots, z_\theta^\star$ are the following values arranged in decreasing order:

- *for $c > 0$, in the* AB*-case $z_i^\star = x_i^\star + y_i^\star$ and in the* WB*-case $z_i^\star = x_i^\star - y_i^\star$;*
- *for $c < 0$, in the* AB*-case $z_i^\star = x_i^\star + y_{\theta+1-i}^\star$ and in the* WB*-case $z_i^\star = x_i^\star - y_{\theta+1-i}^\star$.*

It is natural to take W to have trace zero. Then, from Proposition 6.1.3, for all $\beta < \beta_c$ the only maximiser is ω_0 (6.16) and we have

$$\frac{\partial \Phi}{\partial h}\Big|_{h\downarrow 0} = \frac{\partial \Phi}{\partial h}\Big|_{h\uparrow 0} = 0, \tag{6.32}$$

for both AB- and WB-models and for both $c > 0$ and $c < 0$. This holds also for $\beta = \beta_c$ when $\theta = 2$.

Let us discuss the case $\theta \geq 3$ in Proposition 6.1.5 at $\beta = \beta_c$. Recall that $c \geq 0$ in this case. Calculations with the point ω_1 (6.20) give the following:

- In the AB-case, at ω_1 the values

  $$z_1 = \tfrac{\theta-1}{\theta}, \quad z_2 = \dots = z_r = \tfrac{1}{\theta(\theta-1)} \tag{6.33}$$

  are already decreasing. This gives

  $$\frac{\partial \Phi^{AB}}{\partial h}\Big|_{h\downarrow 0} = \max\big\{0, \tfrac{\theta-2}{\theta-1} w_1\big\} \qquad \frac{\partial \Phi^{AB}}{\partial h}\Big|_{h\uparrow 0} = \min\big\{0, \tfrac{\theta-2}{\theta-1} w_\theta\big\}. \tag{6.34}$$

  For non-trivial W we have $w_1 > 0 > w_\theta$, thus the magnetisation is discontinuous at the point of phase-transition.

- In the WB-case, at ω_1 the ordering of the values $x_i - y_i$ depends on ρ. If $\rho > \frac{1}{2}$ we get

  $$z_1 = (2\rho - 1)\tfrac{\theta-1}{\theta}, \quad z_2 = \dots = z_\theta = \tfrac{2\rho-1}{\theta(\theta-1)}, \tag{6.35}$$

  and from there

  $$\begin{aligned} \frac{\partial \Phi^{WB}}{\partial h}\Big|_{h\downarrow 0} &= \max\big\{0, (2\rho-1)\tfrac{\theta-2}{\theta-1} w_1\big\} \\ \frac{\partial \Phi^{WB}}{\partial h}\Big|_{h\uparrow 0} &= \min\big\{0, (2\rho-1)\tfrac{\theta-2}{\theta-1} w_\theta\big\}. \end{aligned} \tag{6.36}$$

For non-trivial W, this gives a discontinuous magnetisation. In the case $\rho < \frac{1}{2}$, the magnetisation is obtained by exchanging w_1 and w_θ in the latter expressions. For $\rho = \frac{1}{2}$, the magnetisation is continuous at the point of phase-transition.

6.1.4 Multi-block models

We generalize the free energy calculation of Theorem 6.1.1 to a class of models with $p \geq 1$ blocks rather than just the two blocks A and B, and with certain many-body interactions.

We first need some notation. For $\sigma \in S_n$ a permutation of $1, 2, \dots, n$, let T_σ be the linear operator on $\mathbb{V} = V^{\otimes n}$ which permutes the tensor factors according to σ:

$$T_\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(n)}. \tag{6.37}$$

(The mapping T is a representation of S_n - it is the map $\mathfrak{p}^{S_n}$ from (3.8); we use the notation T for the rest of this section.) Let γ be a partition with all parts > 1, that is $\gamma = (\gamma_1, \dots, \gamma_\ell)$ is a sequence of integers $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_\ell \geq 2$. We say that a permutation $\sigma \in S_n$ has cycle-type γ if its non-trivial cycles, ordered from longest to shortest, have lengths $\gamma_1, \dots, \gamma_\ell$. Then $|\gamma| := \gamma_1 + \cdots + \gamma_\ell \leq n$. Let C_n^γ be the set of permutations in S_n with cycle-type γ; this is a conjugacy-class of S_n. For example, if $\gamma = (2)$ then $C_n^\gamma = C_n^{(2)}$ is the set of transpositions in S_n, and if $\gamma = (3)$ then $C_n^\gamma = C_n^{(3)}$ is the set of three-cycles in S_n. Similarly, for $A \subseteq \{1, 2, \dots, n\}$, let C_A^γ denote the set of permutations of the elements of A with cycle-type γ.

Let $A_1, \dots, A_p$ form a partition of $\{1, \dots, n\}$ with $|A_k| = m_k$. Fix a finite set Γ of partitions γ with all parts > 1. We assume that n and all m_k are large enough that $C_n^\gamma \neq \varnothing$ and $C_{A_k}^\gamma \neq \varnothing$ for all $\gamma \in \Gamma$. For $a_1^\gamma, \dots, a_p^\gamma, c^\gamma \in \mathbb{R}$, consider the Hamiltonian

$$H_n^{\mathrm{MB}} = -n \sum_{\gamma \in \Gamma} \Big(\sum_{k=1}^{p} \frac{a_k^\gamma}{|C_{A_k}^\gamma|} \sum_{\sigma \in C_{A_k}^\gamma} T_\sigma + \frac{c^\gamma}{|C_n^\gamma|} \sum_{\sigma \in C_n^\gamma} T_\sigma \Big), \tag{6.38}$$

and the partition function $Z_n^{\mathrm{MB}}(\beta) = \mathrm{Tr}_{\mathbb{V}}[e^{-\beta H_n^{\mathrm{MB}}}]$. Note that we have the scaling factor n in front of (6.38) rather than $\frac{1}{n}$ as in (6.5). This is because the sizes of the conjugacy classes C_A^γ depend on n, for example for transpositions we have $|C_n^{(2)}| = \binom{n}{2}$.

The form of the Hamiltonian (6.38) means that spins at vertices in each block A_k interact with other with the many-body interaction T_σ (as opposed to the pair-interaction $T_{i,j} = T_{(i,j)}$ before), with strength constants a_k^γ dependent on the cycle type γ of σ; as well as this, spins in all blocks together interact with each other similarly, this time with strength constants c^γ.

The operators T_σ appearing in (6.38) may all be written in terms of spin-matrices. Indeed, for transpositions $\sigma = (i, j)$ this was discussed above, and for general σ we may write T_σ as a product of $T_{i,j}$'s. However, we do not pursue an explicit formula for T_σ in terms of spin-matrices.

Our result about the free energy of this model is most compactly expressed in terms of positive semidefinite Hermitian $\theta \times \theta$ matrices X. For such a matrix, having eigenvalues

$x_1, \dots, x_\theta \geq 0$, we use the von Neuman entropy

$$S(X) = -\operatorname{Tr}[X \log X] = -\sum_{i=1}^{\theta} x_i \log x_i. \tag{6.39}$$

We have the following:

Theorem 6.1.9. *Let $p \geq 1$ be fixed, and suppose that for all $k = 1, \dots, p$ we have that $m_k/n \to \rho_k \in (0,1)$ as $n \to \infty$. For the Hamiltonian (6.38), we have that the free energy is given by*

$$\lim_{n\to\infty} \tfrac{1}{n} \log Z_n^{MB}(\beta) = \max \phi_\beta(X_1, \dots, X_p), \tag{6.40}$$

where the maximum is taken over all positive semidefinite Hermitian $\theta \times \theta$ matrices $X_1, \dots, X_p$ with $\operatorname{Tr}[X_k] = \rho_k$, and where

$$\begin{aligned} \phi_\beta(X_1, \dots, X_p) &= \sum_{k=1}^{p} S(X_k) \\ &\quad + \beta \sum_{\gamma \in \Gamma} \Big(\sum_{k=1}^{p} a_k^\gamma \prod_{j\geq 1} \operatorname{Tr}\big[X_k^{\gamma_j}\big] + c^\gamma \prod_{j\geq 1} \operatorname{Tr}[(X_1 + \cdots + X_p)^{\gamma_j}] \Big). \end{aligned} \tag{6.41}$$

Let us now discuss a few specializations of Theorem 6.1.9. If we set $p = 2$, $\Gamma = \{(2)\}$ and $a_1^{(2)} = (a-c)/2$, $a_2^{(2)} = (b-c)/2$ and $c^{(2)} = c/2$, then

$$\phi_\beta(X_1, X_2) = S(X_1) + S(X_2) + \tfrac{\beta}{2} \operatorname{Tr}\big[aX_1^2 + bX_2^2 + 2cX_1X_2\big]. \tag{6.42}$$

In fact, in this case we recover Theorem 6.1.1, i.e. we have $\max \phi_\beta(X_1, X_2) = \Phi_\beta^{\mathrm{AB}}(a, b, c)$. For details, see the discussion around (6.79).

If instead we set $p = 1$ and all $a_k^\gamma = 0$ then (6.38) becomes

$$H_n^{\mathrm{MB}} = -n \sum_{\gamma\in\Gamma} \frac{c^\gamma}{|C_n^\gamma|} \sum_{\sigma \in C_n^\gamma} T_\sigma. \tag{6.43}$$

We thus obtain a homogeneous model of many-body interaction on the complete graph K_n. (In fact, (6.43) is the image of a general central element of $\mathbb{C}[S_n]$ under the representation T.) In this case we get that

$$\tfrac{1}{n} \log Z_{\beta,n}^{\mathrm{MB}} \to \max\Big(-\sum_{i=1}^{\theta} x_i \log x_i + \beta \sum_{\gamma\in\Gamma} c^\gamma p_\gamma(x_1, \dots, x_\theta) \Big), \tag{6.44}$$

where the maximum is over all $x_1, \dots, x_\theta$ satisfying $x_i \geq 0$ and $\sum_{i=1}^{\theta} x_i = 1$, and where $p_\gamma(x_1, \dots, x_\theta)$ denotes the power-sum symmetric polynomial

$$p_\gamma(x_1, \dots, x_\theta) = \prod_{j\geq 1}^{\ell} (x_1^{\gamma_j} + \cdots + x_\theta^{\gamma_j}). \tag{6.45}$$

It seems likely that Theorems 6.1.7 and 6.1.8 can be extended to multi-block cases, though we do not pursue such extensions here.

6.1.5 Heuristics for phase diagrams and extremal Gibbs states

In [14], for several models including the interchange model (6.2), the authors give a heuristic argument which points towards the structure of the set Ψ_β of extremal Gibbs states at inverse temperature β. The extremal Gibbs states in infinite volume are not well-defined on the complete graph, so the working is by analogy. Specifically, their heuristics lead to two expected equalities: first that

$$\lim_{\Lambda\to\mathbb{Z}^d} \langle e^{\frac{h}{n}\sum_i W_i}\rangle_{\beta,\Lambda} = \int_{\Psi_\beta} e^{h\langle W_0\rangle_\psi} d\mu(\psi), \tag{6.46}$$

for $\theta\times\theta$ matrices W, where $\langle\cdot\rangle_\psi$ is an extremal Gibbs state, $d\mu$ is the measure on Ψ_β corresponding to the symmetric Gibbs state, W_0 is the operator W at the lattice site 0, and the left hand side is the limit of successively larger boxes $\Lambda\in\mathbb{Z}^d$; second that

$$\lim_{n\to\infty} \langle e^{\frac{h}{n}\sum_i W_i}\rangle_{\beta,n} = \lim_{\Lambda\to\mathbb{Z}^d} \langle e^{\frac{h}{n}\sum_i W_i}\rangle_{\beta,\Lambda}, \tag{6.47}$$

where the left hand term is the observable on the complete graph. The left hand side of (6.47) is computed rigorously on the complete graph, and then, with the expected structure of Ψ_β inserted, the right hand side of (6.46) is rigorously computed, and the two are shown to be the same. This working is not a proof either of the expected equalities (6.46), (6.47) or of the expected structure of Ψ_β, but it points towards all three statements holding true. Using the results of the present paper, we can provide the same calculations and heuristics for the interchange model and for the nematic model in spin $S = 1$, this time on the complete bipartite graph.

The interchange model on the complete bipartite graph is exactly our AB model with parameters $a = b = 0$. For $c > 0$, Proposition 6.1.3 shows that this model has a phase transition. At low temperatures, the model is expected to have extremal Gibbs states labelled by $\mathbb{CP}^{\theta-1}$, rank 1 projections in $\mathbb{C}^\theta$. With $\langle W_0\rangle_\psi = u^\star$, when W is a rank 1 projection and assuming that Ψ_β is indeed given by $\mathbb{CP}^{\theta-1}$, the right hand side of (6.46) is given by

$$\frac{(2S)!}{(hu^\star)^{2S}} e^{\frac{h}{2S+1}(1-u^\star)} \sum_{j=2S}^{\infty} \frac{(hu^\star)^j}{j!}. \tag{6.48}$$

Now (6.27) (using Proposition 6.5.1) shows that for our general AB model, in the case $c > 0$ and Q not negative semidefinite (which includes the interchange model), the left hand side of (6.46) is the also of the form (6.48), at least when the maximiser of F (6.6) is unique.

In contrast, for the WB model with $a = b = 0$, $c = 1$ and $\rho = 1/2$, we can show that the observable of Theorem 6.1.7 is equal to 1 at all temperatures. Indeed, in the proof of Proposition 6.5.1, we will show that for $a = b = 0$, $c = 1$, and $\rho = 1/2$, the maximiser of F satisfies $x_i^\star = y_i^\star$, for all $i = 1,\dots,\theta$. This gives $z_i^\dagger = 0$ for all $i = 1,\dots,\theta$ ($z^\dagger$ from (6.24)), and after calculations with the function R, the limit in (6.23) is trivial. Note that by our comments below (6.10), the WB-model with $a = b = 0$, $c = 1$, has Hamiltonian unitarily equivalent to

$$-\frac{1}{n}\sum_{1\le i\le m<j\le n} P_{i,j}, \tag{6.49}$$

where $P_{i,j}$ is (θ times) the projection onto the singlet state, given by (6.11).

We can interpret this result to comment on the nematic (or biquadratic) model in spin $S = 1$ ($\theta = 3$). Our AB- and WB-models in spin $S = 1$ with $a = b = 0$, $c = \pm 1$ are special cases of a two-parameter model on the complete bipartite graph known as the bilinear-biquadratic Heisenberg model, which has Hamiltonian

$$-\frac{1}{n}\sum_{1\le i\le m<j\le n}\Big(J_1(\boldsymbol{S}_i\cdot\boldsymbol{S}_j)+J_2(\boldsymbol{S}_i\cdot\boldsymbol{S}_j)^2\Big), \tag{6.50}$$

where $\boldsymbol{S}_i\cdot\boldsymbol{S}_j = \sum_{k=1}^3 S_i^{(k)}S_j^{(k)}$, and $J_1, J_2 \in \mathbb{R}$. Indeed, using the relations $\boldsymbol{S}_i\cdot\boldsymbol{S}_j = T_{i,j} - P_{i,j}$ and $(\boldsymbol{S}_i\cdot\boldsymbol{S}_j)^2 = P_{i,j} + 1$ (see Lemma 7.1 from [101]) one can rewrite (6.50), up to addition of a constant, as

$$-\frac{1}{n}\sum_{1\le i\le m<j\le n}\Big(J_1T_{i,j} + (J_2 - J_1)P_{i,j}\Big). \tag{6.51}$$

Setting $J_1 = J_2 = \pm 1$ gives the AB model with $a = b = 0$, $c = \pm 1$, while setting $J_1 = 0$, $J_2 = \pm 1$ gives the WB model with $a = b = 0$, $c = \pm 1$, in the form (6.49). The case $J_1 = 0$, $J_2 = 1$ (i.e. our WB-model with $a = b = 0$, $c = 1$) is the nematic, or biquadratic, Heisenberg model.

The phase diagram of the bilinear-biquadratic Heisenberg model on $\mathbb{Z}^d$, $d \ge 3$ is given in Ueltschi [101], and we expect that the model on the complete bipartite graph has the same diagram. (The corresponding one-dimensional spin chain has a different phase-diagram, exhibiting dimerization, see [11,67].) The nematic model lies in the nematic phase of that diagram, and for low temperatures its extremal Gibbs states are expected to be indexed by $\mathbb{RP}^2$, projections in $\mathbb{R}^3$. Heuristically, we expect spins at all vertices to be either aligned or anti-aligned. In particular, one obtains that the right hand side of (6.46) is trivial when $W_i = \vec{v}\cdot\boldsymbol{S}_i = \sum_{k=1}^3 v_k S_i^{(k)}$, for any $\vec{v} \in \boldsymbol{S}^2$ (and non-trivial when $W_i = (\vec{v}\cdot\boldsymbol{S}_i)^2$). Now by Theorem 6.1.8, with $W_i = S_i^{(2)}$ (which satisfies $(S_i^{(2)})^\top = -S_i^{(2)}$), the left hand side of (6.46) equals 1. This aligns with the heuristics described above. One can also note that for all $\beta > 0$, the magnetisation from Theorem 6.1.8 is

$$\frac{\partial\Phi^{\mathrm{WB}}}{\partial h}\Big|_{h\downarrow 0} = \frac{\partial\Phi^{\mathrm{WB}}}{\partial h}\Big|_{h\uparrow 0} = 0; \tag{6.52}$$

again this aligns with the picture of $\Psi_\beta = \mathbb{RP}^2$ (we expect something nontrivial when the magnetisation term in the Hamiltonian is $\sum_{1\le i\le n}(S_i^{(2)})^2$).

6.2 Free energy and correlations

In this section we prove Theorems 6.1.1, 6.1.2, 6.1.7 and 6.1.8. Although Theorem 6.1.1 is actually a special case of Theorem 6.1.9, we give a detailed proof of Theorem 6.1.1 and then describe the modifications necessary to obtain Theorem 6.1.9 in Section 6.4.

6.2.1 Interchange model: proof of Theorem 6.1.1

As noted in the introduction, our method is to identify the eigenspaces of the Hamiltonian (6.5). This is facilitated by the classical theory Schur–Weyl duality. We start by recalling a few basic definitions and facts. A partition $\lambda \vdash n$ of n is a non-increasing sequence of non-negative integers summing to n: $\lambda = (\lambda_1, \lambda_2, \dots)$ with $\lambda_1 \ge \lambda_2 \ge \dots \ge 0$ and $\sum_{k\ge 1}\lambda_k = n$. Its length $\ell(\lambda)$ is the number of non-zero entries.

Recall the mapping $T : S_n \to \mathrm{End}(\mathbb{V})$ defined in (6.37). This is a representation of S_n and hence of the group algebra $\mathbb{C}[S_n]$ on $\mathbb{V}$. We may also regard $\mathbb{V}$ as a module for the group $GL(\theta)$ of invertible $\theta \times \theta$ matrices by the diagonal action (3.6)

$$g(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = g(v_1) \otimes g(v_2) \otimes \cdots \otimes g(v_n). \tag{6.53}$$

Classical Schur–Weyl duality [34, Corollary 4.59] states that these actions of S_n and of $GL(\theta)$ are each others' centralizers, so that $\mathbb{V}$ may be regarded as a representation of the direct product $GL(\theta) \times S_n$, and that $\mathbb{V}$ decomposes as a multiplicity-free direct sum of irreducible representations of $GL(\theta) \times S_n$. Specifically, from (3.0.2),

$$\mathbb{V} = \bigoplus_{\lambda \vdash n, \ell(\lambda) \leq \theta} \psi_\lambda^{GL(\theta)} \boxtimes \psi_\lambda^{S_n}. \tag{6.54}$$

Here $\psi_\lambda^{GL(\theta)}$ is the irreducible $GL(\theta)$-representation indexed by (its highest weight) λ (Theorem 2.1.16), and $\psi_\lambda^{S_n}$ is the irreducible S_n-representation (Specht module) indexed by λ (Theorem 2.1.5). We use the same notation T for the representation of $GL(\theta) \times S_n$ on $\mathbb{V}$.

Recall our Hamiltonian H_n^{AB} given in (6.5). We now write this as $H_n^{\mathrm{AB}} = T(h_n^{\mathrm{AB}})$ where

$$h_n^{\mathrm{AB}} = -\tfrac{1}{n}[(a-c)\alpha_A + (b-c)\alpha_B + c\,\alpha_{AB}], \tag{6.55}$$

and where $\alpha_A, \alpha_B, \alpha_{AB}$ are the following elements of $\mathbb{C}[S_n]$:

$$\alpha_A = \sum_{1 \leq i < j \leq m} (i,j), \quad \alpha_B = \sum_{m+1 \leq i < j \leq n} (i,j), \quad \alpha_{AB} = \sum_{1 \leq i < j \leq n} (i,j). \tag{6.56}$$

We have by linearity that $e^{-\beta H_n^{\mathrm{AB}}} = T(e^{-\beta h_n^{\mathrm{AB}}})$. Now let W be an $\theta \times \theta$ matrix over $\mathbb{C}$. Then $e^W \in GL(\theta)$ and we have that $T(e^W) = \exp\big(\sum_{i=1}^n W_i\big)$. Thus we may write

$$\exp\big(\textstyle\sum_{i=1}^n W_i\big) e^{-\beta H_n^{\mathrm{AB}}} = T\big(e^W \times e^{-\beta h_n^{\mathrm{AB}}}\big), \tag{6.57}$$

where $e^W \times e^{-\beta h_n^{\mathrm{AB}}} \in \mathbb{C}[GL(\theta) \times S_n]$.

Let us now consider how $e^W \times e^{-\beta h_n^{\mathrm{AB}}}$ acts on the right-hand-side of (6.54), starting with how $e^{-\beta h_n^{\mathrm{AB}}}$ acts on $\psi_\lambda^{S_n}$. The term α_{AB} is the sum of all elements of a conjugacy class (the transpositions), hence it belongs to the center of $\mathbb{C}[S_n]$. By Schur's Lemma, it therefore acts as a constant multiple of the identity on $\psi_\lambda^{S_n}$. The constant in question is well known [42, p. 52] to equal the contents of the partition λ, defined by

$$\mathrm{ct}(\lambda) = \sum_{j \geq 1} \Big(\frac{\lambda_j(\lambda_j + 1)}{2} - j\lambda_j \Big). \tag{6.58}$$

(This equals the sum of the contents of all boxes in any standard Young tableau of shape λ, where the contents of a box in position (x,y) is $y - x$.) We have

$$\alpha_{AB}|_{\psi_\lambda^{S_n}} = \mathrm{ct}(\lambda)\mathtt{Id}_{\psi_\lambda^{S_n}}. \tag{6.59}$$

Now, to deal with the remaining two terms α_A and α_B, note that as a representation of

$S_m \times S_{n-m}$, from (2.27), the module $\psi_\lambda^{S_n}$ splits as

$$\psi_\lambda^{S_n} = \bigoplus_{\mu \vdash m, \nu \vdash n-m} c_{\mu,\nu}^\lambda \psi_\mu^{S_m} \otimes \psi_\nu^{S_{n-m}}, \tag{6.60}$$

where $c_{\mu,\nu}^\lambda$ are non-negative integers known as the Littlewood–Richardson coefficients. We give more details about these numbers later, for now we just note that $c_{\mu,\nu}^\lambda = 0$ only if $\ell(\mu), \ell(\nu) \leq \ell(\lambda)$. On each term of the sum in (6.60), from (2.1.10), α_A acts as $\mathrm{ct}(\mu)\mathtt{Id}_{\psi_\mu^{S_m}}$ and α_B acts as $\mathrm{ct}(\nu)\mathtt{Id}_{\psi_\nu^{S_{n-m}}}$, consequently h_n^{AB} acts on that term as

$$-\tfrac{1}{n}[(a-c)\mathrm{ct}(\mu) + (b-c)\mathrm{ct}(\nu) + c\,\mathrm{ct}(\lambda)]\mathtt{Id}_{\psi_\mu^{S_m} \otimes \psi_\nu^{S_{n-m}}}, \tag{6.61}$$

and therefore $e^{-\beta h_n^{\mathrm{AB}}}$ acts as

$$\exp\left(\tfrac{\beta}{n}[(a-c)\mathrm{ct}(\mu) + (b-c)\mathrm{ct}(\nu) + c\,\mathrm{ct}(\lambda)]\right)\mathtt{Id}_{\psi_\mu^{S_m} \otimes \psi_\nu^{S_{n-m}}} \tag{6.62}$$

As to the factor e^W, we first note that from Lemma 2.1.21 the character of the module $\psi_\lambda^{GL(\theta)}$ evaluated at $g \in GL(\theta)$ with eigenvalues $x_1, \ldots, x_\theta$ is the Schur-polynomial:

$$\chi_\lambda^{GL(\theta)}[g] = s_\lambda(x_1, \ldots, x_\theta) = \frac{\det\left[x_i^{\lambda_j + \theta - j}\right]_{i,j=1}^{\theta}}{\prod_{1 \leq i < j \leq \theta}(x_i - x_j)}. \tag{6.63}$$

If W has eigenvalues $w_1, \ldots, w_\theta$, then e^W has eigenvalues $e^{w_1}, \ldots, e^{w_\theta}$. Writing $d_\mu^{S_m}, d_\nu^{S_{n-m}}$ for the dimensions of $\psi_\mu^{S_m}, \psi_\nu^{S_{n-m}}$, we may summarize these findings as follows:

Lemma 6.2.1. *Suppose that W has eigenvalues $w_1, \ldots, w_\theta$. Then*

$$\begin{aligned}\mathrm{Tr}_{\mathbb{V}}[\exp\left(\textstyle\sum_{i=1}^n W_i\right)e^{-\beta H_n^{AB}}] = \sum_{\lambda,\mu,\nu} & s_\lambda(e^{w_1}, \ldots, e^{w_\theta}) c_{\mu,\nu}^\lambda d_\mu^{S_m} d_\nu^{S_{n-m}} \\ & \exp\left(\tfrac{\beta}{n}[(a-c)\mathrm{ct}(\mu) + (b-c)\mathrm{ct}(\nu) + c \cdot \mathrm{ct}(\lambda)]\right),\end{aligned} \tag{6.64}$$

where the sum is over $\lambda \vdash n$ with $\ell(\lambda) \leq \theta$, $\mu \vdash m$, and $\nu \vdash n - m$. In particular, setting W to be the zero matrix (so that $e^W = \mathtt{Id}$),

$$Z_{\beta,n}^{AB} = \sum_{\lambda,\mu,\nu} s_\lambda(1, \ldots, 1) c_{\mu,\nu}^\lambda d_\mu^{S_m} d_\nu^{S_{n-m}} \exp\left(\tfrac{\beta}{n}[(a-c)\mathrm{ct}(\mu) + (b-c)\mathrm{ct}(\nu) + c \cdot \mathrm{ct}(\lambda)]\right). \tag{6.65}$$

Here we used the following specialization of s_λ:

$$d_\lambda^{GL(\theta)} = s_\lambda(1, \ldots, 1) = \prod_{1 \leq i < j \leq \theta} \frac{\lambda_i - i - \lambda_j + j}{j - i}. \tag{6.66}$$

As to $d_\mu^{S_m}$, a convenient formula is

$$d_\mu^{S_m} = \dim(\psi_\mu^{S_m}) = \frac{n!}{m_1! \cdots m_\theta!} \prod_{1 \leq i < j \leq \theta} (m_i - m_j) \tag{6.67}$$

where $m_i = \mu_i + \theta - i$, see [42, (4.11)].

In Lemma 6.2.1 we have written the partition function as a sum of terms exponentially large in n, with relatively few summands. Such a sum is dominated by its largest term.

To prove Theorem 6.1.1 we need to understand the asymptotic behavior of each of the factors in (6.65), and since only those terms with $c^\lambda_{\mu,\nu} \neq 0$ appear in the sum, we need a condition for $c^\lambda_{\mu,\nu} \neq 0$.

Proof of Theorem 6.1.1. First, from (6.66) we see that $d^{GL(\theta)}_\lambda = s_\lambda(1,\dots,1)$ is positive whenever $\ell(\lambda) \leq \theta$, and that $d^{GL(\theta)}_\lambda = \exp(o(1))$ where the $o(1)$ is uniform in λ. Now consider the coefficients $c^\lambda_{\mu,\nu}$. These are known (see e.g. [40, Chapter 5, Proposition 3]) to equal the size of a certain subset of semi-standard tableaux with shape $\lambda \setminus \mu$ filled with ν_1 1's, ν_2 2's, etc. In particular, $c^\lambda_{\mu,\nu} > 0$ only if μ is contained in λ, and then $\ell(\mu) \leq \ell(\lambda) \leq \theta$. Since $c^\lambda_{\mu,\nu} = c^\lambda_{\nu,\mu}$ (see [40] again) we also need $\ell(\nu) \leq \theta$ for $c^\lambda_{\mu,\nu} > 0$. The combinatorial description also gives the upper bound $c^\lambda_{\mu,\nu} \leq (n+1)^{\theta^2} = \exp(o(1))$ where the $o(1)$ is uniform in λ, μ, ν.

We now turn to the remaining factors in (6.65). First, as one can see in (6.67), for fixed θ we have that $d^{S_m}_\mu$ is essentially a multinomial coefficient. Thus (see e.g. [13, pp. 14–15] for details), we have

$$\tfrac{1}{n} \log d^{S_m}_\mu = -\textstyle\sum_{j=1}^\theta \frac{\mu_j}{n} \log \frac{\mu_j}{n} + O(\frac{\log n}{n}). \tag{6.68}$$

Next, from (6.58) we have that

$$\mathrm{ct}(\lambda) = \tfrac{n^2}{2} \textstyle\sum_{j=1}^\theta \big(\frac{\lambda_j}{n}\big)^2 + O(n). \tag{6.69}$$

Taken altogether, these facts mean that we can write (6.65) as

$$Z^{\mathrm{AB}}_{\beta,n} = \sum_{\lambda,\mu,\nu} \mathbb{1}\{c^\lambda_{\mu,\nu} > 0\} \exp\Big(n\Big\{\tilde F(\tfrac{\mu}{n}, \tfrac{\nu}{n}, \tfrac{\lambda}{n}) + o(1)\Big\}\Big), \tag{6.70}$$

where $\lambda \vdash n$, $\mu \vdash m$ and $\nu \vdash n-m$, all having $\leq \theta$ rows, and where

$$\begin{aligned} \tilde F(\vec x, \vec y, \vec z) = &- \textstyle\sum_{j=1}^\theta x_j \log x_j - \sum_{j=1}^\theta y_j \log y_j \\ &+ \tfrac{\beta}{2}\big[(a-c)\textstyle\sum_{j=1}^\theta x_j^2 + (b-c)\sum_{j=1}^\theta y_j^2 + c\sum_{j=1}^\theta z_j^2\big]. \end{aligned} \tag{6.71}$$

There is a sufficient condition for $c^\lambda_{\mu,\nu} > 0$ which is very useful for our purposes, known as Horn's inequalities. It is best stated in terms of eigenvalues of Hermitian matrices, as follows: $c^\lambda_{\mu,\nu} > 0$ if and only if there are Hermitian $\theta \times \theta$ matrices X and Y with eigenvalues $\mu_1, \dots, \mu_\theta$ and $\nu_1, \dots, \nu_\theta$, respectively, such that $X+Y$ has eigenvalues $\lambda_1, \dots, \lambda_\theta$. For information about this, see e.g. [41]. We thus have

$$c^\lambda_{\mu,\nu} > 0 \text{ if and only if } (\tfrac{\mu}{n}, \tfrac{\nu}{n}, \tfrac{\lambda}{n}) \in \Omega^+_{m/n} \tag{6.72}$$

where Ω^+_ρ is the set of triples $(\vec x, \vec y, \vec z)$ such that there exist positive semidefinite Hermitian matrices X, Y with $\mathrm{tr}(X) = 1 - \mathrm{tr}(Y) = \rho$ having eigenvalues $x_1, \dots, x_\theta$ and $y_1, \dots, y_\theta$, respectively, such that $Z = X+Y$ has eigenvalues $z_1, \dots, z_\theta$.

From (6.70) and the fact that $\tilde F$ is continuous in its arguments, we conclude that

$$\tfrac{1}{n} \log Z^{\mathrm{AB}}_{\beta,n} \to \max_{(\vec x, \vec y, \vec z) \in \Omega_\rho} \tilde F(\vec x, \vec y, \vec z). \tag{6.73}$$

See e.g. [13, Section 3] for a detailed argument in a similar setting. Now note that if

X, Y, Z are as above, then

$$\textstyle\sum_{j=1}^{\theta} x_j^2 = \mathrm{Tr}\big[X^2\big], \qquad \sum_{j=1}^{\theta} y_j^2 = \mathrm{Tr}\big[Y^2\big], \tag{6.74}$$

and also

$$\textstyle\sum_{j=1}^{\theta} z_j^2 = \mathrm{Tr}\big[Z^2\big] = \mathrm{Tr}\left[(X+Y)^2\right] = \mathrm{Tr}\big[X^2\big] + \mathrm{Tr}\big[Y^2\big] + 2\,\mathrm{Tr}\big[XY\big]. \tag{6.75}$$

Thus

$$(a-c)\sum_{j=1}^{\theta} x_j^2 + (b-c)\sum_{j=1}^{\theta} y_j^2 + c\sum_{j=1}^{\theta} z_j^2 = \mathrm{Tr}\left[aX^2 + bY^2 + 2cXY\right]. \tag{6.76}$$

So for $(\vec{x}, \vec{y}, \vec{z}) \in \Omega_\rho$, we have that

$$\tilde{F}(\vec{x}, \vec{y}, \vec{z}) = S(X) + S(Y) + \tfrac{\beta}{2}\,\mathrm{Tr}\left[aX^2 + bY^2 + 2cXY\right], \tag{6.77}$$

where S is as given in (6.39). It follows that

$$\tfrac{1}{n}\log Z_n^{\mathrm{AB}}(\beta) \to \max_{X,Y}\Big(S(X) + S(Y) + \tfrac{\beta}{2}\,\mathrm{Tr}\left[aX^2 + bY^2 + 2cXY\right]\Big) \tag{6.78}$$

where the maximum is over positive definite Hermitian matrices X, Y with $\mathrm{Tr}[X] = 1 - \mathrm{Tr}[Y] = \rho$.

The final step is to use the fact that for positive semidefinite Hermitian matrices X, Y with fixed spectra $x_1, \dots, x_\theta$ and $y_1, \dots, y_\theta$, respectively, ordered so that $x_1 \geq x_2 \geq \cdots \geq x_\theta$ and $y_1 \geq y_2 \geq \cdots \geq y_\theta$, we have the inequality

$$\sum_{j=1}^{\theta} x_j y_{\theta+1-j} \leq \mathrm{Tr}[XY] \leq \sum_{j=1}^{\theta} x_j y_j. \tag{6.79}$$

We discuss this result in Appendix 6.6. In particular, both the maximum and the minimum of $\mathrm{Tr}[XY]$ are attained when X, Y are simultaneously diagonal. Since the other terms in $F(\vec{x}, \vec{y})$ are symmetric under permuting the x_i or the y_i, the result follows. ∎

6.2.2 Walled Brauer algebra: proof of Theorem 6.1.2

As noted above, our analysis of the model in (6.10) uses the walled Brauer algebra. We will now define this algebra, and collect some facts which allow us to approach a proof in a similar way to that of Theorem 6.1.1. An accessible introduction to the walled Brauer algebra is given in [80], and its Schur–Weyl duality is proved in [8], at least for the range $\theta \geq n$. The extension to all θ, n is a straightforward extension of the work in [8] - this is of course covered in Chapter 3.

Let us first define the (usual) Brauer algebra. Fix $n \in \mathbb{N}, \theta \in \mathbb{C}$. Arrange two rows each of n labelled vertices, one above the other. We call a *diagram* a graph on these $2n$ vertices, with each vertex having degree one. Let B_n be the set of such diagrams. The Brauer algebra $\mathbb{B}_{n,\theta}$ is the formal complex span of B_n. Multiplication of two diagrams is defined as follows. Taking two diagrams g, h, identify the upper vertices of h with the lower of g. Then form a new diagram by concatenation and removing any closed loops, as in Figure 6.1. The product gh is the concatenation, multiplied by $\theta^{\#\mathrm{loops}}$, where #loops is the number of loops removed.

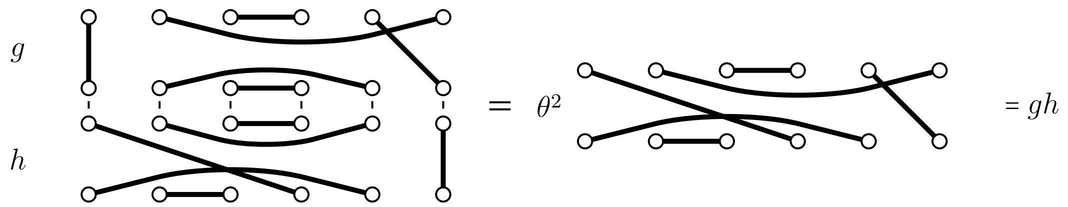


Figure 6.1: Two diagrams g and h (left), and their product (right). The concatenation contains two loops, so we multiply the concatenation with middle vertices removed by θ^2.

The walled Brauer algebra is a subalgebra of $\mathbb{B}_{n,\theta}$. Let $m \le n$. Returning to the $2n$ labelled vertices, draw a line (a "wall") separating the leftmost $2m$ vertices and the rightmost $2(n-m)$. Let $B_{n,m}$ be the set of diagrams in B_n with the condition that any edge connecting two upper vertices or two lower vertices *must* cross the wall, and any edge connecting an upper vertex and a lower vertex *must not* cross the wall. See Figure 6.2. The walled Brauer algebra $\mathbb{B}_{n,m,\theta}$ is the span of $B_{n,m}$, with multiplication as in the Brauer algebra.

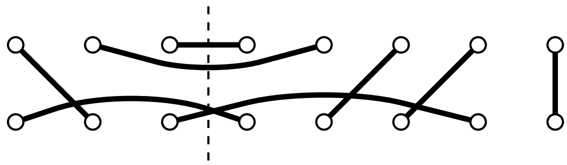


Figure 6.2: A diagram in the basis $B_{8,3}$ of the walled Brauer algebra $\mathbb{B}_{8,3,\theta}$. Notice that all edges connecting two upper vertices (or two lower) cross the wall, and all edges connecting an upper vertex to a lower vertex do not.

Some useful representation-theoretic facts follow. First, the group algebra $\mathbb{C}[S_m \times S_{n-m}]$ is a subalgebra of $\mathbb{B}_{n,m,\theta}$ whose basis $S_m \times S_{n-m}$ consists of those diagrams with no edges crossing the wall. As above, we let (i,j) denote the transposition exchanging i and j. Note that in the walled Brauer algebra, we must have $1 \le i,j \le m$ or $m+1 \le i,j \le n$. For $1 \le i \le m < j \le n$, let $(\overline{i,j})$ denote the diagram with all edges vertical, except that the i^{th} and j^{th} upper vertices are connected, and the i^{th} and j^{th} lower vertices are connected. See Figure 6.3. The elements (i,j) and $(\overline{i,j})$ generate the walled Brauer algebra.

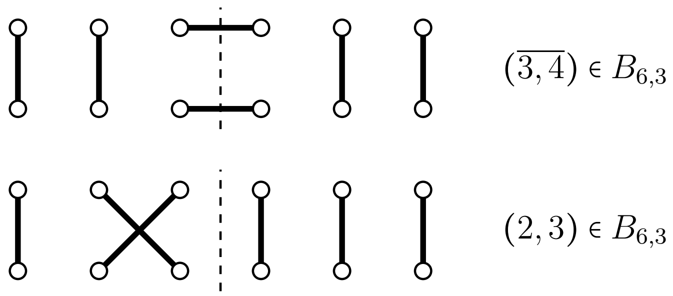


Figure 6.3: Examples of the elements $(\overline{i,j})$ and the transpositions (i,j).

Next, from (2.1.13), the irreducible representations of $\mathbb{B}_{n,m,\theta}$ are indexed by

$$\{(\lambda,\mu) \mid \lambda \vdash m-t,\ \mu \vdash n-m-t,\ t = 0,\dots,\min\{m,n-m\}\ \}, \tag{6.80}$$

where λ and μ are partitions (see Proposition 2.4 of [24]). Henceforth, we will assume without loss of generality that $m \leq n/2$ so that the standing condition on t is that $t \in \{0, 1, \ldots, m\}$. The element

$$J_{n,m} = \sum_{\substack{1 \leq i < j \leq m \\ m < i < j < n}} (i, j) - \sum_{1 \leq i \leq m < j \leq n} (\overline{i, j}). \tag{6.81}$$

is central in $\mathbb{B}_{n,m,\theta}$, and, from (2.15) acts as the scalar $\mathrm{ct}(\lambda) + \mathrm{ct}(\mu) - \theta t$ on the irreducible representation (λ, μ), where $\lambda \vdash m - t$, $\mu \vdash n - m - t$ and $\mathrm{ct}(\cdot)$ denotes the contents defined in (6.58) (a consequence of, for example, Lemma 4.1 of [23]).

The walled Brauer algebra, like the symmetric group algebra, has a Schur–Weyl duality with the general linear group. To describe this, let us first recall some facts about representations of the general linear group $GL(\theta)$. The irreducible (finite-dimensional) rational representations of $GL(\theta)$ are indexed by their highest weights, which are θ-tuples $\nu = (\nu_1 \geq \cdots \geq \nu_\theta) \in \mathbb{Z}^\theta$. Such a tuple can be equivalently written as a pair $\nu = [\lambda, \mu]$ of partitions λ, μ with $\lambda_1^\top + \mu_1^\top \leq \theta$, by letting $\nu_i = [\lambda, \mu]_i = \lambda_i - \mu_{\theta - i + 1}$ for $i = 1, \ldots, \theta$. Note that at most one of the terms λ_i or $\mu_{\theta-i+1}$ is non-zero for each i, due to the constraint $\lambda_1^\top + \mu_1^\top \leq \theta$, thus ν uniquely determines λ and μ. See Figure 6.4 for an illustration.

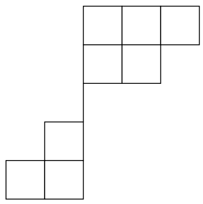


Figure 6.4: The θ-tuple $\nu = (3, 3, 0, -1, -2)$ illustrated in the style of a Young diagram, where negative entries are shown by boxes to the left of the main vertical line. Here $\theta = 5$. From the figure it is straightforward to see that $\nu = [\lambda, \mu]$, where $\lambda = (3, 2)$ and $\mu = (2, 1)$.

We write $\psi_{[\lambda,\mu]}^{GL(\theta)}$ for the corresponding irreducible $GL(\theta)$-module. These rational representations are closely related to the polynomial representations $\psi_\lambda^{GL(\theta)}$ appearing in (6.54); the polynomial representations are the rational representations with non-negative θ-tuple ν. One can also relate the rational and polynomial representations by the Pieri-rule (2.35) [93]. Indeed, writing $\det(\cdot)$ for the determinant representation of $GL(\theta)$, which has highest weight $(1, 1, \ldots, 1)$ and character $x_1 x_2 \cdots x_\theta$, we have that $\det^{\otimes k} \otimes \psi_\nu^{GL(\theta)} = \psi_{\nu + \underline{k}}^{GL(\theta)}$ where $\underline{k} = (k, k, \ldots, k)$. For $k = \mu_1$ we have that $\psi_{[\lambda,\mu]+\underline{\mu_1}}^{GL(\theta)}$ is a polynomial representation. It follows from this and (6.63) that the character of $\psi_{[\lambda,\mu]}^{GL(\theta)}$ is

$$\chi_{[\lambda,\mu]}^{GL(\theta)}[g] = \frac{s_{[\lambda,\mu]+\underline{\mu_1}}(x_1, \ldots, x_\theta)}{(x_1 x_2 \cdots x_\theta)^{\mu_1}} = \frac{\det\left[x_i^{[\lambda,\mu]_j + \theta - j}\right]_{i,j=1}^{\theta}}{\prod_{1 \leq i < j \leq \theta}(x_i - x_j)}, \tag{6.82}$$

where $x_1, \ldots, x_\theta$ are the eigenvalues of g.

Now, let $GL(\theta)$ act on $\mathbb{V} = V^{\otimes n} = V^{\otimes m} \otimes V^{\otimes(n-m)}$ as m tensor powers of its defining representation, and $n - m$ tensor powers of the dual of its defining representation (multiplication by the inverse transpose) (3.15):

$$g(v_1 \otimes \cdots \otimes v_m \otimes v_{m+1} \otimes \cdots \otimes v_n) = g(v_1) \otimes \cdots \otimes g(v_m) \otimes g^{-\top}(v_{m+1}) \otimes \cdots \otimes g^{-\top}(v_n).$$

Let $\mathbb{B}_{n,m,\theta}$ act on $\mathbb{V}$ by sending (i,j) to the transposition operator $T_{i,j}$, and $(\overline{i,j})$ to $Q_{i,j}$ (3.12). Then from Theorem 3.0.5, as a representation of $\mathbb{C}[GL(\theta)] \otimes \mathbb{B}_{n,m,\theta}$,

$$\mathbb{V} = \bigoplus_{t=0}^{m} \bigoplus_{\substack{\lambda \vdash m-t \\ \mu \vdash n-m-t \\ \lambda_1^\top + \mu_1^\top \le \theta}} \psi_{[\lambda,\mu]}^{GL(\theta)} \boxtimes \psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}, \tag{6.83}$$

with $\psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}$ irreducible $\mathbb{B}_{n,m,\theta}$-representations as above (as noted above, this is a straightforward extension of the work in [8]).

Notice now that our Hamiltonian (6.10) can be rewritten as

$$H_n^{\text{WB}} = -\frac{1}{n}\Big((a+c)\sum_{1\le i<j\le m} T_{i,j} + (b+c)\sum_{m+1\le i<j\le n} T_{i,j} - cJ_{n,m}\Big), \tag{6.84}$$

where $J_{n,m}$ is the central element given in (6.81). Now in an identical way to how we developed equation (6.65), we have

$$\begin{aligned} \mathrm{Tr}_{\mathbb{V}}[e^{-\beta H_n^{\text{WB}}}] = &\sum_{t=0}^{m} \sum_{\substack{\lambda \vdash m-t \\ \mu \vdash n-m-t \\ \lambda_1^\top + \mu_1^\top \le \theta}} \sum_{\substack{\pi \vdash m \\ \tau \vdash n-m}} d_{[\lambda,\mu]}^{GL(\theta)} b_{(\lambda,\mu),(\pi,\tau)}^{n,m,\theta} d_\pi^{S_m} d_\tau^{S_{n-m}} \cdot \\ &\cdot \exp\Big(\beta\big[(c+a)\mathrm{ct}(\pi) + (c+b)\mathrm{ct}(\tau) - c(\mathrm{ct}(\lambda) + \mathrm{ct}(\mu) - rt)\big]\Big), \end{aligned} \tag{6.85}$$

where $b_{(\lambda,\mu),(\pi,\tau)}^{n,m,\theta}$ is the branching coefficient (2.40) from $\mathbb{C}[S_m \times S_{n-m}]$ to $\mathbb{B}_{n,m,\theta}$, i.e. the multiplicity of the $\mathbb{C}[S_m \times S_{n-m}]$-module $\psi_\pi^{S_m} \otimes \psi_\tau^{S_{n-m}}$ in $\psi_{(\lambda,\mu)}^{\mathbb{B}_{n,m,\theta}}$ when the latter is regarded as a $\mathbb{C}[S_m \times S_{n-m}]$-module. These branching coefficients play the same role as the Littlewood–Richardson coefficient did in the AB-model. Our next step is to determine when $b_{(\lambda,\mu),(\pi,\tau)}^{n,m,\theta}$ is strictly positive.

Lemma 6.2.2. *The branching coefficient $b_{(\lambda,\mu),(\pi,\tau)}^{n,m,\theta}$ is strictly positive if and only if there exist $\theta\times\theta$ Hermitian matrices X, Y, Z with respective spectra $\pi, \tau, [\lambda,\mu]$, such that $X - Y = Z$.*

Note that the parameter t is encoded the branching coefficient, in the sense that $b_{(\lambda,\mu),(\pi,\tau)}^{n,m,\theta} > 0$ implies that $\lambda \vdash m-t = |\pi|-t$ and $\mu \vdash n-m-t = |\tau|-t$ for some $0 \le t \le \hat{m}$. The same conclusion can be seen to follow from the Hermitian matrices side of Lemma 6.2.2. Indeed, if X, Y, Z are Hermitian with respective spectra $\pi, \tau, [\lambda,\mu]$, such that $X - Y = Z$, then X, Y are simultaneously diagonalisable, so for each i, $[\lambda,\mu]_i = \pi_j - \tau_k$, for some j, k. Figure 6.5 then illustrates via an example how $\lambda \vdash m - t = |\pi| - t$ and $\mu \vdash n - m - t = |\tau| - t$ for some $0 \le t \le \hat{m}$ follows.

The first step to prove Lemma 6.2.2 is another lemma, analogous to the well known fact that the Littlewood–Richardson coefficients are both the branching coefficients from $\mathbb{C}[S_m \times S_{n-m}]$ to $\mathbb{C}[S_n]$, and the coefficients of the decomposition of the tensor product of two irreducible polynomial representations of $GL(\theta)$.

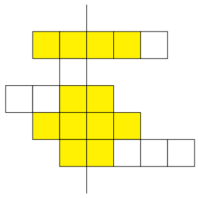


Figure 6.5: The spectra $\pi = (3,0,1,2,4)$ and $\tau = (2,1,3,2,1)$, respectively of X and Y (simultaneously diagonalised), displayed in the style of Young diagrams, either side of the main vertical line. The spectrum of $Z = X - Y$ is $(1,-1,-2,0,3)$ (and so when ordered becomes $[\lambda,\mu] = (3,1,0,-1,-2)$). The yellow boxes are those eliminated in the subtraction. Naturally there are the same number either side of the main vertical - this is the parameter $0 \le t \le \min|\pi|,|\tau|$; in this example, $t = 6$.

Lemma 6.2.3. *Let π,τ,λ,μ be partitions with at most θ parts, with $\lambda_1^\top + \mu_1^\top \le \theta$, and let*

$$\psi_\pi^{GL(\theta)} \otimes \psi_{[\varnothing,\tau]}^{GL(\theta)} = \bigoplus_{\substack{\lambda,\mu \\ \lambda_1^\top+\mu_1^\top\le\theta}} \hat{b}^{n,m,\theta}_{[\lambda,\mu],(\pi,\tau)} \psi^{GL(\theta)}_{[\lambda,\mu]}. \tag{6.86}$$

Then $\hat{b}^{n,m,\theta}_{[\lambda,\mu],(\pi,\tau)} = b^{n,m,\theta}_{(\lambda,\mu),(\pi,\tau)}$.

Proof. This is proved using Schur–Weyl duality. We restrict (6.83) to $\mathbb{C}[GL(\theta)] \otimes \mathbb{C}[S_m \times S_{n-m}]$ to see that

$$\mathbb{V} = \bigoplus_{t=0}^{m} \bigoplus_{\substack{\lambda\vdash m-t \\ \mu\vdash n-m-t \\ \lambda_1^\top+\mu_1^\top\le\theta}} \bigoplus_{\substack{\pi\vdash m \\ \tau\vdash n-m \\ \pi_1^\top,\tau_1^\top\le\theta}} b^{n,m,\theta}_{(\lambda,\mu),(\pi,\tau)} \psi^{GL(\theta)}_{[\lambda,\mu]} \otimes (\psi_\pi^{S_m} \otimes \psi_\tau^{S_{m-n}}). \tag{6.87}$$

On the other hand, the Schur–Weyl duality between $GL(\theta) \times GL(\theta)$ and $\mathbb{C}[S_m \times S_{n-m}]$ is

$$\mathbb{V} = \bigoplus_{\substack{\pi\vdash m \\ \tau\vdash n-m \\ \pi_1^\top,\tau_1^\top\le\theta}} (\psi_\pi^{GL(\theta)} \otimes \psi_{[\varnothing,\tau]}^{GL(\theta)}) \otimes (\psi_\pi^{S_m} \otimes \psi_\tau^{S_{m-n}}). \tag{6.88}$$

Expanding $\psi_\pi^{GL(\theta)} \otimes \psi_{[\varnothing,\tau]}^{GL(\theta)}$ as in (6.86) and equating coefficients from the two equations above, gives the result. ∎

Proof of Lemma 6.2.2. We take equation (6.86) and modify it using the Pieri rule (2.35):

$$\psi_\pi^{GL(\theta)} \otimes \psi_{[\varnothing,\tau]+\underline{\tau_1}}^{GL(\theta)} = \bigoplus_{\substack{\lambda,\mu \\ \lambda_1^\top+\mu_1^\top\le\theta}} \hat{b}^{n,m,\theta}_{[\lambda,\mu],(\pi,\tau)} \psi^{GL(\theta)}_{[\lambda,\mu]+\underline{\tau_1}}. \tag{6.89}$$

Now the highest weights appearing on both sides have no negative parts, so by the previous Lemma and the Littlewood–Richardson Rule,

$$b^{n,m,\theta}_{(\lambda,\mu),(\pi,\tau)} = \hat{b}^{n,m,\theta}_{[\lambda,\mu],(\pi,\tau)} = c^{[\lambda,\mu]+\underline{\tau_1}}_{\pi,[\varnothing,\tau]+\underline{\tau_1}}. \tag{6.90}$$

We know from Horn's inequalities that $c^{[\lambda,\mu]+\underline{\tau_1}}_{\pi,[\varnothing,\tau]+\underline{\tau_1}} > 0$ if and only if there exist $\theta \times \theta$ Hermitian $\bar{X}, \bar{Y}, \bar{Z}$ with respective spectra $\pi, [\varnothing,\tau]+\underline{\tau_1}$ and $[\lambda,\mu]+\underline{\tau_1}$ such that $\bar{X} + \bar{Y} = \bar{Z}$. Now it is straightforward to show that such matrices exist if and only if there exist $\theta \times \theta$

Hermitian X, Y, Z with respective spectra π, τ and $[\lambda, \mu]$ such that $X - Y = Z$. Indeed, let $X = \bar{X}$, $Y = -\bar{Y} + \tau_1 \mathtt{Id}$, and $Z = \bar{Z} - \tau_1 \mathtt{Id}$ for the first implication, and similarly for the reverse implication. ■

We can now return to equation (6.85). Using similar workings as in Section 6.2.1, we let $m, n \to \infty$ such that $m/n \to \rho \in (0, 1/2]$ (recall that we assumed $m \leq n - m$), $\pi/n \to \vec{x}$, $\tau/n \to \vec{y}$ and $[\lambda, \mu]/n \to \vec{z}$. Note that $\vec{z}$ can now have negative entries, and that

$$\frac{\mathrm{ct}(\lambda) + \mathrm{ct}(\mu) - \theta t}{n^2} = \sum_{i=1}^{\theta} \Big(\big(\tfrac{\lambda_i}{n}\big)^2 + \big(-\tfrac{\mu_i}{n}\big)^2 \Big) + o(1) = \sum_{i=1}^{\theta} \Big(\tfrac{[\lambda,\mu]_i}{n} \Big)^2 + o(1). \tag{6.91}$$

We find that

$$Z_n^{\mathrm{WB}}(\beta) = \sum_{\substack{\pi \vdash m \\ \tau \vdash n-m}} \sum_{\substack{\lambda,\mu \\ (\pi/n, \tau/n, [\lambda,\mu]/n) \in \Omega^-_{m/n}}} \exp\Big(n \Big\{ \tilde{G}\big(\tfrac{\pi}{n}, \tfrac{\tau}{n}, \tfrac{[\lambda,\mu]}{n}\big) + o(1) \Big\} \Big) \tag{6.92}$$

where Ω^-_ρ is the set of triples of θ-tuples $\vec{x}, \vec{y}, \vec{z}$ such that $x_1, \ldots, x_\theta \geq 0$, $y_1, \ldots, y_\theta \geq 0$, $\sum_{i=1}^{\theta} x_i = \rho = 1 - \sum_{i=1}^{\theta} y_i$, and there exist $\theta \times \theta$ Hermitian matrices X, Y, Z with respective spectra $\vec{x}, \vec{y}, \vec{z}$ such that $X - Y = Z$, and where

$$\tilde{G}(\vec{x}, \vec{y}, \vec{z}) = \sum_{i=1}^{\theta} \big[\tfrac{\beta}{2} \big((a+c) x_i^2 + (b+c) y_i^2 - c z_i^2 \big) - x_i \log x_i - y_i \log y_i \big]. \tag{6.93}$$

Notice that the sum over t appearing in (6.85) is hidden in (6.92), as it is implicit in the definition of Ω^-_ρ, due to our remark after the statement of Lemma 6.2.2. Therefore

$$\Phi_\beta^{\mathrm{WB}}(a, b, c) := \lim_{n \to \infty} \frac{1}{n} \log Z_n^{\mathrm{WB}}(\beta) = \max_{(\vec{x}, \vec{y}, \vec{z}) \in \Omega^-_\rho} \tilde{G}(\vec{x}, \vec{y}, \vec{z}). \tag{6.94}$$

As in (6.77) and (6.78), we can rewrite this in terms of the matrices X and Y:

$$\Phi_\beta^{\mathrm{WB}}(a, b, c) = \max_{X,Y} \big[S(X) + S(Y) + \tfrac{\beta}{2} \big(a\, \mathrm{tr}[X^2] + b\, \mathrm{tr}[Y^2] + 2c\, \mathrm{tr}[XY] \big) \big], \tag{6.95}$$

where now the maximum is only over $\theta \times \theta$ Hermitian matrices X, Y with respective spectra $\vec{x}, \vec{y}$ as above. This is the same as (6.78), and this completes the proof of Theorem 6.1.2. ■

6.2.3 Correlation functions: proof of Theorem 6.1.7

Let us prove the result for the AB-model first. We use (6.64) and the argument leading up to (6.70) to get that, as $n \to \infty$,

$$\Big\langle \exp\big\{ \tfrac{1}{n} \textstyle\sum_{i=1}^n W_i \big\} \Big\rangle_{\beta,n}^{\mathrm{AB}} = \frac{\sum_{\lambda,\mu,\nu} \mathbb{1}\{c^\lambda_{\mu,\nu} > 0\} \frac{s_\lambda(e^{w_1}, \ldots, e^{w_\theta})}{s_\lambda(1, \ldots, 1)} \exp\Big(n \Big\{ \tilde{F}\big(\tfrac{\mu}{n}, \tfrac{\nu}{n}, \tfrac{\lambda}{n}\big) + o(1) \Big\} \Big)}{\sum_{\lambda,\mu,\nu} \mathbb{1}\{c^\lambda_{\mu,\nu} > 0\} \exp\Big(n \Big\{ \tilde{F}\big(\tfrac{\mu}{n}, \tfrac{\nu}{n}, \tfrac{\lambda}{n}\big) + o(1) \Big\} \Big)}. \tag{6.96}$$

Both sums on the right-hand-side are over $\lambda \vdash n$, $\mu \vdash m$ and $\nu \vdash n - m$, all having at most θ parts, and in the numerator we have multiplied and divided by $d_\lambda^{GL(\theta)} = s_\lambda(1, \ldots, 1)$ in order that the $o(1)$ terms in the exponents are exactly equal. Then the arguments of

[14, Section 6] apply, meaning that

$$\lim_{n\to\infty}\left\langle \exp\left\{\tfrac{1}{n}\textstyle\sum_{i=1}^{n} W_i\right\}\right\rangle_{\beta,n}^{\mathrm{AB}} = \lim_{\lambda/n\to\vec{z}^\star} \frac{s_\lambda(e^{w_1},\ldots,e^{w_\theta})}{s_\lambda(1,\ldots,1)}, \tag{6.97}$$

where $\vec{z}^\star = (z_1^\star,\ldots,z_\theta^\star)$ lists the eigenvalues of $X+Y$ where X,Y are the Hermitian matrices which maximize the right-hand-side of (6.78). But we know from (6.79) that the maximum is attained when X,Y are simultaneously diagonal, with ordering of eigenvalues decreasing for both X and Y if $c>0$, respectively decreasing for X and increasing for Y if $c<0$. Then clearly the eigenvalues of $Z = X+Y$ are the sums of the eigenvalues of X and of Y, ordered appropriately, giving $z^\star$ as in (6.24).

Turning to the WB-model, very similarly to equation (6.96) we have

$$\begin{aligned}&\left\langle \exp\left\{\tfrac{1}{n}\left(\textstyle\sum_{i=1}^{m} W_i - \sum_{i=m+1}^{n} W_i^\top\right)\right\}\right\rangle_{\beta,n}^{\mathrm{WB}} \\ &= \frac{\sum_{\lambda,\mu,\pi,\tau} \mathbb{1}\{b^{n,m,\theta}_{[\lambda,\mu],(\pi,\tau)}>0\} \frac{\chi^{GL(\theta)}_{[\lambda,\mu]}(e^{W/n})}{d^{GL(\theta)}_{[\lambda,\mu]}} \exp\left(n\left\{\tilde{G}(\frac{\pi}{n},\frac{\tau}{n},\frac{[\lambda,\mu]}{n})+o(1)\right\}\right)}{\sum_{\lambda,\mu,\pi,\tau} \mathbb{1}\{b^{n,m,\theta}_{[\lambda,\mu],(\pi,\tau)}>0\} \exp\left(n\left\{\tilde{G}(\frac{\pi}{n},\frac{\tau}{n},\frac{[\lambda,\mu]}{n})+o(1)\right\}\right)},\end{aligned} \tag{6.98}$$

where once again the $o(1)$ terms in the exponents are exactly equal and now

$$\tilde{G}(\vec{x},\vec{y},\vec{z}) = \sum_{i=1}^{\theta}\left[\tfrac{\beta}{2}\left((a+c)x_i^2 + (b+c)y_i^2 - cz_i^2\right) - x_i\log x_i - y_i \log y_i\right]. \tag{6.99}$$

The arguments of [14, Section 6] apply once again, meaning the limit equals

$$\lim_{[\lambda,\mu]/n\to z^\dagger} \frac{\chi^{GL(\theta)}_{[\lambda,\mu]}(e^{W/n})}{d^{GL(\theta)}_{[\lambda,\mu]}}, \tag{6.100}$$

where this time, $(\vec{x}^\star,\vec{y}^\star,\vec{z}^\dagger)$ maximises $\tilde{G}(\vec{x},\vec{y},\vec{z})$, with the conditions that $x_i,y_i\geq 0$, $\sum_{i=1}^{\theta} x_i = \rho = 1-\sum_{i=1}^{\theta} y_i$, and that there exist Hermitian matrices X,Y,Z with respective spectra x,y,z with $X-Y=Z$. Following equation (6.95), we can rewrite $\tilde{G}$ as the function of the matrices X and Y being maximised in (6.95). If the entries of $\vec{x}$ are ordered decreasing, then as before the trace-inequality (6.79) implies that for $c>0$ the entries of $\vec{y}$ should also be ordered decreasing, while for $c<0$ they should be ordered decreasing. This gives the form of $\vec{z}^\dagger$ stated in (6.24).

It remains only to show that

$$\lim_{[\lambda,\mu]/n\to z} \frac{\chi^{GL(\theta)}_{[\lambda,\mu]}(e^{W/n})}{d^{GL(\theta)}_{[\lambda,\mu]}} = R(w_1,\ldots,w_\theta; z_1,\ldots,z_\theta). \tag{6.101}$$

This is proved almost identically to Lemma 6.1 from [14]. Indeed, using (6.82) we get

$$\begin{aligned}\frac{\chi^{GL(\theta)}_{[\lambda,\mu]}(e^{W/n})}{d^{GL(\theta)}_{[\lambda,\mu]}} &= \det\left[e^{w_i[\lambda,\mu]_j/n + w_i(\theta-j)/n}\right] \\ &\cdot \prod_{1\leq i<j\leq\theta} \frac{j-i}{(e^{w_i/n}-e^{w_j/n})([\lambda,\mu]_i-[\lambda,\mu]_j+j-i)},\end{aligned} \tag{6.102}$$

which, noting all the products (including in the determinant) are finite, tends to the function $R(w_1,\dots,w_\theta;z_1,\dots,z_\theta)$ as $[\lambda,\mu]/n \to z$. ∎

6.2.4 Magnetisation term: proof of Theorem 6.1.8

We start by giving expressions for the free energy with a magnetisation term, and then afterwards we will take the appropriate derivatives. We will need the following notation:

- Δ^+ will denote the set of vectors $\vec{z}=(z_1,z_2,\dots,z_\theta)$ that can arise as spectra of $X+Y$ where X and Y are positive semidefinite Hermitian matrices with $\mathrm{tr}[X]=1-\mathrm{tr}[Y]=\rho$, ordered so that $z_1\geq\dots\geq z_\theta$. In fact, Δ^+ consists of all $\vec{z}$ satisfying $z_1\geq\dots\geq z_\theta\geq 0$ and $\sum_{i=1}^\theta z_i=1$. Given $\vec{z}\in\Delta^+$, we write $\mathcal{H}_\rho^+(\vec{z})$ for the set of pairs (X,Y) of such matrices with $X+Y$ having spectrum $\vec{z}$.

- Δ_ρ^- will denote the set of vectors $\vec{z}=(z_1,z_2,\dots,z_\theta)$ that can arise as spectra of $X-Y$ where X and Y are as above, again ordered so that $z_1\geq\dots\geq z_\theta$. Now Δ_ρ^- consists of all $\vec{z}$ satisfying $\rho\geq z_1\geq\dots\geq z_r\geq-(1-\rho)$ and $\sum_{i=1}^\theta z_i=2\rho-1$. Given $\vec{z}\in\Delta_\rho^-$, we write $\mathcal{H}_\rho^-(\vec{z})$ for the set of pairs (X,Y) of such matrices with $X-Y$ having spectrum $\vec{z}$.

Let $\Phi^\#(\beta,h)=\Phi^\#_{\beta,h}(a,b,c,\vec{w})$ be as in (6.30) and recall from (6.77) that

$$\phi(X,Y)=S(X)+S(Y)+\tfrac{\beta}{2}\mathrm{tr}\big[aX^2+bY^2+2cXY\big].$$

Theorem 6.2.4. *Let $a,b,c\in\mathbb{R}$ and $w_1\geq\dots\geq w_\theta$ be fixed. If $n,m\to\infty$ such that $m/n\to\rho\in(0,1)$, then the free energy of the models (6.28) and (6.29) satisfy:*

$$\begin{aligned}\Phi^{AB}(\beta,h)&=\max_{\vec{z}\in\Delta^+}\left(\max_{(X,Y)\in\mathcal{H}_\rho^+(\vec{z})}\phi(X,Y)+\begin{cases}h\sum_{i=1}^\theta z_iw_i, & \text{if } h>0,\\ h\sum_{i=1}^\theta z_iw_{\theta+1-i}, & \text{if } h<0,\end{cases}\right)\\ \Phi^{WB}(\beta,h)&=\max_{\vec{z}\in\Delta_\rho^-}\left(\max_{(X,Y)\in\mathcal{H}_\rho^-(\vec{z})}\phi(X,Y)+\begin{cases}h\sum_{i=1}^\theta z_iw_i, & \text{if } h>0,\\ h\sum_{i=1}^\theta z_iw_{\theta+1-i}, & \text{if } h<0,\end{cases}\right).\end{aligned}\tag{6.103}$$

Proof. Let us start with the AB case. Using the expression (6.64) and arguing similarly to (6.70) we have

$$\begin{aligned}Z_{n,h}^{\mathrm{AB}}&=\sum_{\mu,\nu,\lambda}s_\lambda(e^{hw_1},\dots,e^{hw_\theta})\\&\quad\cdot c_{\mu,\nu}^\lambda d_\mu d_\nu\exp\Big(\tfrac{\beta}{n}[(a-c)\mathrm{ct}(\mu)+(b-c)\mathrm{ct}(\nu)+c\cdot\mathrm{ct}(\lambda)]\Big)\\&=\sum_{(\mu/n,\nu/n,\lambda/n)\in\Omega_{m/n}^+}s_\lambda(e^{hw_1},\dots,e^{hw_\theta})\exp\Big(n\Big\{\tilde{F}(\tfrac{\mu}{n},\tfrac{\nu}{n},\tfrac{\lambda}{n})+o(1)\Big\}\Big),\end{aligned}\tag{6.104}$$

where $\tilde{F}$ is given in (6.71) and Ω_ρ^+ in (6.72). Recall that [40, Section 2.2]

$$s_\lambda(e^{hw_1},\dots,e^{hw_\theta})=\sum_{\mathbb{T}}\prod_{i=1}^\theta e^{hm_iw_i}=\sum_{\mathbb{T}}e^{\sum_{i=1}^\theta hm_iw_i},\tag{6.105}$$

where the sum is over all semistandard Young tableaux $\mathbb{T}$ with shape λ and entries in $\{1,\dots,\theta\}$, and where for each i, m_i is the number of times the number i appears in $\mathbb{T}$. The tableau with each box in the i^{th} row labelled i appears in the sum, and in fact, for

$h > 0$, it maximises the sum in the exponent:

$$e^{\sum_{i=1}^{\theta} hm_i w_i} \le e^{\sum_{i=1}^{\theta} h\lambda_i w_i}, \tag{6.106}$$

for each valid $\mathbb{T}$. Indeed, note that in a semistandard tableau, the entries of row i must be at least i. Then, taking any semistandard $\mathbb{T}$, shape λ, changing an entry $j \ge i$ in row i to i changes the sum in the exponent by $h(w_i - w_j)$, which is non-negative by our ordering of $\vec{w}$ as $w_1 \ge \cdots \ge w_r$. Hence for $h > 0$,

$$e^{\sum_{i=1}^{\theta} h\lambda_i w_i} \le s_\lambda(e^{hw_1}, \dots, e^{hw_\theta}) \le d_\lambda^{S_n} e^{\sum_{i=1}^{\theta} h\lambda_i w_i}. \tag{6.107}$$

Recalling that $\frac{1}{n} \log d_\lambda^{S_n} \to 0$ we get, for $h > 0$,

$$Z_{n,h}^{\mathrm{AB}} = \sum_{(\mu/n, \nu/n, \lambda/n) \in \Omega_{m/n}^{+}} \exp\Big(n\Big\{\tilde{F}(\tfrac{\mu}{n}, \tfrac{\nu}{n}, \tfrac{\lambda}{n}) + h\textstyle\sum_{i=1}^{\theta} \frac{\lambda_i}{n} w_i + o(1)\Big\}\Big), \tag{6.108}$$

In the case $h < 0$, the sum in the exponent in (6.105) is maximised when $m_i = \lambda_{\theta+1-i}$ for each i; indeed, let $h' = -h$, and $w_i' = -w_{\theta+1-i}$, and apply the same reasoning as above. So, for $h < 0$, we have

$$e^{\sum_{i=1}^{\theta} h\lambda_{\theta+1-i} w_i} \le s_\lambda(e^{hw_1}, \dots, e^{hw_\theta}) \le d_\lambda^{S_n} e^{\sum_{i=1}^{\theta} h\lambda_{\theta+1-i} w_i}, \tag{6.109}$$

and so for $h < 0$,

$$Z_{n,h}^{\mathrm{AB}} = \sum_{(\mu/n, \nu/n, \lambda/n) \in \Omega_{m/n}^{+}} \exp\Big(n\Big\{\tilde{F}(\tfrac{\mu}{n}, \tfrac{\nu}{n}, \tfrac{\lambda}{n}) + h\textstyle\sum_{i=1}^{\theta} \frac{\lambda_i}{n} w_{\theta+1-i} + o(1)\Big\}\Big). \tag{6.110}$$

The result for the AB-case then follows by arguing as in (6.73) and [13, Lemma 3.4].

For the WB-case, a very similar argument as for (6.104) gives

$$Z_n^{\mathrm{WB}}(\beta, h) = \sum_{(\pi/n, \tau/n, [\lambda,\mu]/n) \in \Omega_{m/n}^{-}} \chi_{[\lambda,\mu]}^{GL(\theta)}(e^{hw_1}, \dots, e^{hw_\theta}) \exp\Big(n\Big\{\tilde{G}(\tfrac{\mu}{n}, \tfrac{\nu}{n}, \tfrac{\lambda}{n}) + o(1)\Big\}\Big), \tag{6.111}$$

where $\tilde{G}$ is given in (6.93), Ω_ρ^- is defined just above (6.93), and $\chi_{[\lambda,\mu]}^{GL(\theta)}$ is given in (6.82). In particular, from (6.82), we see that upper and lower bounds from (6.107) and (6.109) extend to this case. The result for the WB-case then follows by arguing as in (6.95) and [13, Lemma 3.4] again. ∎

Proof of Theorem 6.1.8. The proof closely follows that of Theorem 4.1 from [13]. We start from the expressions (6.103) where, for ease of notation, we drop the superscript. We give details only in the AB-case with $h > 0$ as the other cases are very similar.

Let $F_{\max} = \Phi(\beta, 0) = \max_{\vec{z} \in \Delta^+}\big(\max_{(X,Y) \in \mathcal{H}_\rho^+(\vec{z})} \phi(X, Y)\big)$ and let

$$K = \Big\{\vec{z} \in \Delta^+ : \max_{(X,Y) \in \mathcal{H}_\rho^+(\vec{z})} \phi(X, Y) = F_{\max}\Big\} \tag{6.112}$$

denote the set of maximisers. Clearly,

$$\begin{aligned}\frac{\Phi(\beta,h)-\Phi(\beta,0)}{h} &= \max_{\vec{z}\in\Delta^+}\Big[\sum_{i=1}^{\theta} z_i w_i + \frac{\max_{(X,Y)\in\mathcal{H}_\rho^+(\vec{z})}\phi(X,Y)-F_{\max}}{h}\Big] \\ &\geq \max_{\vec{z}\in K}\sum_{i=1}^{\theta} z_i w_i.\end{aligned} \tag{6.113}$$

We want to prove that the left-hand side of (6.113) tends to the right-hand side as $h\to 0$. For a contradiction, assume that there is a sequence $h_n\to 0$ such that the corresponding limit exists and is strictly larger than the right-hand side. For each h_n, pick an element $\vec{z}(h_n)\in\Delta^+$ that achieves the first maximum in (6.113). Since Δ^+ is compact, we can assume after passing to a subsequence if necessary that $\vec{z}(h_n)\to\vec{z}^\star$ as $h_n\to 0$. We claim that $\vec{z}^\star\in K$. Otherwise, $\max_{(X,Y)\in\mathcal{H}_\rho^+(\vec{z}^\star)}\phi(X,Y)<F_{\max}$, which would mean that the left-hand side of (6.113) tends to $-\infty$ as $h=h_n\to 0$, contradicting the lower bound on the right. It follows that

$$\begin{aligned}\frac{\Phi(\beta,h_n)-\Phi(\beta,0)}{h_n} &= \sum_{i=1}^{\theta} z_i(h_n)w_i + \frac{\max_{(X,Y)\in\mathcal{H}_\rho^+(\vec{z}(h_n))}\phi(X,Y)-F_{\max}}{h_n} \\ &\leq \sum_{i=1}^{\theta} z_i(h_n)w_i \to \sum_{i=1}^{\theta} z_i^\star w_i \leq \max_{\vec{z}\in K}\sum_{i=1}^{\theta} z_i^\star w_i,\end{aligned} \tag{6.114}$$

as required.

In the WB-case, we follow the same reasoning but with Δ^+ replaced by Δ_ρ^-, with $\mathcal{H}_\rho^+$ replaced by $\mathcal{H}_\rho^-$, and the maxima in (6.113) replaced by minima (as well as $w_i\leftrightarrow w_{\theta+1-i}$).

It remains to show that the z_i may be expressed as in the statement of the Theorem. Indeed, we know from (6.79) that $\phi(X,Y)$ is maximised when X and Y are simultaneously diagonal, with entries $x_1,\dots,x_\theta$ and $y_1,\dots,y_\theta$, respectively, ordered as follows:

- if $c>0$, if $x_1\geq\cdots\geq x_\theta\geq 0$ then $y_1\geq\cdots\geq y_\theta\geq 0$;
- if $c<0$, if $x_1\geq\cdots\geq x_\theta\geq 0$ then $0\leq y_1\leq\cdots\leq y_\theta$.

This gives the result. ∎

6.3 The phase-transition

In this section we prove Propositions 6.1.3, 6.1.4, 6.1.5 and 6.1.6. Let us start by recalling the basic quantities of interest: we wish to maximize the function

$$F(\omega)=F(\vec{x};\vec{y})=\textstyle\sum_{i=1}^{\theta} f(x_i,y_i), \tag{6.115}$$

over the domain

$$\Omega=\Big\{\omega=(\vec{x};\vec{y}):x_1,\dots,x_\theta,y_1,\dots,y_\theta\geq 0,\ \textstyle\sum_{i=1}^{\theta} x_i=1-\sum_{i=1}^{\theta} y_i=\rho\Big\}. \tag{6.116}$$

Here

$$f(x,y)=-x\log x-y\log y+\tfrac{\beta}{2}\big(ax^2+by^2+2cxy\big), \tag{6.117}$$

and we write $Q(x,y) = \frac{1}{2}\big(ax^2 + by^2 + 2cxy\big)$ for the quadratic form appearing in $f(x,y)$. In this section we will write $\rho' = 1 - \rho$ to lighten the notation.

We are particularly interested in whether the maximum of F is attained at the point

$$\omega_0 = \big(\tfrac{\rho}{\theta}, \tfrac{\rho}{\theta}, \dots, \tfrac{\rho}{\theta}; \tfrac{\rho'}{\theta}, \tfrac{\rho'}{\theta}, \dots, \tfrac{\rho'}{\theta}\big), \tag{6.118}$$

or at some other point in Ω. We defined β_{c} to be the supremume of those values of β for which F is maximised at ω_0.

6.3.1 Existence of a phase transition: proof of Proposition 6.1.3

We start with two elementary lemmas about quadratic forms.

Lemma 6.3.1. *If Q is a quadratic form of two variables, then*

$$\theta \sum_{j=1}^{\theta} Q(x_j, y_j) = Q(x_1 + \dots + x_\theta, y_1 + \dots + y_\theta) + \sum_{1 \le i < j \le r} Q(x_j - x_i, y_j - y_i). \tag{6.119}$$

Proof. When $Q(x,y) = xy$ we need to prove that

$$\theta \sum_{j=1}^{\theta} x_j y_j = (x_1 + \dots + x_\theta)(y_1 + \dots + y_\theta) + \sum_{1 \le i < j \le \theta} (x_j - x_i)(y_j - y_i). \tag{6.120}$$

This is easy to see by comparing the coefficient of each monomial on the two sides. Specializing $x_j = y_j$ proves the result for $Q(x,y) = x^2$ and $Q(x,y) = y^2$, and the general case then follows by linearity. ∎

Lemma 6.3.2. *Assume that $Q(x,y) = \frac{1}{2}(ax^2 + by^2 + 2cxy)$ is not negative semidefinite and that β, A, $B > 0$. Then, the form*

$$\beta Q(x,y) - \tfrac{1}{2}\big(Ax^2 + By^2\big) \tag{6.121}$$

is negative semidefinite if and only if $\beta \le \beta_0$, and negative definite if and only if $\beta < \beta_0$, where β_0 is the smallest positive solution to the equation

$$(\beta a - A)(\beta b - B) = \beta^2 c^2 \tag{6.122}$$

or, more explicitly,

$$\beta_0 = \begin{cases} \dfrac{aB + bA - \sqrt{(aB - bA)^2 + 4c^2 AB}}{2(ab - c^2)}, & ab \ne c^2, \\[2ex] \dfrac{AB}{aB + bA}, & ab = c^2. \end{cases} \tag{6.123}$$

Proof. By assumption, the first term in (6.121) can assume positive values, and the second term is always non-positive. It follows that the range of β for which (6.121) is negative semidefinite is of the form $\beta \le \beta_0$ and that it is negative definite if and only if $\beta < \beta_0$. The precise conditions for (6.121) to be negative semidefinite are

$$(\beta a - A)(\beta b - B) \ge \beta^2 c^2, \qquad \beta a \le A, \qquad \beta b \le B. \tag{6.124}$$

By continuity, β_0 solves the equation (6.122). If $ab = c^2$, this is a linear equation with a unique solution. Otherwise, it has two solutions

$$\beta_\pm = \frac{aB + bA \pm \sqrt{(aB - bA)^2 + 4c^2 AB}}{2(ab - c^2)}, \tag{6.125}$$

which satisfy $(ab - c^2)\beta_+\beta_- = AB > 0$. If $ab > c^2$, both solutions are positive and β_0 equals the smallest solution β_-. If $ab < c^2$ the solutions have opposite sign. In this case β_0 is the largest solution, which is again β_-. ∎

We are now ready to prove our result on the existence of a critical point. Recall that we want to prove that β_c exists (is positive and finite) if and only if Q is not negative semidefinite, where β_c is the supremum of the β for which ω_0 is a maximiser of F.

Proof of Proposition 6.1.3. We can write

$$F(\omega) - F(\omega_0) = \beta\mathcal{E}(\omega) + \mathcal{H}(\omega), \tag{6.126}$$

where

$$\mathcal{H}(\vec{x};\vec{y}) = \sum_{j=1}^{\theta}(-x_j \log x_j - y_j \log y_j) + \rho \log \tfrac{\rho}{\theta} + \rho' \log \tfrac{\rho'}{\theta}, \tag{6.127}$$

and

$$\mathcal{E}(\vec{x};\vec{y}) = \sum_{j=1}^{\theta} Q(x_j, y_j) - rQ\big(\tfrac{\rho}{\theta}, \tfrac{\rho'}{\theta}\big). \tag{6.128}$$

Then, F is maximized at ω_0 if and only if $\beta\mathcal{E}(\omega) + \mathcal{H}(\omega) \le 0$ on Ω.

On Ω, we can write

$$\tfrac{1}{\theta}\mathcal{H}(\vec{x};\vec{y}) = -h\left(\frac{x_1 + \cdots + x_\theta}{\theta}\right) + \frac{h(x_1) + \cdots + h(x_\theta)}{\theta} - h\left(\frac{y_1 + \cdots + y_\theta}{\theta}\right) + \frac{h(y_1) + \cdots + h(y_\theta)}{\theta},$$

where $h(x) = -x \log x$. Since h is strictly concave, $\mathcal{H}(\omega) \le 0$ with equality only at the point ω_0. Moreover, by Lemma 6.3.1,

$$\mathcal{E}(\vec{x};\vec{y}) = \frac{1}{\theta} \sum_{1 \le i < j \le \theta} Q(x_j - x_i, y_j - y_i). \tag{6.129}$$

Thus, if Q is negative semidefinite, we have $\mathcal{E}(\omega) \le 0$ and consequently ω_0 is the unique maximum point of F.

Assume now that Q is not negative semidefinite. We claim that $\mathcal{E}$ assumes strictly positive values in Ω. To see this, it suffices to consider the case when $x_2 = \cdots = x_\theta$, $y_2 = \cdots = y_\theta$. Then

$$\mathcal{E}(\vec{x};\vec{y}) = \frac{\theta - 1}{\theta} Q(\xi, \eta), \tag{6.130}$$

where $\xi = x_2 - x_1$ and $\eta = y_2 - y_1$. Here (ξ, η) can take any value in $\big[-\rho, \frac{\rho}{\theta-1}\big] \times \big[-\rho', \frac{\rho'}{\theta-1}\big]$. By assumption, Q assumes positive values in parts of this rectangle. Then it is clear that $\mathcal{E}$ takes positive values, hence that $\mathcal{H}(\omega) + \beta\mathcal{E}(\omega)$ assumes positive values for β large enough, and that the set of $\beta > 0$ for which this is true is an interval $\beta > \beta_c$. To see that

ω_0 is the unique maximiser for $\beta < \beta_c$, take $\omega \in \Omega \setminus \{\omega_0\}$. Then either $\mathcal{E}(\omega) > 0$, in which case $\mathcal{H}(\omega) + \beta\mathcal{E}(\omega) < \mathcal{H}(\omega) + \beta_c\mathcal{E}(\omega) \le 0 = \mathcal{H}(\omega_0) + \beta\mathcal{E}(\omega_0)$, or $\mathcal{E}(\omega) \le 0$, in which case $\mathcal{H}(\omega) + \beta\mathcal{E}(\omega) \le \mathcal{H}(\omega) < 0 = \mathcal{H}(\omega_0) + \beta\mathcal{E}(\omega_0)$.

It remains to show that $\beta_c \ne 0$, that is, that F assumes its maximum value at ω_0 for β close to zero. We will show that this is in fact true if we maximize F over the larger set

$$U = \big\{(\vec{x};\vec{y}) : 0 \le x_j \le \rho,\, 0 \le y_j \le \rho',\, j = 1,\dots,\theta\big\}. \tag{6.131}$$

To do this we will show that the Hessian $H(F)$ is negative definite in U for β close to 0, meaning that F is concave in U for such β and that ω_0 is a global maximum in U. The Hessian $H(F)$ is a direct sum of the Hessians

$$H(f) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} = \begin{pmatrix} \beta a - \frac{1}{x} & \beta c \\ \beta c & \beta b - \frac{1}{y} \end{pmatrix}, \tag{6.132}$$

which is negative definite if and only if

$$\big(\beta a - \tfrac{1}{x}\big)\big(\beta b - \tfrac{1}{y}\big) > \beta^2 c^2, \qquad \tfrac{1}{x} > \beta a, \qquad \tfrac{1}{y} > \beta b. \tag{6.133}$$

By monotonicity, when $x \le \rho$ and $y \le \rho'$ the inequalities (6.133) are implied by

$$\big(\beta a - \tfrac{1}{\rho}\big)\big(\beta b - \tfrac{1}{\rho'}\big) > \beta^2 c^2, \qquad \tfrac{1}{\rho} > \beta a, \qquad \tfrac{1}{\rho'} > \beta b. \tag{6.134}$$

But (6.134) holds for $\beta = 0$, hence by continuity also for small β, as required. ∎

From the proof above we note that $\beta \le \beta_c$ if and only if $\mathcal{H}(\omega) + \beta\mathcal{E}(\omega) \le 0$ for all $\omega \in \Omega$, and secondly that we have the expression

$$\beta_c = \inf_{\omega\in\Omega^+}\Big(-\frac{\mathcal{H}(\omega)}{\mathcal{E}(\omega)}\Big), \quad \text{where } \Omega^+ = \big\{\omega \in \Omega : \mathcal{E}(\omega) > 0\big\}. \tag{6.135}$$

6.3.2 Formulas for β_c: proofs of Propositions 6.1.4 and 6.1.5

We now turn to the proofs of our formulas for β_c, Propositions 6.1.4 for the case $\theta = 2$ and 6.1.5 for the case $\theta \ge 3$, $c \ge 0$ and $(a-c)\rho = (b-c)\rho' =: t$ (6.14).

Our strategy is to obtain general lower and upper bounds on $\beta_c(\theta)$, given in Propositions 6.3.3 and 6.3.4 respectively, which are tight in the two cases that we consider. Both bounds are given in terms of the critical temperature $\beta_c^{\mathrm{h}}(\theta)$ of the homogeneous case $a = b = c = 1$; here $Q(x,y) = \frac{1}{2}(x+y)^2$ is not negative semidefinite and (6.14) holds with $t = 0$. This gives the result [13, Theorem 4.2]

$$\beta_c^{\mathrm{h}}(\theta) = \begin{cases} 2, & \theta = 2, \\ \dfrac{2(\theta-1)\log(\theta-1)}{\theta-2}, & \theta \ge 3, \end{cases} \tag{6.136}$$

where the superscript h is for 'homogeneous' and is reserved for the case $a = b = c = 1$.

For the case $\theta = 2$, it is useful to note that the formula (6.18) for $\beta_c(2)$ is the smallest positive solution to (6.122) with $A = 2/\rho$ and $B = 2/\rho'$. To get a better understanding of Proposition 6.1.5, i.e. the case $\theta \ge 3$, recall our condition (6.14) which says that $(a-c)\rho =$

$(b-c)\rho' =: t$. This condition implies the explicit diagonalization

$$Q(x,y) = \frac{t\rho\rho'}{2}\left(\frac{x}{\rho} - \frac{y}{\rho'}\right)^2 + \frac{c+t}{2}(x+y)^2. \tag{6.137}$$

That Q is not negative semidefinite means that at least one of t and $c+t$ are positive, or equivalently that either $c \geq 0$ and $c+t > 0$ or $c < 0$ and $t > 0$. This shows that the expression for $\beta_{\mathrm{c}}(\theta)$ in Proposition 6.1.5 is always positive.

Let us now obtain the lower bound for $\beta_{\mathrm{c}}(\theta)$. We deduce from (6.135) and [13, Theorem 4.2] with $\rho = 1$ that $-\mathcal{H}(\vec{x};\vec{0}) \geq \beta_{\mathrm{c}}^{\mathrm{h}}(\theta)\mathcal{E}(\vec{x};\vec{0})$. This inequality takes the form

$$\sum_{j=1}^{\theta} x_j \log x_j - \log \tfrac{1}{\theta} \geq \tfrac{\beta_{\mathrm{c}}^{\mathrm{h}}(\theta)}{2\theta} \sum_{1\leq i<j\leq\theta} (x_j - x_i)^2, \quad \text{where } \textstyle\sum_{j=1}^{\theta} x_j = 1. \tag{6.138}$$

Replacing each x_j by x_j/ρ gives

$$\sum_{j=1}^{\theta} x_j \log x_j - \rho\log \tfrac{\rho}{\theta} \geq \tfrac{\beta_{\mathrm{c}}^{\mathrm{h}}(\theta)}{2\rho\theta} \sum_{1\leq i<j\leq\theta} (x_j - x_i)^2, \quad \text{where } \textstyle\sum_{j=1}^{\theta} x_j = \rho. \tag{6.139}$$

As was observed in [13], equality in (6.139) holds both at the point $x_1 = \cdots = x_\theta = \rho/\theta$ and at (6.20a). (These are the same point if $\theta = 2$.)

Proposition 6.3.3. *Assume that Q is not negative semidefinite, so that β_{c} exists. Then,*

$$\beta_{\mathrm{c}}(\theta) \geq \tfrac{1}{2}\beta_{\mathrm{c}}^{h}(\theta)\beta_{\mathrm{c}}(2), \tag{6.140}$$

where $\beta_{\mathrm{c}}^{h}(\theta)$ denotes the expression (6.136) and $\beta_{\mathrm{c}}(2)$ the expression (6.18).

Proof. Using the estimate (6.139) in (6.127) gives

$$-\mathcal{H}(\omega) \geq \frac{\beta_{\mathrm{c}}^{\mathrm{h}}(\theta)}{2\theta} \sum_{1\leq i<j\leq\theta} \left(\frac{(x_i - x_j)^2}{\rho} + \frac{(y_i - y_j)^2}{\rho'}\right). \tag{6.141}$$

It follows that

$$\mathcal{H}(\omega) + \beta\mathcal{E}(\omega) \leq \frac{1}{\theta} \sum_{1\leq i<j\leq\theta} \tilde{Q}(x_j - x_i, y_j - y_i), \tag{6.142}$$

where

$$\tilde{Q}(x,y) = \beta Q(x,y) - \tfrac{\beta_{\mathrm{c}}^{\mathrm{h}}(\theta)}{2}\left(\tfrac{x^2}{\rho} + \tfrac{y^2}{\rho'}\right). \tag{6.143}$$

Recall that $\beta_{\mathrm{c}}(2)$ is the smallest positive solution to (6.122) with $A = 2/\rho$ and $B = 2/\rho'$. Thus by Lemma 6.3.2, if $\beta \leq \frac{1}{2}\beta_{\mathrm{c}}^{\mathrm{h}}(\theta)\beta_{\mathrm{c}}(2)$, then $\tilde{Q}$ is negative semidefinite and $\mathcal{H}(\omega) + \beta\mathcal{E}(\omega) \leq 0$ on Ω. This gives the desired bound on β_{c}. ∎

Let us now move to upper bounds for $\beta_{\mathrm{c}}(\theta)$. We need to find a value of β such that $F(\omega) > F(\omega_0)$ for some points $\omega \in \Omega$. We want to find upper bounds that in some case equal the lower bound in Proposition 6.3.3. We can only expect this to work if we used the inequality (6.139) in cases when it holds with equality. By the results of [13] mentioned above, it is natural to take ω either close to ω_0, or ω_1 as in (6.20). This leads to the following two upper bounds.

Proposition 6.3.4. *Assume that Q is not negative semidefinite, so that β_c exists. Then,*

$$\beta_c(\theta) \leq \tfrac{1}{2}\theta\beta_c(2). \tag{6.144}$$

If, in addition, $Q(\rho,\rho') > 0$ and $\theta \geq 3$, then

$$\beta_c(\theta) \leq \frac{\beta_c^h(\theta)}{2Q(\rho,\rho')}. \tag{6.145}$$

In fact, (6.145) holds also when $\theta = 2$, but in that case it is weaker than (6.144) and will not be needed.

Proof. We first consider the behaviour of F near ω_0. More precisely, consider the points

$$\omega_{t,u} = \omega_0 + (t, -t, 0, \dots, 0; u, -u, 0, \dots, 0), \tag{6.146}$$

which belong to Ω for t, u close to 0. We have the Taylor expansion

$$\begin{aligned} F(\omega_{t,u}) - F(\omega_0) &= f\big(\tfrac{\rho}{\theta} + t, \tfrac{\rho'}{\theta} + u\big) + f\big(\tfrac{\rho}{\theta} - t, \tfrac{\rho'}{\theta} - u\big) - 2f\big(\tfrac{\rho}{\theta}, \tfrac{\rho'}{\theta}\big) \\ &= \big(t^2 f_{xx} + u^2 f_{yy} + 2tu f_{xy}\big)\big(\tfrac{\rho}{\theta}, \tfrac{\rho'}{\theta}\big) + \mathcal{O}((t^2+u^2)^{3/2}). \end{aligned}$$

By (6.132), the quadratic term is

$$2\beta Q(t,u) - \theta\big(\tfrac{t^2}{\rho} + \tfrac{u^2}{\rho'}\big). \tag{6.147}$$

By Lemma 6.3.2, if $\beta > \theta\beta_c(2)/2$, this form is not negative semidefinite. It follows that ω_0 is not a local maximum of F. This gives the first result.

Next, we consider the point ω_1 from (6.20) and assume $\theta \geq 3$. By a straightforward computation,

$$\mathcal{H}(\omega_1) = -\tfrac{\theta-2}{\theta}\log(\theta-1) \tag{6.148}$$

and, by (6.130),

$$\mathcal{E}(\omega_1) = \tfrac{\theta-1}{\theta} Q\big(\tfrac{\rho(\theta-2)}{\theta-1}, \tfrac{\rho'(\theta-2)}{\theta-1}\big) = \tfrac{(\theta-2)^2}{\theta(\theta-1)} Q(\rho,\rho'). \tag{6.149}$$

The second upper bound now follows from (6.135). ∎

We can now put our upper and lower bounds together to prove Propositions 6.1.4 and 6.1.5.

Proof of Proposition 6.1.4. When $\theta = 2$, (6.140) and (6.144) reduce to $\beta_c(2) \leq \beta_c \leq \beta_c(2)$ (where β_c is the critical point and $\beta_c(2)$ the explicit expression (6.18)). This proves the formula for β_c. For the statement about uniqueness of the maximiser, note that if $\beta = \beta_c(2) = \frac{1}{2}\beta_c^h(2)\beta_c(2)$ and $\omega = (\vec{x};\vec{y})$ is a maximiser, then the left-hand-side of (6.142) equals zero. Then also the right-hand-side of (6.142) equals zero, since $\tilde{Q} \leq 0$ for $\beta \leq \frac{1}{2}\beta_c^h(2)\beta_c(2)$ by the proof of Proposition 6.3.3. Hence (6.141) holds with equality and therefore (6.139) holds with equality, as does the corresponding statement for $\vec{y}$. But it follows from the proof of Theorem 4.2 in [13] that (for $\theta = 2$) equality in (6.139) holds only at the point ω_0. ∎

Proof of Proposition 6.1.5. Note that the lower bound in (6.140) and the upper bound in (6.145) are equal if $\beta_c(2) = Q(\rho,\rho')^{-1}$. We need to check that this is implied by $(a-c)\rho = (b-c)\rho'$, which is (6.14) (in fact, it also implies (6.14)). Assuming (6.14), we can parametrize

$$a = c + \tfrac{t}{\rho}, \qquad b = c + \tfrac{t}{\rho'}. \tag{6.150}$$

It is then straight-forward to check that

$$(\rho a - \rho' b)^2 + 4\rho\rho' c^2 = c^2, \quad \text{and} \quad ab - c^2 = \tfrac{t(c+t)}{\rho\rho'}, \tag{6.151}$$

which gives

$$\beta_c(2) = \frac{2t + c - \sqrt{c^2}}{t(c+t)} = \begin{cases} \dfrac{2}{c+t}, & c \ge 0, \\ \dfrac{2}{t}, & c < 0. \end{cases} \tag{6.152}$$

By (6.137),

$$Q(\rho,\rho') = \frac{c+t}{2}, \tag{6.153}$$

which shows the expression $\beta_c(\theta) = \beta_c^h(\theta)/2Q(\rho,\rho')$ when $\theta \ge 3$ (and $c \ge 0$).

To see that the point ω_1 in (6.20) gives another maximiser at $\beta = \beta_c$, take $\beta = \beta_c(\theta) = \beta_c^h(\theta)/2Q(\rho,\rho')$ to see from (6.148) and (6.149) that $\mathcal{H}(\omega_1) + \beta\mathcal{E}(\omega_1) = 0$ which is also the maximum value of $\mathcal{H}(\omega) + \beta\mathcal{E}(\omega)$. To see that ω_1 is the only other maximiser we argue as at the end of the proof of Proposition 6.1.4. Namely, for $\beta = \beta_c(\theta) = \frac{1}{2}\beta_c^h(\theta)\beta_c(2)$, we have that (6.139) holds with equality, as does the corresponding statement for $\vec{y}$. From [13], equality in (6.139) holds only at the points ω_0 and ω_1. ∎

We can now complete the final proof of this section, that of Proposition 6.1.6, that the maximiser is unique for $\beta > \beta_c$ close to β_c under the conditions in Proposition 6.1.5, that is, $\theta \ge 3$, $c \ge 0$, $(a-c)\rho - (b-c)\rho'$ (6.14) and Q not negative semidefinite. For this we use that there are two maximum points at β_c and that they are local maxima.

Proof of Proposition 6.1.6. We first show that F is strictly concave in neighbourhoods of ω_0 and ω_1 in Ω. More generally, consider $F(\vec{x} + \vec{t}; \vec{y} + \vec{u})$, where $(\vec{x};\vec{y}) \in \Omega$ is a point with $x_2 = \cdots = x_\theta$ and $y_2 = \cdots = y_\theta$ and $(\vec{t};\vec{u})$ a small perturbation with

$$\sum_{j=1}^{\theta} t_j = \sum_{j=1}^{\theta} u_j = 0. \tag{6.154}$$

By (6.132), the quadratic term in the Taylor expansion of F is

$$Q_1(t_1,u_1) + \sum_{j=2}^{\theta} Q_2(t_j,u_j), \tag{6.155}$$

where

$$Q_k(t,u) = \beta Q(t,u) - \frac{t^2}{2x_k} - \frac{u^2}{2y_k}.$$

At the point ω_0, we have

$$Q_1(t,u) = Q_2(t,u) = \beta Q(t,u) - \Big(\frac{\theta t^2}{2\rho} + \frac{\theta u^2}{2\rho'}\Big).$$

It follows from Lemma 6.3.2 that this is negative definite if $\beta < \beta_0 = \theta\beta_c(2)/2$. By continuity, it follows that F is strictly concave near ω_0. Since ω_0 is a stationary point it must then be a local maximum, that is, $F(\vec{x};\vec{y}) \le F(\omega_0) = 0$ for $(\vec{x};\vec{y})$ near ω_0 and $\beta < \beta_0$. Using that

$$\beta_c = \frac{(\theta-1)\log(\theta-1)}{\theta-2}\,\beta_c(2),$$

it is easy to check that $\beta_c < \beta_0$, so this applies in particular to β near β_c.

The point ω_1 cannot be handled as easily since Q_1 is then not negative definite. Instead, we use Lemma 6.3.1 and (6.154) to write

$$(\theta-1)\sum_{j=2}^{\theta} Q_2(t_j,u_j) = Q_2(t_1,u_1) + \sum_{2\le i<j\le r} Q_2(t_i-t_j,u_i-u_j).$$

It follows that (6.155) equals

$$Q_1(t_1,u_1) + \frac{1}{\theta-1}\,Q_2(t_1,u_1) + \frac{1}{\theta-1}\sum_{2\le i<j\le\theta} Q_2(t_i-t_j,u_i-u_j).$$

We compute

$$Q_1(t,u) + \frac{1}{\theta-1}\,Q_2(t,u) = \frac{\theta}{\theta-1}\left(\beta Q(t,u) - \left(\frac{\theta t^2}{2\rho} + \frac{\theta u^2}{2\rho'}\right)\right).$$

As before, this is negative definite for $\beta < \beta_0$. Moreover,

$$Q_2(t,u) = \beta Q(t,u) - \frac{\theta(\theta-1)t^2}{2\rho} - \frac{\theta(\theta-1)u^2}{2\rho'}$$

is negative definite for $\beta < (\theta-1)\beta_0$, which is a weaker condition. We conclude that F is strictly concave for $\beta < \beta_0$ and $(\vec{x};\vec{y})$ near ω_1. We also note that

$$F(\omega_1) = \mathcal{H}(\omega_1) + \beta_c\mathcal{E}(\omega_1) + (\beta-\beta_c)\mathcal{E}(\omega_1),$$

where the sum of the first two terms vanish and the last term is computed by (6.149) and (6.153). This gives

$$F(\omega_1) = (\beta-\beta_c)\frac{(\theta-2)^2(c+t)}{2\theta(\theta-1)},$$

which is clearly positive for $\beta > \beta_c$.

For each $\beta > \beta_c$, let $\omega(\beta)$ be a maximiser of F in Ω. Permute the coordinates so that (6.9) holds. We claim that then $\omega(\beta) \to \omega_1$ as $\beta \searrow \beta_c$. Otherwise, there exists a sequence $\omega(\beta_n)$, $\beta_n \searrow \beta_c$, that avoids a neighbourhood of ω_1. Since Ω is compact we may assume that this sequence converges. It must then converge to a maximiser of F for $\beta = \beta_c$ that satisfies (6.9). There are only two such points, ω_0 and ω_1, by Proposition 6.1.5. However, we have seen that for $\beta_c < \beta < \beta_0$ we have $F(\vec{x};\vec{y}) \le 0$ for $(\vec{x};\vec{y})$ near ω_0 whereas $F(\omega_1) > 0$. Thus, a sequence of global maximisers cannot converge to ω_0. This is a contradiction, and we conclude that $\omega(\beta) \to \omega_1$. These points must then enter a region where F is strictly concave and hence maximisers are unique. This completes the proof. ■

6.4 Multi-block models

Here we prove Theorem 6.1.9. The proof follows a similar pattern to that of Theorem 6.1.1. We start by writing

$$H_n^{\mathrm{MB}} = -nT\Big(\sum_{\gamma\in\Gamma}\Big[\sum_{k=1}^{p} a_k^\gamma \alpha_{A_k}^\gamma + c^\gamma \alpha_n^\gamma\Big]\Big) = -n\sum_{\gamma\in\Gamma}\Big[\sum_{k=1}^{p} a_k^\gamma T(\alpha_{A_k}^\gamma) + c^\gamma T(\alpha_n^\gamma)\Big]\Big), \tag{6.156}$$

where T is the representation of $\mathbb{C}[S_n]$ on $\mathbb{V}$ given in (6.37), and

$$\alpha_{A_k}^\gamma = \frac{1}{|C_{A_k}^\gamma|}\sum_{\sigma\in C_{A_k}^\gamma}\sigma \in \mathbb{C}[S_{A_k}], \qquad \alpha_n^\gamma = \frac{1}{|C_n^\gamma|}\sum_{\sigma\in C_n^\gamma}\sigma\in\mathbb{C}[S_n]. \tag{6.157}$$

As in (6.54) we have a decomposition

$$\mathbb{V} \cong \bigoplus_{\lambda\vdash n,\ell(\lambda)\le\theta} d_\lambda^{GL(\theta)}\psi_\lambda^{S_n}. \tag{6.158}$$

Here we consider $\mathbb{V}$ as an $\mathbb{C}[S_n]$-module only (we do not need the $GL(\theta)$-part since we consider only the free energy and not correlations). As a $\mathbb{C}[S_{m_1}\times\cdots\times S_{m_p}]$-module, we have the decomposition

$$\psi_\lambda^{S_n} \cong \bigoplus_{\mu(1),\dots,\mu(p)} c^\lambda_{\mu(1),\dots,\mu(p)}\psi_{\mu(1)}^{S_{m_1}}\otimes\cdots\otimes\psi_{\mu(p)}^{S_{m_p}}, \tag{6.159}$$

which generalizes (6.60). Here $\mu(k)\vdash m_k$ for each k and the multiplicities $c^\lambda_{\mu(1),\dots,\mu(p)}$ are analogs of the Littlewood–Richardson coefficients $c^\lambda_{\mu,\nu}$ and have many similar properties. In particular, a full analog of Horn's inequalities holds: $c^\lambda_{\mu(1),\dots,\mu(p)}>0$ if and only if there are Hermitian matrices $M(1),\dots,M(p)$ with spectra $\mu(1),\dots,\mu(p)$ such that $M(1)+\cdots+M(p)$ has spectrum λ (see Theorem 17 of [41]).

Let us next see how $T(\alpha_{A_k}^\gamma)$ and $T(\alpha_n^\gamma)$ act on these subspaces $\psi_{\mu(k)}^{S_{m_k}}$. For $m\le n$ and $C=C_m^\gamma$ the conjugacy class of γ in S_m, consider $\alpha=\frac{1}{|C|}\sum_{\sigma\in C}\sigma\in\mathbb{C}[S_m]$. For $\mu\vdash m$, since α is central in $\mathbb{C}[S_m]$, it acts on the irreducible $\psi_\mu^{S_m}$ as a scalar, and in fact we have

$$\alpha|_{\psi_\mu^{S_m}} = \frac{\chi_\mu^{S_m}(\alpha)}{d_\mu^{S_m}}\mathtt{Id}_{\psi_\mu^{S_m}} = \frac{\chi_\mu^{S_m}(\gamma)}{d_\mu^{S_m}}\mathtt{Id}_{\psi_\mu^{S_m}}, \tag{6.160}$$

where $\chi_\mu^{S_m}(\gamma)$ is the character of $\psi_\mu^{S_m}$ evaluated at any permutation of cycle-type γ. This leads to the following expression analogous to (6.65):

$$\begin{aligned} Z_n^{\mathrm{MB}} = \sum_{\lambda\vdash n,\ell(\lambda)\le\theta} d_\lambda^{GL(\theta)} &\sum_{\mu(1),\dots,\mu(p)} c^\lambda_{\mu(1),\dots,\mu(p)} d_{\mu(1)}^{S_{m_1}}\cdots d_{\mu(p)}^{S_{m_1}} \\ &\cdot\exp\Big(n\beta\sum_{\gamma\in\Gamma}\Big[\sum_{k=1}^{p} a_k^\gamma\frac{\chi_{\mu(k)}^{S_{m_k}}(\gamma)}{d_{\mu(k)}^{S_{m_k}}} + c^\gamma\frac{\chi_\lambda^{S_n}(\gamma)}{d_\lambda^{S_n}}\Big]\Big).\end{aligned} \tag{6.161}$$

As before, the relevant scaling for the limit $\lim_{n\to\infty}\frac{1}{n}\log Z_n^{\mathrm{MB}}$ is given by letting $\lambda/n\to\vec{z}$ and $\mu(k)/n\to\vec{x}(k)$ for all k. Also as before, $d_\lambda^{GL(\theta)}$ is negligible on the relevant scale, and the $d_{\mu(k)}^{S_{m_k}}$ obey the asymptotics of (6.68). Below, we prove that $c^\lambda_{\mu(1),\dots,\mu(p)}\le(n+1)^{p\theta^2}$ which is also too small to contribute to the limit.

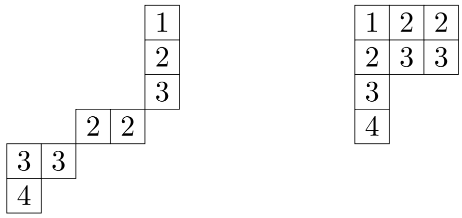


Figure 6.6: Left: A skew tableau with shape ν formed from the three partitions $\mu(1) = (2,1)$, $\mu(2) = (2)$ and $\mu(3) = (1,1,1)$. Right: its rectification.

What remains is to identify the limits of the expressions of the form $\frac{\chi_\mu^{S_m}(\gamma)}{d_\mu^{S_m}}$. The latter limits are well-known in the asymptotic representation theory of the symmetric group: Thoma's Theorem and the Vershik–Kerov Theorem (see e.g. [18, Corollary 4.2 and Theorem 6.16]) imply that if $\mu/n \to \vec{x} = (x_1, \ldots, x_\theta)$, then

$$\frac{\chi_\mu^{S_m}(\gamma)}{d_\mu^{S_m}} \to p_\gamma(x_1, \ldots, x_\theta), \tag{6.162}$$

where $p_\gamma(\cdot)$ is the power-sum symmetric polynomial given in (6.45). Writing $\vec{x}(k) = \lim_{n\to\infty} \mu(k)/n$ and $\vec{z} = \lim_{n\to\infty} \lambda/n$, we conclude that the contributing $\vec{x}(k)$ and $\vec{z}$ are eigenvalues of Hermitian matrices $X_1, \ldots, X_p$ and $Z = X_1 + \cdots + X_p$, respectively, where $\mathrm{Tr}[X_k] = \rho_k$. Re-writing the free energy in terms of these matrices, as in (6.78) and (6.95), we obtain the claim (6.41).

It remains to verify the bound $c^\lambda_{\mu(1),\ldots,\mu(p)} \le (n+1)^{p\theta^2}$. We use the following combinatorial description of $c^\lambda_{\mu_1,\ldots,\mu_p}$ which is mentioned just after Proposition 13 of [41]. Form a skew shape ν by stacking $\mu(1), \ldots, \mu(p)$ from bottom left to top right, such that the lower left corner of $\mu(k)$ just touches the upper right corner of $\mu(k-1)$ as in Figure 6.6. Fix any semistandard tableau τ_λ of shape λ, to be concrete let us say that the first row of τ_λ consists of λ_1 1's, the second row of λ_2 2's etc. Then $c^\lambda_{\mu(1),\ldots,\mu(p)}$ is the number of semistandard tableaux σ_ν of skew shape ν whose *rectification* equals τ_λ. For a full description of the rectification, see [40, Section 1.2], but in brief terms the rectification is obtained by 'sliding' the numbered boxes of σ_ν until a non-skew shape is obtained. To see the claimed bound, note that in order to obtain the tableau τ_λ, the number of boxes labelled 1 in ν must equal the number of boxes labelled 1 in λ, and similarly for labels 2, 3, etc. Thus, for each row of ν we have at most

$$(\lambda_1 + 1)(\lambda_2 + 1)\cdots(\lambda_\theta + 1) \le (n+1)^\theta$$

choices of entries (from 0 to λ_1 1's, from 0 to λ_2 2's etc). Since ν has at most $p\theta$ rows, the total number of choices is $\le [(n+1)^\theta]^{p\theta}$, as claimed. ∎

6.5 Form of the maximiser of F for $c > 0$

In this section we prove that for $c > 0$, the maximiser of F (6.6) is of the form (6.163). We assume thoughout this section that $\vec{x}$ is ordered as in (6.9), that is $x_1 \ge x_2 \ge \cdots \ge x_\theta$. Recall from the discussion after (6.9) that, for $c > 0$, F is maximised when the orders of $\vec{x}$ and $\vec{y}$ match, that is when also $y_1 \ge \cdots \ge y_\theta$. We will adapt the arguments in [13] and in the appendix of [14] to show the following.

Proposition 6.5.1. *For $c > 0$, the maximiser $(\vec{x}^\star;\vec{y}^\star)$ of F in the set Ω (6.15) is of the form*

$$\begin{aligned} x_1^\star &\geq x_2^\star = \cdots = x_\theta^\star, \\ y_1^\star &\geq y_2^\star = \cdots = y_\theta^\star. \end{aligned} \tag{6.163}$$

Moreover for the special case $a = b = 0$, $c > 0$, $\rho = 1/2$, and $\beta \neq \beta_c$ we have that the maximiser is unique, and $x_i^\star = y_i^\star$ for all $i = 1, \dots, \theta$.

The proof of this proposition is divided into several steps. We first prove that a maximum point $(\vec{x};\vec{y})$ only has positive coordinates, and that $x_j = x_k$ if and only if $y_j = y_k$ (this holds also for $c < 0$). Then we prove that, when $c > 0$, the entries x_i (and therefore y_i) can take at most two distinct values. This reduces the number of variables we need to consider, leading to (6.163) and the uniquenes statement via direct calculations.

Lemma 6.5.2. *For any $\theta \geq 2$ and $a, b, c \in \mathbb{R}$, if $(\vec{x};\vec{y})$ is a maximum point of F in Ω, then*

1. *all x_j and y_j are strictly positive,*
2. *$x_j = x_k$ if and only if $y_j = y_k$.*

Proof. In this proof we write e_j for the unit vector with a 1 in the x_j coordinate and remaning entries $= 0$. For the first part, suppose that $\omega = (\vec{x};\vec{y}) \in \Omega$ is a maximum point such that $x_j = 0$ for some j, and that j is the smallest index with this property. Then, $\omega(t) = \omega + t(e_j - e_{j-1}) \in \Omega$ for small enough $t > 0$. By a direct computation, $F(\omega(t)) - F(\omega) = -t\log t + O(t)$ as $t \to 0$. It follows that $F(\omega(t)) > F(\omega)$ for small t, which contradicts ω being a maximum point. The same argument works for the variables y_j.

For the second part, suppose that $x_j = x_k$ and $y_j \neq y_k$. If necessary, redefine j and k so that $\{l : x_l = x_k\} = \{j, j+1, \dots, k\}$. We still have $y_j \neq y_k$. Then $\omega(t) := (\vec{x};\vec{y}) + t(e_j - e_k) \in \Omega$ for small enough $t > 0$. (Here we use the first part of the lemma in the case $k = \theta$.) We have that $\frac{\partial}{\partial t}F(\omega(t))|_{t=0} = c(y_j - y_k) > 0$. This contradicts ω being a maximum point. The same argument proves the reverse implication. ■

Lemma 6.5.2 shows that at a maximum point there is a composition $\theta = k_1 + \cdots + k_m$ so that

$$(x_1^\star, \dots, x_\theta^\star) = (\underbrace{\xi_1, \dots, \xi_1}_{k_1}, \dots, \underbrace{\xi_m, \dots, \xi_m}_{k_m}), \tag{6.164a}$$

$$(y_1^\star, \dots, y_\theta^\star) = (\underbrace{\eta_1, \dots, \eta_1}_{k_1}, \dots, \underbrace{\eta_m, \dots, \eta_m}_{k_m}), \tag{6.164b}$$

where $\xi_j \neq \xi_k$ and $\eta_j \neq \eta_k$ for $j \neq k$. This leads to the problem of maximizing

$$\bar{F}(\xi;\eta) = k_1 f(\xi_1, \eta_1) + \cdots + k_m f(\xi_m, \eta_m) \tag{6.165}$$

over the set $\Omega^{(m)}$ defined by

$$\xi_1 > \xi_2 > \cdots > \xi_m > 0, \qquad k_1\xi_1 + \cdots + k_m\xi_m = \rho, \tag{6.166a}$$

$$\eta_1 > \eta_2 > \cdots > \eta_m > 0, \qquad k_1\eta_1 + \cdots + k_m\eta_m = 1 - \rho. \tag{6.166b}$$

For $m \geq 1$, the set $\Omega^{(m)}$ is open, so we may find local extreme points by using Lagrange multipliers. At any such point we have

$$\nabla \bar{F}(\xi;\eta) = \lambda\nabla(k_1\xi_1 + \cdots + k_m\xi_m) + \mu\nabla(k_1\eta_1 + \cdots + k_m\eta_m), \tag{6.167}$$

for some $\lambda, \mu \in \mathbb{R}$. Equivalently

$$\frac{\partial f}{\partial x}(\xi_i, \eta_i) = \lambda, \quad \frac{\partial f}{\partial y}(\xi_i, \eta_i) = \mu, \qquad 1 \leq i \leq m. \tag{6.168}$$

The system (6.168) can in turn be rewritten in the form

$$\eta_i = \phi_\lambda(\xi_i), \qquad \xi_i = \psi_\mu(\eta_i), \qquad 1 \leq i \leq m, \tag{6.169}$$

where

$$\phi_\lambda(x) = \frac{\lambda + 1 + \log(x) - ax}{c}, \qquad \psi_\mu(y) = \frac{\mu + 1 + \log(y) - by}{c}. \tag{6.170}$$

If we let $P_{\lambda,\mu}$ denote the intersection of the graphs $y = \phi_\lambda(x)$ and $x = \psi_\mu(y)$, we can summarize these findings as follows: the maximum of F in Ω is attained either at the point ω_0 (6.16), or at a point of the form (6.164), where $2 \leq m \leq \theta$, $(\xi, \eta) \in \Omega^{(m)}$ and $(\xi_i, \eta_i) \in P_{\lambda,\mu}$ for $1 \leq i \leq m$. Note that $\phi_\lambda''(x) = -1/cx^2$, $\psi_\mu''(y) = -1/cy^2$, so for $c > 0$ the graphs are convex. We can now prove that for $c > 0$, a maximiser of F can have at most two distinct entries x_i (and therefore the same for y_i). Henceforth we suppress the indices λ, μ from ϕ, ψ.

Proposition 6.5.3. *If $c > 0$ then the m of (6.164) satisfies $m \leq 2$.*

Proof. Suppose first that $b < 0$. Then, ψ is increasing and concave, so ψ^{-1} is increasing and convex. The graph of ψ^{-1} can intersect the graph of the concave function ϕ in at most two points. If $a < 0$ the same argument works with ϕ and ψ interchanged.

This leaves the case when $a > 0$ and $b > 0$. In the region

$$\mathcal{R} = \{(x, y) : 0 < x < 1/a,\ 0 < y < 1/b\}, \tag{6.171}$$

ϕ is increasing and concave whereas the local inverse ψ^{-1} is increasing and convex. Thus, there are at most two crossing points in $\mathcal{R}$. If there are zero or two crossing points in $\mathcal{R}$, then an elementary convexity argument shows that there are no crossing points outside $\mathcal{R}$.

In all the cases considered so far there are at most two crossing points, which implies $m \leq 2$. In the remaining case, when there is exactly one crossing point in $\mathcal{R}$, there can be several crossing points outside $\mathcal{R}$. They can be ordered as a sequence (x_j, y_j) with x_j decreasing and y_j increasing. We are only interested in subsequences of crossing points with x_j and y_j decreasing. The maximum length of such a subsequence is 2, where we may pick the unique crossing point in $\mathcal{R}$ and an arbitrary crossing point outside $\mathcal{R}$. This proves that $m \leq 2$ also in this case. ∎

We can now prove the main thrust of Proposition 6.5.1, that for $c > 0$, the maximiser of F in Ω is of the form (6.163).

Proof of (6.163). We absorb β in a, b, c, effectively setting $\beta = 1$. By Proposition 6.5.3 we can write $\vec{x}$ and $\vec{y}$ as

$$\begin{aligned} x_1 = \dots = x_k = s, \quad & x_{k+1} = \dots = x_\theta = \tfrac{\rho - ks}{\theta - k}, \\ y_1 = \dots = y_k = t, \quad & y_{k+1} = \dots = y_\theta = \tfrac{\rho' - kt}{\theta - k}, \end{aligned} \tag{6.172}$$

where $\rho' = 1 - \rho$ and where the range for s, t is: $\rho/\theta \le s \le \rho/k$ $\rho'/\theta \le t \le \rho'/k$. Recall that $Q(x, y) = \frac{1}{2}(ax^2 + by^2 + 2cxy)$. Then F evaluated at such an $(\vec{x}; \vec{y})$ can be written as

$$\begin{aligned} F(s,t,k) =& k\big(- s\log s - t\log t + Q(s,t)\big) - (\rho - ks)\log\big(\tfrac{\rho-ks}{\theta-k}\big) \\ & - (\rho' - kt)\log\big(\tfrac{\rho'-kt}{\theta-k}\big) + (\theta - k)Q\big(\tfrac{\rho-ks}{\theta-k}, \tfrac{\rho'-kt}{\theta-k}\big). \end{aligned} \tag{6.173}$$

We regard k as a continuous variable satisfying $1 \le k \le \theta$. The plan is to show that $F(s, t, k)$ does not have any stationary points in the interior of the relevant domain for s, t, k, and then that on the boundary it is largest for $k = 1$.

First consider the local maxima. We get

$$\begin{aligned} \tfrac{\partial F}{\partial s} &= k\Big[a\big(\tfrac{\theta s-\rho}{\theta-k}\big) + c\big(\tfrac{\theta t-\rho'}{\theta-k}\big) - \log\big(\tfrac{s(\theta-k)}{\rho-ks}\big)\Big], \\ \tfrac{\partial F}{\partial t} &= k\Big[b\big(\tfrac{\theta t-\rho'}{\theta-k}\big) + c\big(\tfrac{\theta s-\rho}{\theta-k}\big) - \log\big(\tfrac{t(\theta-k)}{\rho'-kt}\big)\Big]. \end{aligned} \tag{6.174}$$

We now introduce the notation:

$$\xi = \tfrac{\theta s-\rho}{\theta-k} \in [0, \rho/k], \qquad \eta = \tfrac{\theta t-\rho'}{\theta-k} \in [0, \rho'/k]. \tag{6.175}$$

Note that $\xi = 0$ if and only if $s = \rho/\theta$, in which case all x-variables are the same, i.e. $(\vec{x}; \vec{y}) = \omega_0$. Similarly if $\eta = 0$. So we need to see if there is a stationary point with $\xi > 0$ and $\eta > 0$. Setting $\frac{\partial F}{\partial s} = \frac{\partial F}{\partial t} = 0$ we get the equations

$$\log\big(\tfrac{s(\theta-k)}{\rho-ks}\big) = a\xi + c\eta, \qquad \log\big(\tfrac{t(\theta-k)}{\rho'-kt}\big) = b\eta + c\xi. \tag{6.176}$$

It is useful to solve these for s, t:

$$s = \xi \frac{e^{a\xi+c\eta}}{e^{a\xi+c\eta} - 1}, \qquad t = \eta \frac{e^{b\eta+c\xi}}{e^{b\eta+c\xi} - 1}. \tag{6.177}$$

Next, computing the k-derivative we get

$$\begin{aligned} \tfrac{\partial F}{\partial k} =& \tfrac{\theta s-\rho}{\theta-k} - s\log\big(\tfrac{s(\theta-k)}{\rho-ks}\big) + \tfrac{\theta t-\rho'}{\theta-k} - t\log\big(\tfrac{t(\theta-k)}{\rho'-kt}\big) \\ & + \tfrac{1}{2}\Big[a\big(\tfrac{\theta s-\rho}{\theta-k}\big)^2 + b\big(\tfrac{\theta t-\rho'}{\theta-k}\big)^2 + 2c\big(\tfrac{\theta s-\rho}{\theta-k}\big)\big(\tfrac{\theta t-\rho'}{\theta-k}\big)\Big], \end{aligned} \tag{6.178}$$

which simplifies to

$$\tfrac{\partial F}{\partial k} = \xi\Big(1 - \frac{e^{a\xi+c\eta}(a\xi + c\eta)}{e^{a\xi+c\eta} - 1}\Big) + \eta\Big(1 - \frac{e^{b\eta+c\xi}(b\eta + c\xi)}{e^{b\eta+c\xi} - 1}\Big) + Q(\xi, \eta). \tag{6.179}$$

We show below that $\frac{\partial F}{\partial k} < 0$ for $\xi, \eta > 0$, except for a certain case which does not have any relevance.

Now recall that the domain in question constists of those (k, s, t) such that $1 \le k \le \theta$,

$\rho/\theta \le s \le \rho/k$, and $\rho'/\theta \le t \le \rho'/k$. The boundary consists of points where at least one of the inequalities is in fact an equality. Start with $s = \rho/\theta$: then all x-variables are equal (to ρ/θ) and then in fact $(\vec{x};\vec{y}) = \omega_0$. Similarly if $t = \rho'/\theta$. Next, if $s = \rho/k$ then $x_{k+1} = \frac{\rho-ks}{\theta-k} = 0$. But we know from Lemma 6.5.2 that F is not maximized at such a point. Similarly, we may exclude the possibility $t = \rho'/k$. Finally, $k = \theta$ also gives $(\vec{x};\vec{y}) = \omega_0$, so the only possibility for $m = 2$ is if $k = 1$.

We now show that $\frac{\partial F}{\partial k} < 0$, for all $\xi, \eta > 0$, unless $-a/c = -c/b = \alpha > 0$, in which case it is equal to zero. We first reparametrise $\frac{\partial F}{\partial k}$ by setting $\eta = \alpha\xi$, for some $\alpha > 0$. This gives:

$$\begin{aligned}\frac{\partial F}{\partial k} = \xi\Big(1 - \frac{e^{(a+c\alpha)\xi}(a+c\alpha)\xi}{e^{(a+c\alpha)\xi}-1}\Big) + \alpha\xi\Big(1 - \frac{e^{(c+b\alpha)\xi}(c+b\alpha)\xi}{e^{(c+b\alpha)\xi}-1}\Big)\\ + \frac{1}{2}(a+c\alpha+\alpha(c+b\alpha))\xi^2.\end{aligned} \tag{6.180}$$

Letting $\delta = a + c\alpha$, $\gamma = c + b\alpha$, and $G_\delta(\xi) = \xi + \delta\xi^2\big(\frac{1}{2} - \frac{e^{\delta\xi}}{e^{\delta\xi}-1}\big)$, we have

$$\frac{\partial F}{\partial k} = G_\delta(\xi) + \alpha G_\gamma(\xi). \tag{6.181}$$

It now suffices to analyse $G_\delta(\xi)$, $\xi > 0$, $\delta \in \mathbb{R}$. We can rewrite this function as

$$G_\delta(\xi) = \frac{2\xi(e^{\delta\xi}-1) - \delta\xi^2(e^{\delta\xi}+1)}{2(e^{\delta\xi}-1)}. \tag{6.182}$$

For $\delta > 0$, the denominator is positive, and rearranging shows the numerator is negative if and only if $\tanh\big(\frac{1}{2}\delta\xi\big) < \frac{1}{2}\delta\xi$, which holds for all $\xi > 0$. Similarly if $\delta < 0$, then the numerator is positive if and only if $\tanh\big(\frac{1}{2}\delta\xi\big) > \frac{1}{2}\delta\xi$, which holds for all $\xi > 0$. Lastly, if $\delta = 0$, then $G_\delta(\xi) = 0$ for all $\xi > 0$.

Hence using (6.181), we see that $\frac{\partial F}{\partial k} < 0$ unless $\eta = \alpha\xi$, and both $\delta = a + c\alpha = 0$, and $\gamma = c + b\alpha = 0$. But this case is not relevant, since substituting these three equations into (6.177) gives $s = \infty$ and $t = \infty$. ∎

To finish the proof of Proposition 6.5.1, it remains to prove that in the case $a = b = 0$, $c > 0$, $\rho = \frac{1}{2}$, and $\beta \neq \beta_c$, the maximiser is unique and satisfies $x_i = y_i$ for all $i = 1, \dots, \theta$. Without loss of generality we can let $c = 1$. Using the fact that the maximiser must be of the form (6.163), and setting $x_1 = x$, $y_1 = y$, we can write

$$\begin{aligned}F(\vec{x};\vec{y}) = F_0(x,y) :=& \beta\Big(xy + \tfrac{(\frac{1}{2}-x)(\frac{1}{2}-y)}{\theta-1}\Big) - x\log x - y\log y\\ &- \big(\tfrac{1}{2}-x\big)\log\tfrac{\frac{1}{2}-x}{\theta-1} - \big(\tfrac{1}{2}-y\big)\log\tfrac{\frac{1}{2}-y}{\theta-1}.\end{aligned} \tag{6.183}$$

We are maximising F_0 in the box $[\frac{1}{2\theta}, \frac{1}{2}]^2$. Calculations yield that when $x > y$, $\frac{\partial F_0}{\partial x} < \frac{\partial F_0}{\partial y}$, and vice-versa, so that the maximum points of F_0 must satisfy $x = y$ or lie on the boundary. Lemma 6.5.2 shows that they cannot lie on the boundary unless $(\vec{x};\vec{y}) = \omega_0$. So, substituting $x = y$, and reparametrising with $z = 2x$, we have

$$F_0\big(\tfrac{z}{2}, \tfrac{z}{2}\big) = \tfrac{\beta}{4}\big(z^2 + \tfrac{(1-z)^2}{\theta-1}\big) - z\log z - (1-z)\log\tfrac{1-z}{\theta-1} + \log 2. \tag{6.184}$$

Now, apart from the constant $\log 2$, this is precisely the function maximised in [13, Theorem 1.1], with β in that paper replaced with $\beta/2$ here, and $\vec{x}$ in that paper of the form $x_1 \geq x_2 = \cdots = x_\theta$. By the working in that paper and the Appendix of [14], the maximiser is unique for all $\beta \neq \beta_c = \frac{4(\theta-1)\log(\theta-1)}{\theta-2}$ from (6.19). This concludes the proof of Proposition 6.5.1. ∎

6.6 The trace-inequality (6.79)

The inequality (6.79) appears e.g. in [69, Prop. 9.H.1.g-h], but we give here an almost self-contained proof based on Birkhoff's theorem, adapted from the discussion at [105]. The problem is to maximize (respectively, minimize) $\mathrm{Tr}[XY]$ subject to the condition that X, Y are nonnegative definite Hermitian matrices with fixed spectra $x_1 \geq x_2 \geq \cdots \geq x_\theta \geq 0$ and $y_1 \geq y_2 \geq \cdots \geq y_\theta \geq 0$. Equivalently, since there are unitary matrices U and V such that $U^*XU = D_x = \mathrm{diag}(x_1, \ldots, x_\theta)$ and $V^*YV = D_x = \mathrm{diag}(x_1, \ldots, x_\theta)$, the goal is to to extremize

$$\mathrm{Tr}[UD_xU^*VD_yV^*] = \mathrm{Tr}[D_xU^*VD_yV^*U] \tag{6.185}$$

over unitaries U, V. Writing $W = U^*V$ we may equivalently extremize over the unitary W,

$$\mathrm{Tr}[D_xWD_yW^*] = \sum_{i,j=1}^{\theta} x_i w_{i,j} y_j w_{j,i}^* = \sum_{i,j=1}^{\theta} x_i y_j |w_{i,j}|^2. \tag{6.186}$$

Define the matrix $P = (p_{i,j})_{i,j=1}^{\theta}$ where $p_{i,j} = |w_{i,j}|^2$. Since W is unitary, P is doubly stochastic (rows and columns sum to 1). We have by the above

$$\max\nolimits_W \ \mathrm{Tr}[D_xWD_yW^*] \geq \max\nolimits_P \sum_{i,j=1}^{\theta} x_i y_j p_{i,j}, \tag{6.187}$$

where the second max is over doubly-stochastic matrices P (and similarly for the min). The function to be maximized on the right-hand-side is linear in P and the set of doubly-stochastic matrices is convex and compact. Thus the maximum (as well as the minimum) is attained at an extreme point of the set of doubly-stochastic matrices. By Birkhoff's theorem [69, Theorem 2.A.2], the extreme points are the permutation matrices Π. Since permutation matrices are real orthogonal (hence unitary) it follows that

$$\max\nolimits_W \ \mathrm{Tr}[D_xWD_yW^*] = \max\nolimits_\Pi \ \mathrm{Tr}[D_x\Pi D_y\Pi^*] \tag{6.188}$$

and similarly for the minimum. Thus, we must only find the permutation π which maximizes or minimizes the function

$$\sum_{j=1}^{\theta} x_j y_{\pi(j)}. \tag{6.189}$$

The maximum is obtained for the identity permutation and the minimum for the reversal of $12\ldots\theta$.

6.7 Equivalence of $Q_{i,j}$ and $P_{i,j}$ in the wb-model

In this second appendix we study two representations of the walled Brauer algebra $\mathbb{B}_{n,m,\theta}$. We will prove that they are isomorphic for all $\theta \geq 2$. This will in particular give the equivalence of our WB-model with the same model, but with each $Q_{i,j}$ replaced with $P_{i,j}$. More generally Lemma 6.7.1 gives the same statement on general graphs. To be precise, if $G = A \cup B$ is any graph (with $A \cap B = \varnothing$), with E_A the set of edges between two vertices in A, E_B similar, and E_{AB} those between a vertex of A and a vertex of B, then for all $a, b, c \in \mathbb{R}$, the following two Hamiltonians are unitarily equivalent:

$$\begin{aligned} H &= -\sum_{\{i,j\}\in E_A} aT_{i,j} - \sum_{\{i,j\}\in E_B} bT_{i,j} - \sum_{\{i,j\}\in E_{AB}} cP_{i,j}, \\ H' &= -\sum_{\{i,j\}\in E_A} aT_{i,j} - \sum_{\{i,j\}\in E_B} bT_{i,j} - \sum_{\{i,j\}\in E_{AB}} cQ_{i,j}. \end{aligned} \tag{6.190}$$

This in particular shows that the models with interactions $P_{i,j}$ and $Q_{i,j}$ are equivalent on any bipartite graph; the equivalence of partition functions was proved by Aizenman and Nachtergaele in [1]. The same statement (and in fact slightly stronger) holds on non-bipartite graphs, but only for r odd. Indeed, (6.190) is very similar to a statement on the model (6.3): for any graph G with edge set E, for any $L_1, L_2 \in \mathbb{R}$, the following two Hamiltonians are unitarily equivalent for r odd:

$$\begin{aligned} H &= -\sum_{\{i,j\}\in E} L_1 T_{i,j} + L_2 P_{i,j}, \\ H' &= -\sum_{\{i,j\}\in E} L_1 T_{i,j} + L_2 Q_{i,j}. \end{aligned} \tag{6.191}$$

This is proved with Lemma B.1 of [89], which is the equivalent of our Lemma 6.7.1 below, but for the full Brauer algebra.

The representations we consider are defined as follows. First, we let $|a\rangle$ denote the standard basis for $\mathbb{C}^\theta$, indexed using $a \in \{-S, -S+1, \dots, S\}$ where $S = (\theta-1)/2$, and recall that $\mathbb{V} = V^{\otimes n}$. Let $T : \mathbb{B}_{n,m,\theta} \to \mathrm{End}(\mathbb{V})$ satisfy

$$T(\overline{i,j}) = Q_{i,j}, \qquad T(i,j) = T_{i,j}, \tag{6.192}$$

where we recall that $T_{i,j}$ is the transposition operator, and $\langle a_i, a_j | Q_{i,j} | b_i, b_j\rangle = \delta_{a_i,a_j}\delta_{b_i,b_j}$ This T is just $\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ (3.12). Similarly, define $\tilde{T} : \mathbb{B}_{n,m,\theta} \to \mathrm{End}(\mathbb{V})$ by

$$\tilde{T}(\overline{i,j}) = P_{i,j}, \qquad \tilde{T}(i,j) = T_{i,j}, \tag{6.193}$$

where we recall that $\langle a_i, a_j | P_{i,j} | b_i, b_j\rangle = (-1)^{a_i - b_i}\delta_{a_i,-a_j}\delta_{b_i,-b_j}$.

Lemma 6.7.1. *For all $\theta \geq 2$, and all n, the representations T and $\tilde{T}$ of $\mathbb{B}_{n,m,\theta}$ are isomorphic via a unitary transformation.*

Proof. The proof follows closely that of Lemma B.1 of [89]. For θ odd, the lemma actually follows from that Lemma B.1 by restricting the two representations there to the walled Brauer algebra. So let θ be even. The elements (i,j) and $(\overline{i,j})$ generate the algebra $\mathbb{B}_{n,m,\theta}$, so we aim to find an invertible linear function $A : \mathbb{V} \to \mathbb{V}$ such that

$$A^{-1} T_{i,j} A = T_{i,j}, \tag{6.194}$$

for all $1 \le i < j \le m$ and $m < i < j \le n$, and

$$A^{-1} Q_{i,j} A = P_{i,j}, \tag{6.195}$$

for all $1 \le i \le m < j \le n$. By the Schur–Weyl duality for the general linear and symmetric groups (6.54), the first condition holds if and only if $A = \alpha^{\otimes m} \otimes \gamma^{\otimes n-m}$ for some $\alpha, \gamma \in GL(\theta)$. Then the second condition also holds if and only if $(\alpha \otimes \gamma)^{-1} Q_{i,j} (\alpha \otimes \gamma) = P_{i,j}$ for all $1 \le i \le m < j \le n$, which holds if and only if:

$$\begin{aligned}(-1)^{a_i - b_i} \delta_{a_i,-a_j} \delta_{b_i,-b_j} &= \sum_{c_i,c_j,d_i,d_j} (\alpha^{-1})_{a_i,c_i} (\gamma^{-1})_{a_j,c_j} \delta_{c_i,c_j} \delta_{d_i,d_j} \alpha_{d_i,b_i} \gamma_{d_j,b_j} \\ &= \sum_{c,d} (\alpha^{-1})_{a_i,c} (\gamma^{-1})_{a_j,c} \alpha_{d,b_i} \gamma_{d,b_j} \\ &= (\alpha^{-1} \gamma^{-\top})_{a_i,a_j} (\alpha^\top \gamma)_{b_i,b_j}.\end{aligned} \tag{6.196}$$

Now recall that we assumed θ to be even, meaning that S and all the indices a_i, a_j, b_i, b_j are odd multiples of $\frac{1}{2}$. Thus $(-1)^{a_i} = -(-1)^{-a_i}$ and (6.196) holds if

$$\alpha^\top \gamma = -(\gamma^\top \alpha)^{-1} = \begin{bmatrix} & & & & (-1)^{-S} \\ & & & (-1)^{1-S} & \\ & & \iddots & & \\ & (-1)^{S-1} & & & \\ (-1)^{S} & & & & \end{bmatrix}. \tag{6.197}$$

The matrix on the right on the right in (6.197) is an involution whose transpose is its negative, so it suffices to check this for $\alpha^\top \gamma$. Further, the matrix consists of the block matrices $(-1)^{\theta/2} \left[\begin{smallmatrix} 0 & i \\ -i & 0 \end{smallmatrix}\right]$ aligned along the antidiagonal, where $i = \sqrt{-1}$.

Such a pair α, γ exists: for example let

$$g_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} i & i \\ -1 & 1 \end{bmatrix}, \qquad g_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ -i & -i \end{bmatrix},$$

take α to be block-antidiagonal with blocks g_1, and take γ to be block-diagonal with blocks $(-1)^{\theta/2} g_2$. Since $g_1^\top g_2 = \left[\begin{smallmatrix} 0 & i \\ -i & 0 \end{smallmatrix}\right]$, $\alpha^\top \gamma$ is as required. Further, since both α and γ are unitary, so is A. ∎

Glossary

algebra 17

centre denoted $Z(A)$ 17

ideal 17

homomorphism 17

representation (or module) 17

End(M) Endomorphisms on M 17

bimodule 17

submodule (or subrepresentation) 17

irreducible (representation) 17

$M_1 \oplus M_2$ The direct sum of representations 18

indecomposable (representation) 18

${}_AA$ The regular representation of the algebra A 18

$M_1 \boxtimes M_2$ The box-tensor product 18

$M \otimes_B N$ The tensor product over B 18

semisimple module 18

semisimple algebra 18

head (of a representation) 19

idempotent 19

orthogonal (idempotent) 19

primitive (idempotent) 19

$M_1 \otimes M_2$ The tensor product 20

trivial representation 20

$\mathbb{C}G$ The group algebra of the given group G 20

dual representation 20

character 20

conjugacy classes 20

$\langle \alpha, \beta \rangle$ Inner product of class functions on a group 20

Lie algebra 21

ψ_ρ^G The irreducible representation of the given group or algebra G corresponding to the partition or tuple ρ 21

χ_ρ^G The character of ψ_ρ^G 21

d_ρ^G The dimension of ψ_ρ^G 21

S_n The symmetric group 21

$\mathcal{N}$ The set $\{1, \ldots, n\}$ 21

(i, j) The transposition in S_n 21

$\lambda, \rho, \mu, \pi, \xi$ Partitions 21

Young diagram 22

λ^{T} The transpose of a Young diagram 22

tableau 22

$\mathcal{T}(U)$ The tableaux with entries in U 22

standard tableau 22

$\mathcal{ST}(U)$ The standard tableaux with entries in U 22

z_τ The Young symmetriser 23

$\mathrm{ct}(\lambda)$ The sum of contents of the Young diagram λ 23

$\mathbb{B}_{n,\theta}$ The Brauer algebra 24

B_n The basis of the Brauer algebra 24

$(\overline{i, j})$ The "Brauer" transposition in B_n 24

$\Delta_\lambda^{\mathbb{B}_{n,\theta}}$ The cell module of the Brauer algebra corresponding to the partition λ 26

$GL(\theta)$ The general linear group 29

$(\cdot, \cdot)$ The (non-degenerate, symmetric, bilinear) inner product 29

$O(\theta)$ The orthogonal group 29

$SO(\theta)$ The special orthogonal group 29

rational (representation) 30

polynomial (representation) 30

$\mathfrak{gl}(\theta)$ The general linear lie algebra 30

$\mathfrak{so}(\theta)$ The special orthogonal lie algebra 30

$[\lambda, \mu]$ Tuple from two partitions 32

res Restriction of a representation 33

Littlewood-Richardson rule 34

$c^{\xi}_{\pi,\mu}$ The Littlewood-Richardson coefficient 34

s_ρ The Schur polynomial 34

semistandard tableau 35

$\mathcal{SS}_\lambda(U)$ The semistandard tableaux with entries in U, shape λ 35

Pieri rule 35

$\tilde{b}^{n,\theta}_{\lambda,\rho}$ The cell module $\mathbb{B}_{n,\theta}$-S_n branching coefficient 36

$b^{n,\theta}_{\lambda,\rho}$ The $\mathbb{B}_{n,\theta}$-S_n branching coefficient 36

$b^{n,m,\theta}_{(\lambda,\mu),(\rho,\xi)}$ The walled Brauer algebra-symmetric group branching coefficient 36

$[V]^G$ Invariants on V with respect to the action of G 39

$\mathfrak{p}^{GL(\theta)}$ The diagonal action of $GL(\theta)$ on $V^{\otimes n}$ 40

$\mathfrak{p}^{S_n}$ The representation of $\mathbb{C}S_n$ on $V^{\otimes n}$ by permuting the tensor factors 40

$\mathfrak{p}^{O(\theta)}$ The diagonal action of $O(\theta)$ on $V^{\otimes n}$ 41

$\mathfrak{p}^{\mathbb{B}_{n,\theta}}$ The representation of $\mathbb{B}_{n,\theta}$ on $V^{\otimes n}$ 41

$T_{i,j}$ The transposition operator 41

$Q_{i,j}$ The projection operator 41

$\mathfrak{q}^{GL(\theta)}$ The action of $GL(\theta)$ on $V^{\otimes n}$, m tensor multiples of the natural action and $n-m$ multiples of its dual 42

$\mathfrak{p}^{\mathbb{B}_{n,m,\theta}}$ The representation of $\mathbb{B}_{n,m,\theta}$ on $V^{\otimes n}$ 42

$\mathcal{Q}(k)$ Set of pairings 49

$\mathcal{Q}'(k)$ Set of pairings 49

W_n Alternative symbol for $V^{\otimes n}$ 49

W^k_n Subset of tensor space 49

$[W_n]^k$ Subset of tensor space 49

classical spin system 70

$\mathbb{Z}^d$ The d-dimensional lattice 70

$\mathbb{S}^2$ The two-sphere 70

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ A graph (and its vertices and edges) 70

σ_i Classical spin 70

β Inverse temperature 70

Z Partition function 71

H Hamiltonian 71

$\langle f(\sigma) \rangle_{H,\mathcal{G},\beta}$ Expectation of classical observable 72

$\Phi(\beta)$ Free energy 73

quantum spin systems 73

quantum Heisenberg model 74

$\boldsymbol{S}_i$ Quantum spin operator 74

antiferromagnet 75

XXZ model 76

bilinear-biquadratic model 76

$S^{(j)}$ Component of the quantum spin operator 79

$\langle \cdot | \cdot \rangle$ Bra-ket notation 79

$\mathbb{V}$ Alternative notation for $V^{\otimes n}$ 80

L_1, L_2 Parameters of the general model of Chapter 5 80

Δ^*_θ Domain for function in a free energy 80

ϕ Function in several free energy functions in Chapter 5 80

Δ_θ Domain for function in a free energy 81

K_1, K_2 Parameters of the spin $S = \frac{1}{2}$ model of Chapter 5 82

J_1, J_2 Parameters of the spin $S = 1$ model of Chapter 5 85

$\Lambda_n(\theta)$ Set of pairs of partitions 92

$P_n(\theta)$ Set of pairs of partitions 93

d_H Hausdorff distance 93

$P_{i,j}$ The spin singlet projection operator 103

$\mathcal{A}$ Subset of the $J_1 - J_2$ plane 104

$\mathcal{A}'$ Subset of the $J_1 - J_2$ plane 106

$\gamma_\lambda^{\mathbb{B}_{n,\theta}}$ The character of $\Delta_\lambda^{\mathbb{B}_{n,\theta}}$ 110

interchange model 117

a, b, c Parameters of the general model of Chapter 6 119

$H_n^{\mathbf{ab}}$ Hamiltonian of the AB-model 119

F Function in several free energy functions in Chapter 6 119

$f(x, y)$ Function in several free energy functions in Chapter 6 119

$\Phi_\beta^{\mathbf{ab}}$ Free energy of the AB-model 120

$H_n^{\mathbf{wb}}$ Hamiltonian of the WB-model 120

$\Phi_\beta^{\mathbf{wb}}$ Free energy of the WB-model 120

Ω Set of vectors in Chapter 6 121

ω_0 Specific vector in Chapter 6 121

$Q(x, y)$ Function in several free energy functions in Chapter 6 121

ω_1 Specific vector in Chapter 6 122

R Function in Chapter 6 122

$H_n^{\mathbf{mb}}$ Hamiltonian of the MB-model 125

p_γ power-sum symmetric polynomial 126

Horn's inequalities 131

Bibliography

[1] Michael Aizenman and Bruno Nachtergaele, *Geometric Aspects of Quantum Spin States*, Comm. Math. Phys. **164** (1994), no. 1, 17–63.

[2] D. Aldous and J. A. Fill, *Reversible Markov Chains and Random Walks on Graphs*, Vol. 21, 2002. Preprint, available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[3] Gil Alon and Gady Kozma, *The probability of long cycles in interchange processes*, Duke Mathematical Journal **162** (2013), no. 9, 1567–1585.

[4] ———, *The mean-field quantum Heisenberg ferromagnet via representation theory* (2018), available at arXiv:1811.10530.

[5] Omer Angel, *Random infinite permutations and the cyclic time random walk*, Discrete random walks (Paris, 2003), 2003, pp. 9–16.

[6] E. J. Beamond, J. Cardy, and J. T. Chalker, *Quantum and classical localisation, the spin quantum Hall effect and generalisations*, Physical Review B **65** (2002), no. 21.

[7] E. J. Beamond, A. L. Owczarek, and J. Cardy, *Quantum and classical localization and the Manhattan lattice*, J. Phys. A **36** (2003), no. 41, 10251–10267.

[8] G. Benkart, M. Chakrabarti, T. Halverson, R. Leduc, C. Y. Lee, and J. Stroomer, *Tensor Product Representations of General Linear Groups and Their Connections with Brauer Algebras*, Journal of Algebra **166** (1994), 529–567.

[9] Nathanaël Berestycki and Gady Kozma, *Cycle structure of the interchange process and representation theory*, Bulletin de la Société Mathématique de France **143** (2015), no. 2, 265–280.

[10] Volker Betz, Johannes Ehlert, and Benjamin Lees, *Phase transition for loop representations of quantum spin systems on trees*, Journal of Mathematical Physics **59** (2018), no. 11, 113302.

[11] J. E. Björnberg, P. Mühlbacher, B. Nachtergaele, and D. Ueltschi, *Dimerization in Quantum Spin Chains with $O(n)$ Symmetry*, Commun. Math. Phys. (2021).

[12] Jakob Björnberg, Hjalmar Rosengren, and Kieran Ryan, *Heisenberg models and Schur–Weyl duality* (2021). Preprint.

[13] J. E. Björnberg, *The free energy in a class of quantum spin systems and interchange processes*, Journal of Mathematical Physics **57** (2016), no. 7.

[14] Jakob E. Björnberg, Jürg Fröhlich, and Daniel Ueltschi, *Quantum Spins and Random Loops on the Complete Graph*, Commun. Math. Phys. **375** (2019), no. 3, 1629–1663.

[15] Jakob E. Björnberg, Michał Kotowski, Benjamin Lees, and Piotr Miłoś, *The interchange process with reversals on the complete graph*, Electronic Journal of Probability **24** (2019).

[16] Jakob E. Björnberg and Daniel Ueltschi, *Decay of transverse correlations in quantum Heisenberg models*, Journal of Mathematical Physics **56** (2015), no. 4.

[17] ———, *Critical Temperature of Heisenberg Models on Regular Trees, via Random Loops*, Journal of Statistical Physics **173** (2018), no. 5, 1369–1385.

[18] A. Borodin and G. Olshanski, *Representations of the Infinite Symmetric Group*, Cambridge University Press, 2017.

[19] Richard Brauer, *On Algebras Which are Connected with the Semisimple Continuous Groups*, Annals of Mathematics **38** (1937), no. 4, 857–872.

[20] William P. Brown, *The semisimplicity of ω_n^f*, Annals of Mathematics **63** (1956), no. 2, 324–335.

[21] M. Correggi, A. Giuliani, and R. Seiringer, *Validity of the Spin-Wave Approximation for the Free Energy of the Heisenberg Ferromagnet*, Commun. Math. Phys. **339** (2015), 279–307.

[22] Anton Cox, Paul Martin, Alison Parker, and Changchang Xi, *Representation theory of towers of recollement: Theory, notes, and examples*, Journal of Algebra **302** (2006), 340–360.

[23] Anton Cox, Maud De Visscher, Stephen Doty, and Paul Martin, *On the blocks of the walled Brauer algebra*, Journal of Algebra **320** (2008), 169–212.

[24] Anton Cox, Maud De Visscher, and Paul Martin, *The blocks of the Brauer algebra in characteristic zero*, Representation Theory **13** (2009), 272–308.

[25] Nicholas Crawford, Stephen Ng, and Shannon Starr, *Emptiness formation probability*, Communications in Mathematical Physics **345** (2016), no. 3, 881–922.

[26] Charles Curtis and Irving Reiner, *Representation Theory of Finite Groups and Associative Algebras*, Interscience Publishers, 1962.

[27] P. Di Francesco and P. Zinn-Justin, *Around the Razumov-Stroganov conjecture: proof of a multi-parameter sum rule*, Electron. J. Combin. **12** (2005).

[28] ———, *Inhomogenous Model of Crossing Loops and Multidegrees of Some Algebraic Varieties*, Comm. Math. Phys. **262** (2006), no. 2, 459–487.

[29] Persi Diaconis, *Group representations in probability and statistics*, Harvard University, 1988.

[30] Persi Diaconis and Mehrdad Shahshahani, *Generating a Random Permutation with Random Transpositions*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **57** (1981), no. 2, 159–179.

[31] William Doran, Phil Hanlon, and David Wales, *On the semisimplicity of Brauer's Centralizer Algebras*, Journal of Algebra **211** (1999), 647–685.

[32] Stephen Doty, Aaron Lauve, and George Seelinger, *Canonical idempotents of multiplicity-free families of algebras*, L'Enseignement Mathématique **64** (2019), no. 1, 23–63.

[33] Freeman J. Dyson, Elliott H. Lieb, and Barry Simon, *Phase Transitions in Quantum Spin Systems with Isotropic and Nonisotropic Interactions*, Journal of Statistical Physics **4** (1978), 335–383.

[34] Pavel Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, Elena Yudovina, and Slava Gerovitch, *Introduction to Representation Theory*, American Mathematical Society, 2011.

[35] Matt Fayers, *LTCC Representation theory.* Lecture notes given at the London Taught Course Centre, available at http://www.ltcc.ac.uk/media/london-taught-course-centre/documents/Representation-Theory-Notes-(updated-2017-18).pdf.

[36] Yu.A. Fridman, O.A. Kosmachev, and Ph.N. Klevets, *Spin nematic and orthogonal nematic states in S=1 non-Heisenberg magnet*, Journal of Magnetism and Magnetic Materials **325** (2013), 125–129.

[37] S. Friedli and Y. Velenik, *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*, Cambridge University Press, 2017.

[38] J. Fröhlich, B. Simon, and T. Spencer, *Infrared Bounds, Phase Transitions and Continuous Symmetry Breaking*, Commun. Math. Phys. **50** (1976), 79–85.

[39] Jürg Fröhlich and Elliott Lieb, *Phase Transitions in Anisotropic Lattice Spin Systems*, Commun. Math. Phys. **60** (1978), 233–267.

[40] W. Fulton, *Young tableaux: with applications to representation theory and geometry*, Cambridge University Press, 1997.

[41] ______, *Eigenvalues, invariant factors, highest weights, and Schubert calculus*, Bulletin of the AMS **137** (2000), no. 3.

[42] W. Fulton and J. Harris, *Representation Theory*, Springer, 1991.

[43] G. M. Benkart, D. J. Britten and F. W. Lemire, *Stability in modules for classical Lie algebras - A constructive approach*, Memoirs of the American Mathematical Society **85** (1990), no. 430.

[44] Roe Goodman and Nolan R. Wallach, *Symmetry, Representations, and Invariants*, Springer, 2009.

[45] J. J. Graham and G. I. Lehrer, *Cellular algebras*, Inventiones mathematicae **123** (1996), 1–34.

[46] Geoffrey Grimmett, *Percolation*, Springer, 1989.

[47] Tom Halverson, *Characters of the centralizer algebras of mixed tensor representations of $GL(r,\mathbb{C})$ and the quantum group $q(gl(r,\mathbb{C}))$*, Pacific Journal of Mathematics **174** (1996), no. 2, 359–410.

[48] Alan Hammond, *Sharp phase transition in the random stirring model on trees*, Probability Theory and Related Fields **161** (2015), no. 3-4, 429–448.

[49] Alan Hammond and Milind Hegde, *Critical point for infinite cycles in a random loop model on trees*, The Annals of Applied Probability **29** (2019), no. 4, 2067–2088.

[50] Phil Hanlon and David Wales, *Eigenvalues connected with Brauer's Centralizer Algebras*, Journal of Algebra **121** (1989), 446–476.

[51] ———, *On the Decomposition of Brauer's Centralizer Algebras*, Journal of Algebra **121** (1989), 409–445.

[52] ———, *Computing the Discriminants of Brauer's Centralizer Algebras*, Mathematics of Computation **54** (1990), no. 190, 771–796.

[53] ———, *A tower construction for the radical in Brauer's Centralizer Algebras*, Journal of Algebra **164** (1994), 773–830.

[54] Kenji Harada and Naoki Kawashima, *Quadrupolar order in isotropic Heisenberg models with biquadratic interaction*, Phys. Rev. B **65** (2002), 052403.

[55] T. E. Harris, *Nearest-neighbor Markov interaction processes on multidimensional lattices*, Advances in Mathematics **9** (1972), 66–89.

[56] W. Heisenberg, *Zur Theorie des Ferromagnetismus*, Z. Physik **49** (1928), 619–636.

[57] G. D. James, *The Representation Theory of the Symmetric Groups*, Springer-Verlag, 1978.

[58] Tom Kennedy, *Long Range Order in the Anisotropic Quantum Ferromagnetic Heisenberg Model*, Commun. Math. Phys. **100** (1985), 447–46.

[59] Tom Kennedy, Elliott H. Lieb, and B. Sriram Shastry, *Existence of Néel order in some spin-1/2 Heisenberg antiferromagnets*, Journal of Statistical Physics **53** (1988), 1019–1030.

[60] Kazuhiko Koike, *On the Decomposition of Tensor Products of the Representations of the Classical Groups: By Means of the Universal Characters*, Advances in Mathematics **74** (1989), 57–86.

[61] Roman Kotecký, Piotr Miłoś, and Daniel Ueltschi, *The random interchange process on the hypercube*, Electronic Communications in Probability **21** (2016).

[62] Gady Kozma and Vladas Sidoravicius, *Lower bound for the escape probability in the Lorentz Mirror Model on the lattice*, 2013.

[63] H. Kraft and C. Procesi, *Classical invariant theory: a primer*, 1996. available at `http://www.math.iitb.ac.in/~shripad/Wilberd/KP-Primer.pdf`.

[64] Benjamin Lees, *Existence of néel order in the s=1 bilinear-biquadratic heisenberg model via random loops*, Communications in Mathematical Physics **347** (2016), no. 1, 83–101.

[65] Linjun Li, *On the Manhattan pinball problem* (2020), available at arXiv:2006.10797.

[66] ______, *Polynomial bound for the localization length of Lorentz mirror model on the 1d cylinder* (2020), available at arXiv:2010.05900.

[67] S. Warzel M. Aizenman H. Duminil-Copin, *Dimerization and Néel Order in Different Quantum Spin Chains Through a Shared Loop Representation*, Ann. Henri Poincaré **21** (2020), 2737–2774.

[68] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd ed., Oxford Mathematical Monographs, Oxford University Press, USA, 1995.

[69] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, Springer, 1979.

[70] M. J. Martins, B. Nienhuis, and R. Rietman, *Intersecting Loop Model as a Solvable Super Spin Chain*, Phys. Rev. Lett. **81** (1998), no. 3, 504–507.

[71] Oliver McBryan and Tom Spencer, *On the Decay of Correlations in $SO(n)$-symmetric Ferromagnets*, Commun, Math. Phys. **53** (1977), 299–302.

[72] Elizabeth S. Meckes, *The Random Matrix Theory of the Classical Compact Groups*, Cambridge University Press, 2019.

[73] N. D. Mermin, *Absence of ordering in certain classical systems*, Journal of Mathematical Physics **8** (1967), no. 5, 1061–1064.

[74] N. D. Mermin and H. Wagner, *Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models*, Phys. Rev. Lett. **17** (1966), 1133–1136.

[75] Pascal Molin, *Intégration numérique par la méthode double-exponentielle*, 2010. Available at `https://hal.archives-ouvertes.fr/hal-00491561`.

[76] Peter Mühlbacher, *Critical Parameters for Loop and Bernoulli Percolation*, ALEA, Lat. Am. J. Probab. Math. Stat. **18** (2021), 289–308.

[77] Bruno Nachtergaele, *A Stochastic Geometric Approach to Quantum Spin Systems*, Probability and Phase Transition, 1994, pp. 237–246.

[78] Maxim Nazarov, *Young's Orthogonal Form for Brauer's Centralizer Algebra*, Journal of Algebra **182** (1996), 664–693.

[79] Bernard Nienhuis and Ronald Rietman, *A Solvable Model for Intersecting Loops* (1993), available at arXiv:hep-th/9301012.

[80] P. P. Nikitin, *The centralizer algebra of the diagonal action of the group $GL_n(\mathbb{C})$ in a mixed tensor space*, Journal of Mathematical Sciences **141** (2007), no. 4.

[81] Soichi Okada, *Pieri Rules for Classical Groups and Equinumeration between Generalized Oscillating Tableaux and Semistandard Tableaux*, Electron. J. Combin. **23** (2016), no. 4, Paper 4.43, 27.

[82] O. Penrose, *Bose-Einstein condensation in an exactly soluble system of interacting particles*, Journal of Statistical Physics **63** (1991), no. 3-4, 761–781.

[83] R. T. Powers, *Heisenberg model and a random walk on the permutation group*, Physics 1 (1976), 125–130.

[84] Arun Ram, *Characters of Brauer's centralizer algebras*, Pacific J. Math. **169** (1995), no. 1, 173–200.

[85] A. V. Razumov and Yu. G. Stroganov, *Combinatorial Nature of the Ground-State Vector of the $O(1)$ Loop Model*, Teoret. Mat. Fiz. **138** (2004), no. 3, 395–400.

[86] David Ruelle, *Statistical Mechanics: Rigorous Results*, W.A. Benjamin, 1969.

[87] Hebing Rui, *A criterion on the semisimple Brauer algebras*, Journal of Combinatorial Theory **111** (2005), 78–88.

[88] Th. W. Ruijgrok and E. G. D. Cohen, *Deterministic lattice gas models*, Phys. Lett. A **133** (1988), no. 7–8, 415–418.

[89] Kieran Ryan, *The free energy of a class of $O_{2S+1}(\mathbb{C})$-invariant spin 1/2 and 1 quantum spin systems on the complete graph* (2020), available at arXiv:2011.07007.

[90] ______, *The Manhattan and Lorentz Mirror Models – A result on the Cylinder with low density of mirrors* (2020), available at arXiv:2010.03609.

[91] O. Schramm, *Compositions of random transpositions*, Israel Journal of Mathematics **147** (2004), 571–593.

[92] Thomas Spencer, *Duality, statistical mechanics, and random matrices*, Current developments in mathematics 2012, 2013, pp. 229–260.

[93] John R. Stembridge, *Rational Tableaux and the tensor algebra of gl_n*, Journal of Combinatorial Theory **Series A 46** (1987), 79–120.

[94] Sheila Sundaram, *Tableaux in the Representation Theory of the Classical Lie Groups*, Institute for Mathematics and Its Applications **19** (1990).

[95] Kengo Tanaka, Akinori Tanaka, and Toshihiro Idogaki, *Long-range order in the ground state of the $S = 1$ isotropic bilinear-biquadratic exchange Hamiltonian* **34** (2001), no. 42, 8767–8780.

[96] Bálint Tóth, *Phase transition in an interacting Bose system. An application of the theory of Ventsel' and Friedlin*, J. Stat. Phys. **61** (1990), 749–764.

[97] ———, *Improved Lower Bound on the Thermodynamic Pressure of the Spin 1/2 Heisenberg Ferromagnet*, Letters in Mathematical Physics **28** (1993).

[98] Tamás A. Tóth, Andreas M. Läuchli, Frédéric Mila, and Karlo Penc, *Competition between two- and three-sublattice ordering for $S = 1$ spins on the square lattice*, Phys. Rev. B **85** (2012), 140403.

[99] V. G. Turaev, *Operator invariants of tangles, and R-matrices*, Izvestiya: Mathematics **35** (1990), 411–444.

[100] Daniel Ueltschi, *Quantum Spin Systems and Phase Transitions*, 2013. available at http://www.ueltschi.org/articles/13-Ua.pdf.

[101] ———, *Random loop representations for quantum spin systems*, Journal of Mathematical Physics **54** (2013), no. 8.

[102] ———, *Ferromagnetism, antiferromagnetism, and the curious nematic phase of $S = 1$ quantum spin systems*, Physical Review E **91** (2015), no. 4.

[103] Hans Wenzl, *On the structure of Brauer's centralizer algebras*, Annals of Mathematics **128** (1988), no. 1, 173–193.

[104] Hermann Weyl, *The Classical Groups: Their Invariants and Representations*, Princeton University Press, 1946.

[105] E. Yang, *Maximizing the trace in an elegant way*, Vol. 61, 1990. https://math.stackexchange.com/q/2558297, version: 2020-08-25.

[106] Robert M. Ziff, X. P. Kong, and E. G. D. Cohen, *Lorentz lattice-gas and kinetic-walk model*, Phys. Rev. A **44** (1991), 2410–2428.