Statistical mechanics on the Square Lattice

by

Jordan C. Moodie

A thesis submitted to the
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Physics and Astronomy
College of Engineering and Physical Sciences
University of Birmingham
January 7, 2021

ABSTRACT

We investigate the statistical mechanical properties of spin models on the square lattice, using the technique of transfer matrices. With the aid of exact diagonalisation we obtain results which indicate that excitations of the transfer matrix can be thought of as physical objects which we term topological excitations, which are domain walls in the case of the Ising model at low temperature. Inspired by these results, we extend the usefulness of Baker-Campbell-Hausdorff formula by finding and proving a new representation. This mathematical result is the major contribution of this thesis to the wider literature. Applying perturbation theory to it allows us to perturbatively find the eigenvalues of a square lattice transfer matrix in a way reminiscent of a high-temperature expansion. We then do so for the Ising model, comparing our results to known formulae, and extend the calculation to Potts models.

STATEMENT OF RELATION TO PUBLISHED WORK

Please note that this thesis contains text lifted from published work. In particular Chapter 2 Section 3 is lifted from [1] and Chapter 3 from [2]. Both of these papers were written by the author of this thesis.

ACKNOWLEDGEMENTS

Thanks, Martin.

CONTENTS

1 Introduction **3**

1.1 The Ising and Potts models 5

1.2 Transfer matrices 9

1.3 The power method 20

2 Motivation **25**

2.1 Numerical technique 26

2.2 Rotated lattice 36

2.3 Bilayer lattice 43

2.4 Summary 51

3 The Baker-Campbell-Hausdorff formula **55**

3.1 Derivation of the main result 59

3.1.1 Expanding M^m in powers of B 61

3.1.2 Rewriting F_N in terms of fundamental sums S_N 63

3.1.3 Calculating S_N 65

3.1.4 Rewriting F_N as a partition sum in terms of f_r 71
3.1.5 A partition formula for f_r 74
3.1.6 Resumming the partition formula 76
3.1.7 Revisiting M^m and implementing the fundamental mathematical approach 78
3.1.8 Final form 80
3.2 Representation as a sum of commutators 81
3.3 Finite examples 85
3.4 Apparent singularities and an alternative representation 90
3.5 Choice of basis 95
3.6 Summary 98

4 Perturbation theory **101**

4.1 General perturbation result 102
4.2 Applying perturbation theory to the Baker-Campbell-Hausdorff series 107
4.3 Understanding the perturbation result 113
4.4 Summary 120

5 Applying perturbation theory to the Potts model **121**

5.1 Low-temperature expansion 122
5.2 Perturbation results for $q = 2$ 138
5.3 Perturbation results for $q \geq 3$ 147

6 Conclusion and outlook **151**

6.1 What is to be done? 153

6.1.1 Proving the non-singularity of the Baker-Campbell-Hausdorff formula 153

6.1.2 Making perturbation theory rigorous 154

6.1.3 Proving cancellation 154

6.1.4 Calculate higher eigenvalues 155

6.1.5 Use the results to say something meaningful 155

6.1.6 Apply the results to more models 156

A Algorithmically removing apparent singularities **157**

References **165**

LIST OF FIGURES

1.1 Transfer matrices for the square lattice 12

2.1 Helical boundary conditions for the square lattice 27

2.2 Numerically obtained free-energy gap for the square lattice Ising model 31

2.3 Depiction of fluctuating domain walls 32

2.4 Comparisons of numerical data with modelling 35

2.5 The rotated lattice 37

2.6 Comparison of free-energy gaps for different rotations of the square lattice 41

2.7 The bilayer lattice 44

2.8 Depiction of topological excitations for the bilayer lattice 45

2.9 Changes of behaviour 48

2.10 Excitations on a finite bilayer lattice 49

3.1 The parameter space of n_1 and n_2 67

3.2 a) Boltzmann suppresion factor; b) the function g_2 96

5.1 The heat capacity of the Ising model 146

5.2 The heat capacity of the 3-state Potts model 149

5.3 Accuracy of the perturbative formula for various Potts models . 150

LIST OF TABLES

3.1 Bounds of the functions g_N, obtained using the procedure outlined in the text and numerically verified 98

CHAPTER 1

INTRODUCTION

Statistical mechanics is a firmly established pillar of modern physics. It arises from physicists' and mathematicians' attempts to connect microscopic behaviour to observed macroscopic trends. This work will focus on magnetic models wherein microscopic spins on a lattice contribute to macroscopic observables like the magnetisation or the heat capacity.

This thesis is structured as follows. First, this chapter contains some basic background material: initially introducing the Ising model then generalising it to Potts models; discussing the analytic tool of transfer matrices; and explaining this work's primary numerical technique of exact diagonalisation via the power method.

Chapter 2 will then show some of these ideas in action. We shall discuss nu-

merical results concerning the Ising model on different variations of a square lattice. We will argue that the corresponding transfer matrix contains information that is useful beyond the partition function and present evidence in favour of this view. In particular, we will show how the differences between eigenvalues of a transfer matrix can be viewed through the physical lens of domain walls, presenting detailed modelling [1]. As far as we are aware this physical interpretation is novel, though of course these excitations are well known in terms of correlation lengths.

We will then move on to chapter 3. This chapter is the major mathematical contribution that this work makes to the statistical mechanics literature and physics more generally [2]. It presents a proof of an alternative representation of the Baker-Campbell-Hausdorff formula, a theorem which is fundamental in the study of Lie groups and Lie algebras [3–5]. It is also highly relevant when considering transfer matrices, as we shall discuss. The chapter also goes into detail on how best to think naturally about this formula, presenting additional representations which may prove much more useful.

In chapter 4 we will apply the time-honoured technique of perturbation theory to the aforementioned Baker-Campbell-Hausdorff formula. This allows us access to the eigenvalues of a given transfer matrix through our new representation. This chapter is a much more recent development than the previous and so should not be thought of as being held to the same level of rigour. Instead,

we shall exactly calculate the formula for the first few orders and discuss trends observed.

Chapter 5 will then be the culmination of this work; we shall apply the previously calculated perturbation theory to the Potts model and see how useful it is. We will take some time to do this first for the Ising model, where we can compare our results to the exact formula, before moving on to the more general and non-integrable Potts model. The result is essentially a new analytic technique for interrogating square lattice spins models, similar to high-temperature expansion [6].

Finally, chapter 6 will conclude this work. We will briefly discuss what has been learned before moving on to the more important topic of what is still to be discovered. We will attempt to provide some hints as to what future directions this project may take.

1.1 The Ising and Potts models

We will first discuss the basic models we will be dealing with and their histories. First, likely the most famous and widely known magnetic model in physics; the Ising model [7].

The Ising model was first written down by Wilhelm Lenz in 1920 [8] and later given as a problem to his student Ernst Ising, who solved it in one-dimension

in his thesis [9]. The model itself is alluringly simple, describing a set of spins which each have two states; up or down. In the absence of a field its energy may be written as

$$H = -\frac{1}{2}\sum_{\langle ij \rangle} J_{ij}\sigma_i\sigma_j \,, \tag{1.1}$$

where $\sigma_i \in \{1, -1\}$ denotes the spin on site i and $\langle ij \rangle$ denotes pairs of adjacent sites. The factor of a half is to account for double counting. The variable J_{ij} describes the interaction between two spins. Throughout this thesis we will be assuming a uniform ferromagnetic interaction, that is that $J_{ij} \equiv J$ and $J > 0$ to make the spins prefer to align.

The binary nature of this Ising model makes it particularly pleasant to work with computationally but also permits great mathematical breakthroughs [10–23]. The square lattice Ising model was solved by Onsager in 1944 [11, 12] and we shall discuss a solution of it in more detail later in this chapter. This solution is incredibly important as, unlike the 1D version, the two-dimensional Ising model contains a phase transition where at a certain non-zero temperature the spins cease to be ferromagnetically aligned and instead enter a disordered phase. It is one of the simplest models to exhibit a phase transition.

While some materials do appear to have Ising like properties, the primary reason the Ising model is studied is the possibility of theoretical advancement. [24] This has been further highlighted by the concept of universality [25]. In

basic terms, universality is the fact that many seemingly different models, and thus materials, exhibit the same behaviour within some critical region surrounding a phase transition. What matters is the models symmetries, rather than any other microscopic features. One would then just need to solve the simplest model in any given universality class to gain insight into all models in that class. Familiar concepts such as critical exponents arise from such discussions. As the 2D Ising model is likely to be the simplest model in its class, studying it is the sensible route to understanding the wider universality class.

Of course, not all phase transitions are in the 2D Ising universality class. The Potts model [26], not to be confused with the planar Potts or clock model, is a natural generalisation of the Ising model whose universality class(es) are typically distinct. Essentially, the Potts model, or more accurately the q-state Potts model, contains spins which can point in any of q directions. One may think of these different directions as being on a $q-1$ dimensional simplex, though typically the energy is written as the sum of Kronecker delta functions

$$H = -J \sum_{\langle ij \rangle} \delta(s_i \, , s_j). \tag{1.2}$$

Here, $s_i \in \mathbb{Z}_q$ denotes the spin on the site i. Again we are assuming a uniform ferromagnetic interaction in the absence of a magnetic field. Of course, under such a formulation the obvious coincidence of the Ising model and the 2-state

Potts model is slightly obscured by a rescaling. Consequently some authors prefer to assign the energy $-J$ if the spins are aligned and $J/(q-1)$ if they are not. This formulation would make the aforementioned simplex analogy less tenuous, but in the interests of mathematical simplicity this document shall prefer the prescription as written above.

This mathematical preference should not be unexpected. The Potts model is not one that is particularly motivated by real-life materials [26]. A more realistic model is the planar Potts model mentioned previously. This model, also known as the clock model, contains spins who can point in any of q evenly spaced directions around a clock-face whose interactions are dependent of the cosine of their relative orientation. It was suggested to Renfrey Potts by his supervisor Cyril Domb [27], a noticeable trend in the histories of these magnetic models. It is, however, a much more complicated model than the regular Potts model that we will be dealing with in this thesis. Generically it lacks duality for example, that is the ability to map low- to high-temperature, which is something we shall discuss in slightly more detail later. It also leads to a more complicated final expression for some of our later results in chapter 5 [28]. In that chapter, however, it will be stressed that all results we present for the simpler model can also be obtained for the more complicated one and how they differ will be hinted at. For now assume that any further mention of the Potts model is referring to the regular q-state Potts model as has been

presented.

1.2 Transfer matrices

Transfer matrices are an analytic tool which may be deployed against problems with translational symmetry [29]. They make solving one-dimensional problems particularly simple and also generalise neatly to two dimensions. They form the mathematical backbone of this thesis so a thorough understanding of them is crucial. Such an understanding is best formed through examples. As such, this section shall present the solution to the 1D Ising model, generalise that to the 1D Potts model, then move on to the more formidable 2D Ising model.

First, the 1D Ising model. As transfer matrices require translational invariance let us assume our spins are on a ring of size N such that

$$H = -J \sum_{j=1}^{N} \sigma_j \sigma_{j+1} \,, \tag{1.3}$$

is the energy of the system, with $\sigma_{N+1} = \sigma_1$. We may rewrite this as the sum of energies of pairs of spins, that is

$$H = \sum_{j=1}^{N} E_{\sigma_j \sigma_{j+1}} \,, \tag{1.4}$$

where

$$E_{\sigma_j\sigma_{j+1}} = -J\sigma_j\sigma_{j+1}\,. \tag{1.5}$$

The partition function is thus

$$Z = \sum_{\sigma_1,\sigma_2,\ldots,\sigma_N} e^{-\beta E_{\sigma_1\sigma_2}} e^{-\beta E_{\sigma_2\sigma_3}} \cdots e^{-\beta E_{\sigma_{N-1}\sigma_N}} e^{-\beta E_{\sigma_N\sigma_1}}\,. \tag{1.6}$$

Next, we introduce the transfer matrix $\hat{T}$ with elements

$$T_{\sigma_i\sigma_j} = e^{-\beta E_{\sigma_i\sigma_j}}\,, \tag{1.7}$$

or explicitly

$$\hat{T} = \begin{bmatrix} T_{\uparrow\uparrow} & T_{\uparrow\downarrow} \\ T_{\downarrow\uparrow} & T_{\downarrow\downarrow} \end{bmatrix} = \begin{bmatrix} e^{\beta J} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J} \end{bmatrix}\,. \tag{1.8}$$

This allows us to rewrite the partition function as

$$Z = \sum_{\sigma_1,\sigma_2,\ldots,\sigma_N} T_{\sigma_1\sigma_2} T_{\sigma_2\sigma_3} \cdots T_{\sigma_{N-1}\sigma_N} T_{\sigma_N\sigma_1}\,, \tag{1.9}$$

where we can immediately recognise matrix multiplication and thus write

$$Z = \sum_{\sigma_1} \hat{T}^N_{\sigma_1\sigma_1} \equiv \mathrm{tr}\left(\hat{T}^N\right)\,. \tag{1.10}$$

Note, if the matrix $\hat{T}$ is diagonalisable then there exists a matrix M such that

$\hat{T} = M\Lambda M^{-1}$ where Λ is a diagonal matrix. The cyclical properties of the trace thus allow us to write

$$Z = \mathrm{tr}\left(\Lambda^N\right) = \lambda_1^N + \lambda_2^N = \lambda_1^N \left(1 + \frac{\lambda_2}{\lambda_1}\right) , \tag{1.11}$$

where λ_1 and λ_2 are the two entries of the matrix Λ, with $\lambda_1 \geq \lambda_2$ chosen without loss of generality, also known as the eigenvalues of the transfer matrix $\hat{T}$. Of course, if the two eigenvalues are distinct then in the thermodynamic limit, that is $N \to \infty$, the partition function depends on λ_1 alone. This is indeed the case here. As such, we have found that the problem of solving the 1D Ising model reduces to finding the largest eigenvalue of its transfer matrix.

Generalising this result to the Potts model is relatively straightforward. One merely needs to replace the energy $E_{\sigma_j \sigma_{j+1}}$ with

$$E_{s_j s_{j+1}} = -J\delta(s_j,\, s_{j+1})\,. \tag{1.12}$$

The transfer matrix is $q \times q$ rather than 2×2, but the approach remains the same. The matrix is very simple with each diagonal element being $\exp(\beta J)$ while all other elements are unity. As such this is not a particularly difficult matrix to diagonalise.

Let us now turn to a non-trivial problem; the square lattice Ising model. We

will offer a whistle-stop tour of the solution based on excellent paper by Shultz, Mattis, and Lieb [13]. First, we begin by writing down the energy,

$$H = -J \sum_{i,\,j} \left(\sigma_{i,\,j} \sigma_{i,\,j+1} + \sigma_{i,\,j} \sigma_{i+1,\,j} \right) . \tag{1.13}$$

We now have two indicies to indicate the two dimensions of the square lattice. Again, we insist of having periodic boundary conditions with both $\sigma_{i,\,j} = \sigma_{i+N,\,j}$ and $\sigma_{i,\,j} = \sigma_{i,\,j+M}$.

Under this toroidal geometry we shall consider the transfer matrix as being split into perpendicular and parallel parts, which do not commute. These are depicted in figure 1.1.

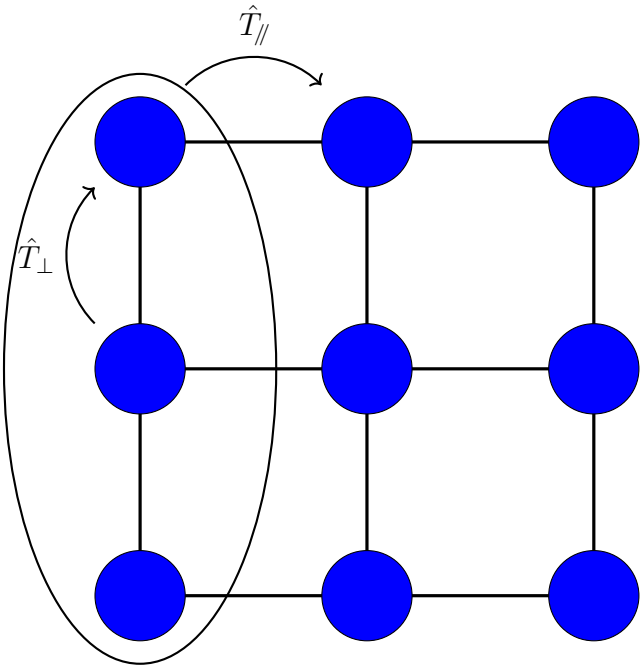


Figure 1.1: Depiction of the perpendicular and parallel parts of the transfer matrix acting on the square lattice, as described in the text.

Considering the perpendicular component as a ring around the torus, we have

$$\hat{T}_{\perp} = \exp\left(\beta J \sum_j \tau_j^z \tau_{j+1}^z\right) . \tag{1.14}$$

Here the operators $\tau_j^\alpha \equiv \mathbb{I} \otimes \cdots \otimes \mathbb{I} \otimes \tau^\alpha \otimes \mathbb{I} \otimes \cdots \otimes \mathbb{I}$ are $2^N \times 2^N$ matrices written in terms the standard Pauli matrices τ^α, with $\alpha = x,\, y,\, z$. This matrix is essentially doing the same thing as those we have already discussed, albeit in a larger space.

The parallel operator, which acts to transfer between each 1D chain, is given by

$$\hat{T}_{/\!/} = \prod_j \left[e^{\beta J} + e^{-\beta J}\tau_j^x\right] . \tag{1.15}$$

The first exponential is for the case where the spin on the first ring is meant to be aligned with the spin on the second, while the second exponential is for when they are not aligned. The spin-flip operator τ_j^x acts to enforce this non-alignment, connecting states where spins are opposite.

Next note that Pauli matrices have the property $\exp(b_0\tau^\alpha) = \cosh b_0 + \sinh b_0 \tau^\alpha$ and so we may recast equation (1.15) as

$$\hat{T}_{/\!/} = e^{b_0 N} \exp\left(\beta \tilde{J} \sum_j \tau_j^x\right) , \tag{1.16}$$

where

$$\sinh 2\beta\tilde{J} \sinh 2\beta J = 1 \,. \tag{1.17}$$

This relation is known as the Kramers-Wannier duality [30–32] and is incredibly important: it relates high-temperature to low-temperature, essentially saying that by relabelling J as $\tilde{J}$ one can map physics below the transition temperature to that above. We will have more to say on this later in this work.

Returning to the calculation, we can combine these two parts into the full transfer matrix,

$$\hat{T} \equiv \hat{T}_{/\!/}^{\frac{1}{2}} \, \hat{T}_{\perp} \, \hat{T}_{/\!/}^{\frac{1}{2}} \,. \tag{1.18}$$

Our aim is to diagonalise this transfer matrix and our approach will be to write it as a single exponential. First we shall apply the rotation $z \to -x$ in order to make the bilinear part in the exponents $\tau_j^x \tau_{j+1}^x$ and the linear part τ_j^z. Upon performing the Jordan-Wigner transformation this will leave us with a quadratic expression in fermionic operators, with which we may proceed.

To that end, let us introduce the raising and lowering operators

$$\tau_j^{\pm} = \frac{1}{2} \left(\tau_j^x \pm i\tau_j^y \right) \,. \tag{1.19}$$

Using the inverse Jordan-Wigner transformation [33],

$$\tau_j^+ = \exp\left(-i\pi \sum_{n=1}^{j-1} f_n^\dagger f_n\right) f_j^\dagger \,, \tag{1.20}$$

$$\tau_j^- = \exp\left(i\pi \sum_{n=1}^{j-1} f_n^\dagger f_n\right) f_j \,, \tag{1.21}$$

we may represent the transfer matrix in terms of a collection of standard fermionic second-quantised operators $f_j^\dagger$. We note that

$$\tau_j^z = f_j^\dagger f_j - f_j f_j^\dagger \,, \tag{1.22}$$

$$\tau_j^x \tau_{j+1}^x = (f_j^\dagger - f_j)(f_{j+1}^\dagger + f_{j+1}) \,. \tag{1.23}$$

However, things are slightly more complicated for the case $j = N$. There we have

$$\tau_N^x \tau_1^x = -(-1)^{\mathcal{N}} (f_N^\dagger - f_N)(f_1^\dagger + f_1) \,, \tag{1.24}$$

where $\mathcal{N} = \sum_j f_j^\dagger f_j$ is the number of fermions in the state. The parity of this operator is a conserved quantity which splits the overall space in two. In terms of the spin operators this is given by

$$\hat{\Sigma} = \prod_j \tau_j^z \qquad \Longrightarrow \qquad \hat{\Sigma}^2 = 1 \,. \tag{1.25}$$

Returning to our transfer matrix, upon performing a Bloch transform we

have

$$\hat{T}_{/\!/} = e^{b_0 N} \exp\left(\beta\tilde{J}\sum_k \left(f_k f_k^\dagger - f_k^\dagger f_k\right)\right) , \tag{1.26}$$

and

$$\hat{T}_\perp = \exp\left(\beta J \sum_k \left[\cos k \, \left(f_k^\dagger f_k + f_{-k}^\dagger f_{-k}\right) - \sin k \left(f_k^\dagger f_{-k}^\dagger - f_k f_{-k}\right)\right]\right) . \tag{1.27}$$

The splitting of the space from the conserved quantity $\hat{\Sigma}$ manifests itself in the definition of the Bloch k's. For an odd number of fermions we are in the periodic subspace and k is an Nth root of unity, as usual. However, for an even number of fermions k is instead an Nth root of negative unity and we are in the anti-periodic subspace.

We can represent this problem in terms of new operators

$$\sigma_k^x = f_k^\dagger f_{-k}^\dagger - f_k f_{-k} , \tag{1.28}$$

$$\sigma_k^y = i\left(f_k^\dagger f_{-k}^\dagger + f_k f_{-k}\right) , \tag{1.29}$$

$$\sigma_k^z = \frac{1}{2}\left(f_k^\dagger f_k - f_k f_k^\dagger + f_{-k}^\dagger f_{-k} - f_{-k} f_{-k}^\dagger\right) \tag{1.30}$$

where the expressions for σ_k^x and σ_k^z come from the exponents of the transfer matrix, noting that $\sum_k \cos k = 0$ in order to give us the extra 1, while σ^y is obtained from $i\sigma_k^x\sigma_k^z$. These operators automatically satisfy the commutation

relations

$$[\sigma_k^\alpha,\, \sigma_k^\beta] = 2i\varepsilon_{\alpha\beta\gamma}\,\sigma_k^\gamma\,, \tag{1.31}$$

and as such provide a representation of spin. They also satisfy

$$(\sigma_k^x)^2 = (\sigma_k^y)^2 = (\sigma_k^z)^2 = 1 - \hat{P}_1\,, \tag{1.32}$$

where

$$\hat{P}_0 = f_{-k} f_k f_k^\dagger f_{-k}^\dagger\,, \tag{1.33}$$

$$\hat{P}_2 = f_k^\dagger f_{-k}^\dagger f_{-k} f_k\,, \tag{1.34}$$

$$\hat{P}_0 + \hat{P}_1 + \hat{P}_2 = 1\,. \tag{1.35}$$

These operators $\hat{P}_n$ project onto the states which have n fermions collectively on the sites k and $-k$. This means that if we have zero or two fermions, the application of $1 - \hat{P}_1$ gives unity and as such we are in a spin-half subspace, while if we have just one fermion it gives zero and we are in a spin-zero subspace.

We may now write our transfer matrix as

$$\hat{T} = e^{b_0 N} \prod_k e^{\frac{\beta\tilde{J}}{2}\sigma_k^z} \prod_k e^{-\beta J\left(\cos k\,\sigma_k^z + \sin k\,\sigma_k^y\right)} \prod_k e^{\frac{\beta\tilde{J}}{2}\sigma_k^z}\,, \tag{1.36}$$

where we have noted that, for different values of k, bilinear fermionic opera-

tors commute and so we may write the sum in the exponent as a product of exponentials. Due to this, we may consider each k component of the transfer matrix separately. Noting that $\sigma_k^z = \sigma_{-k}^z$ and $\sigma_k^x = -\sigma_{-k}^x$ to give us the factors of two in the exponents, we may write

$$\begin{aligned} e^{\beta\tilde{J}\sigma_k^z} e^{-2\beta J\left(\cos k\sigma_k^z + \sin k\sigma_k^y\right)} e^{\beta\tilde{J}\sigma_k^z} &= \left[\cosh\beta\tilde{J} + \sinh\beta\tilde{J}\sigma_k^z\right] \\ &\times \left[\cosh 2\beta J - \sinh 2\beta J\left(\cos k\sigma_k^z + \sin k\sigma_k^y\right)\right] \\ &\times \left[\cosh\beta\tilde{J} + \sinh\beta\tilde{J}\sigma_k^z\right] \qquad (1.37) \\ &\equiv e^{-2\beta\mu_k\left[\cos p_k\sigma_k^z + \sin p_k\sigma_k^y\right]}\,. \end{aligned}$$

Using our relation $\sinh 2\beta\tilde{J}\sinh 2\beta J = 1$, we see that

$$\cosh 2\beta\mu_k = \frac{\cosh^2 2\beta J}{\sinh 2\beta J} - \cos k\,, \qquad (1.38)$$

$$\sinh 2\beta\mu_k \cos 2p_k = \cosh 2\beta J\left(\cos k - \frac{1}{\sinh 2\beta J}\right)\,, \qquad (1.39)$$

$$\sinh 2\beta\mu_k \sin 2p_k = \sinh 2\beta J \sin k\,. \qquad (1.40)$$

While this was done for $k \notin \{0, \pi\}$, those two special values give the same results as above.

From this we can find that

$$\hat{T} = e^{b_0 N} \prod_k e^{-\mu_k \left[\cos 2p_k \sigma_k^z + \sin 2p_k \sigma_k^x\right]} \tag{1.41}$$

At this point we have essentially finished. We have achieved our aim in writing the transfer matrix as a single exponential and the subsequent extraction of its eigenvalues is not too onerous a task. We will end the calculation at this point as its main message for the purpose of this work was never what the actual eigenvalues were.

Writing it as a single exponential is not merely a mathematical convience but also of physical significance. If the largest eigenvalue of the transfer matrix is the partition function, then it is also related to the free-energy via

$$z = e^{-\beta F} \,. \tag{1.42}$$

As such, we could write

$$\hat{T} = e^{-\beta \hat{F}} \,, \tag{1.43}$$

defining a free-energy operator $\hat{F}$. We have effectively just found this operator in the case of the square lattice Ising model. One ought to consider this operator as a $(d-1)$-dimensional quantum mechanical Hamiltonian, if the corresponding transfer matrix was for a d-dimensional statistical mechanics

model. This idea is well known in the literature [25]. Standard ideas concerning the spectra of Hamiltonians apply to this model, with energy being replaced by free-energy [1]. We shall have more to say on such operators in the next chapter, chapter 2.

These messages should also be taken into account when reading the mathematical chapters 3 and 4 as well as the culminating chapter, chapter 5, where we shall be trying to find the partition function of the Potts model. Of course, we shall not be as successful as in this Ising calculation because the model is non-integrable, but lessons, particularly from the early parts of this calculation, may be learned.

1.3 The power method

The previous section introduced the concept of transfer matrices and argued they are useful tools in solving certain translationally invariant problems. One may question if they can be useful for other, non-solvable, problems as well. The answer is that they can [1,34–38] and this section shall provide the framework for how transfer matrices may be used as a numerical tool as well as analytic. Note we will be assuming the diagonalisability of the transfer matrices we work with throughout this thesis.

The concept of exact diagonalisation is well known in the quantum mechanical

literature, so with the formulation of a transfer matrix as the exponential of a free-energy operator we ought to be able to apply the same techniques here. Exact diagonalisation just means expressing a finite system as a matrix and diagonalising that matrix with a computer. As such, there are a range of algorithms that one can choose from to achieve this task, each with their own list of benefits and drawbacks.

All numerical calculations in this work have been done using a very basic algorithm called the power method. The power method, when given a matrix, will return its largest eigenvalue, which is ideal for our purposes. Its benefits are simple: you do not need to store the entire matrix, but rather just encode how it acts on a given state; it is robust to numerical error; and it is simple to code. Its main drawback is that it is extremely slow when employed on a dense matrix [39–41]. Some choose to use a sparse matrix approximation to alleviate this issue, essentially crossing out matrix elements deemed too small. However, that is not an issue for us as we will discuss in chapter 2 and we shall not need to make any such approximations.

The idea is to repeatedly apply the matrix to a vector, with the hope that it eventually converges to the eigenvector of the dominant eigenvalue. This is ensured through the same logic we previously used to claim that we only care about the largest eigenvalue of the transfer matrix. Let us begin with a

random vector

$$u = \sum_{i=1}^{N} a_i u_i \tag{1.44}$$

where $a_i \neq 0$ are random numbers while each u_i are the eigenvectors of our $N \times N$ matrix M. In practice one does not know the eigenvectors beforehand so instead one chooses the elements of the vector u at random, in any convenient basis, to ensure that the resulting weights a_i are almost never zero. In particular, we desire there to be a non-zero overlap between this vector and the eigenvector associated with the largest eigenvalue. A random choice is just a probabilistically guaranteed way of achieving this.

Repeatedly applying this matrix to the vector gives

$$M^N u = \sum_{i=1}^{N} a_i \lambda_i^N u_i \,, \tag{1.45}$$

where we assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$. As before, we may write

$$M^N u = \lambda_1^N \left[a_1 u_1 + \sum_{i=2}^{N} a_i \left(\frac{\lambda_i}{\lambda_1} \right)^N u_i \right] . \tag{1.46}$$

If $\lambda_1 > \lambda_2$ then each of these fractions will tend to zero and we may read off the eigenvalue.

If we wish to find higher eigenvalues then this is just as simple. We merely replace the matrix M with the same matrix minus the outer product of the

largest eigenvector with itself multiplied by the eigenvector. That is,

$$M \to M - \lambda_1 u_1 u_1^{\mathrm{T}} . \tag{1.47}$$

This essentially sets the previous largest eigenvalue to zero and makes the previous second-largest eigenvalue the new largest, allowing the power method to be employed again. This process can be repeated as many times as you desire, computational memory permitting.

CHAPTER 2

MOTIVATION

Note that this chapter contains blocks of text taken from reference [1].

This chapter is intended to provide initial motivation and justification for the major driving force of this work, namely that the spectra of transfer matrices correspond to what we term topological excitations [1]. One can also think of these excitations as quasi-particles, though we will not pursue that interpretation here. The necessarily empirical flavour of this section is something to be embraced; it is hoped that if someone finds themselves entangled in the mathematics of the following chapters they can refer back to here to convince themselves that there is indeed light at the end of the tunnel.

Numerical techniques may perhaps provide such a light. Exact diagonalisation, as the name suggests, provides the exact answer to a finite problem. As

such we can use it to ask questions of analytically tricky models which are physically interesting. This chapter will discuss two such sets of models: simple square lattice Ising models, where the lattice has been rotated with respect to the direction of transfer; and bilayer Ising models. The former is interesting because of course the physics ought to be no different from the exactly solvable case, but the transfer matrix changes. This will give insights as to the nature of its excitations and what they describe, even if the partition function must remain unaltered. The latter will prove a curious case, where changes of behaviour in the excitations will occur. In this case, the text is taken from reference [1]. First, however, we shall turn our attention to the numerical technique itself.

2.1 Numerical technique

For all models on which we shall perform calculations, the fundamental numerical technique will remain the same. Firstly, we begin with the one-dimensional spin Hamiltonian

$$H = -\sum_{n=1}^{N} a_n \sum_j \mathbf{S}_j \cdot \mathbf{S}_{j+n} \,, \tag{2.1}$$

where for the moment we shall choose $\mathbf{S}_j \equiv \sigma_j \hat{\mathbf{z}}$, with $\sigma_j \in \{-1, 1\}$, to be Ising spins on a chain whose sites are labelled by the index j. By appropriately choosing the coupling constants a_n and taking the limit $N \to \infty$, this may be

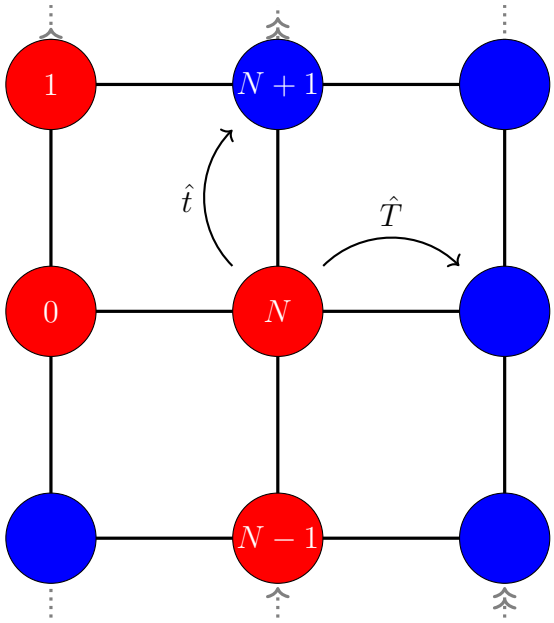


Figure 2.1: Depiction of helical boundary conditions as described in the text. In the thermodynamic limit this leads to the square lattice, as described by equation (2.2). The sub-matrix $\hat{t}$ acts to ratchet the chain of spins highlighted in red along one. That is, spin zero transfers to spins one, spin one to spin two, and so on until spin N tranfers onto spin $N+1$. The transfer matrix which is equivalent to the cylindrical transfer matrix, which transfers one ring to another, is thus N applications of the sub transfer matrix $\hat{t}$. This is described in equation (2.3).

used to probe two-dimensional geometries. For example, setting $a_1 = a_N \equiv J$ with all other matrix elements vanishing leads to the square lattice Ising model

$$H = -\frac{J}{2}\sum_{\langle jj'\rangle}\sigma_j\sigma_{j'}\,. \tag{2.2}$$

This geometry is depicted in figure 2.1 and will be the object of discussion for this section. The notation $\langle jj'\rangle$ implies the sum is over all neighbouring j and j', and hence the factor of one half resolves double counting. Further examples, also accompanied by pictorial aides, will appear later in this chapter.

By employing helical boundary conditions we have introduced an infinitesimal spiral to our system. This has no effect on bulk quantities in the thermodynamic limit, meaning that in practice we may examine the two-dimensional system by extrapolating from relatively modest system sizes. The true benefit of these boundary conditions lies in the sparse Hamiltonian matrix; there are only as many elements in a given row or column as spin degrees of freedom, two for the Ising model. This contrasts with the more widely-used cylindrical geometry, for which the resulting matrix is dense. While the matrix itself is larger in the helical case, as there are fewer symmetries to extract, there is a large computational advantage to working with sparse rather than dense matrices. This advantage allows us to perform calculations on much larger systems than would otherwise be accessible. A representative maximum system size for the Ising model using a cylindrical geometry may be $N \sim 15$ [34, 35, 38] In more numerically demanding works we have managed to calculate on systems up to $N = 29$ [1]. The benefits of doing so in this chapter though are minimal.

The global Ising symmetry, however, still remains. As an example, consider an Ising model in its ferromagnetic ground state, where all spins are aligned. The labelling of the orientation of each spin, either up or down, is irrelevant. This irrelevance halves the state-space, and therefore the size of any matrix, and removes the double ground-state. In practice this is done by introducing a *floating basis* and rewriting the Hamiltonian in terms of variables $\tau_j = \sigma_j \sigma_{j+1}$.

These new variables then describe the relative orientation of each spin, rather than the absolute. Of course, one fewer variable is now required which accounts for the reduction of state-space. The symmetric- and antisymmetric-subspaces then are defined as the collections of states with eigenvalue 1 or -1 respectively with respect to the parity operator $\prod_j \sigma_j^x$, where σ_j^x is the spin-flip Pauli matrix acting on site j. This was seen previously as part of the exact solution, in equation (1.25).

From the Hamiltonian we may construct a transfer matrix

$$\hat{T} = \hat{t}^N \equiv e^{-\beta \hat{F}}\,, \tag{2.3}$$

where explicitly

$$\langle \sigma_0,\, \sigma_1,\, \ldots,\, \sigma_N |\, \hat{t}\, | \sigma_0',\, \sigma_1',\, \ldots,\, \sigma_N' \rangle = e^{-\beta J(\sigma_0\sigma_1 + \sigma_0\sigma_N)} \prod_{j=0}^{N-1} \delta_{\sigma_{j+1},\,\sigma_j'}\,. \tag{2.4}$$

As shown in figure 2.1, the submatrix $\hat{t}$ acts to ratchet along the spiral to the following spin site while the full transfer matrix $\hat{T}$ transfers one to the right of the starting site. The free-energy operator $\hat{F}$ should be interpreted as a quantum mechanical Hamiltonian which, in the thermodynamic limit, is one-dimensional. Standard ideas concerning the spectra of Hamiltonians, in particular energy gaps (think of the metal insulator dichotomy, for example), shall apply to this operator with energy replaced by free-energy. Note, there

is no reason to believe this quantum mechanical operator to be local even if the underlying statistical mechanics model is.

As discussed earlier, the free-energy operator is the prime focus of this work and so we shall spend some time in this section viewing eigenvalues of the submatrix $\hat{t}$ through the lens of the free-energy operator. We may define a free-energy gap in terms of these eigenvalues as

$$\Delta F_m \equiv \frac{N}{\beta} \log \frac{t_0}{t_m} \,, \tag{2.5}$$

where t_0 is the largest eigenvalue and thus becomes the partition function in the thermodynamic limit, while m labels the eigenvalues in the symmetric subspace in order.

In practice we tend to calculate the largest two eigenvalues of this submatrix using the power method for a variety of system sizes. From these calculations we may obtain $\Delta F_1(N, T)$, where $T \equiv 1/\beta$ is the temperature. We then employ polynomial extrapolation for these finite systems at a fixed temperature to effectively obtain $\Delta F_1(\infty, T) \equiv \Delta F(T)$. This is directly analogous to the finite-size scaling of exact diagonalisation results common in the quantum mechanics literature. However, it performs exceptionally well in this thermodynamic context, as discussed in detail in [1] where astonishingly accurate numerical results were obtained. As previously mentioned, the finite data is

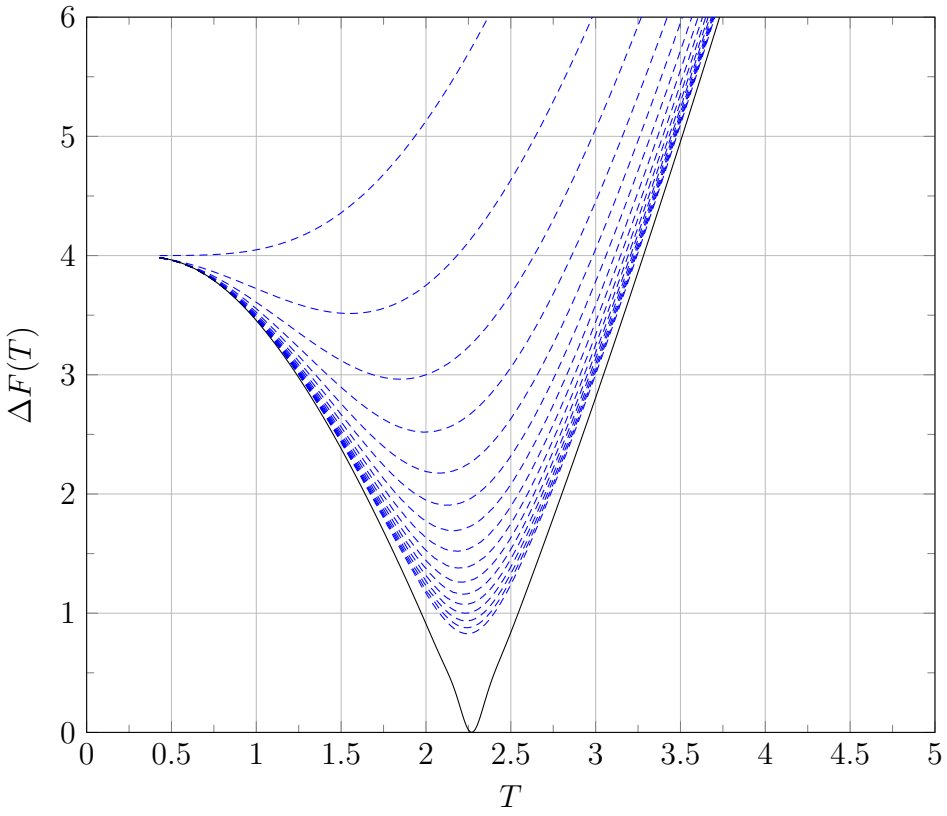


Figure 2.2: The free-energy gap between the two largest eigenvalues in the symmetric subspace for the square lattice Ising model. Dashed lines indicate data for finite systems with $N \in \{2, 3, \ldots, 17\}$. The black solid line is a polynomial extrapolation to the infinite system.

exact to machine precision and so the lone source of error must be the polynomial extrapolation.

Figure 2.2 shows this approach in action for the square lattice Ising model. The dashed lines, the exact finite data, are extrapolated at each temperature point to provide the solid line, which is purported to provide the answer for the infinite system. A quick note ought to be made about convergence. At low temperature several higher lying eigenvalues are almost degenerate with the

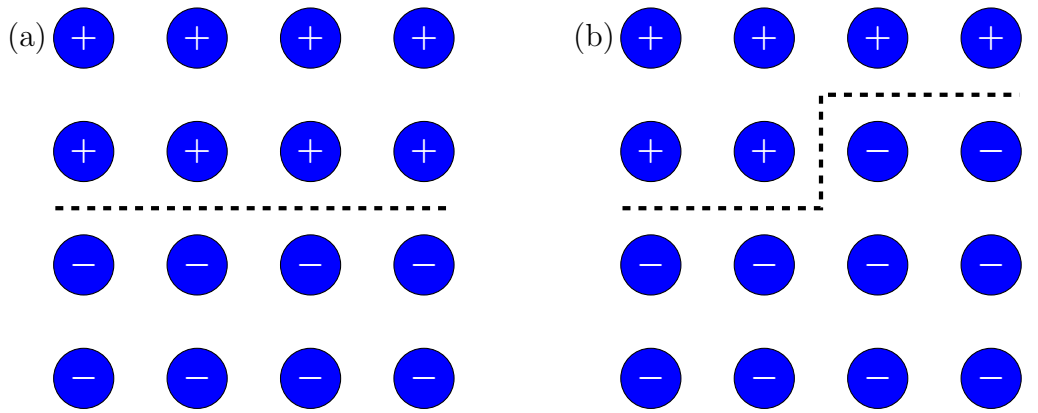


Figure 2.3: Depiction of a physical interpretation of the free-energy gap. (a) a domain wall with no fluctuation. (b) a domain wall with the energetically cheapest fluctuation.

first excitation. As such, convergence for low temperature requires many more iterations than at more moderate temperatures, making it a practical impossibility to obtain results in this region. Fortunately, however, it is physically fairly obvious what one would expect these calculations to show; successive lines of $\Delta F(T) = 4J$, in figure 2.2 for example.

Let us turn now to the interpretation of this figure. We shall talk solely of the extrapolated result, which is an attempt to approximate the infinite system, rather than the finite data. We interpret the function $\Delta F(T)$ as describing how the free-energetic cost of a topological excitation changes with temperature, while the other $\Delta F_m(T)$ describe topological excitations of higher cost. For the Ising model at low temperature this ought to be thought of as a domain wall of infinite extent, propagating in direction of transfer. This is depicted in figure 2.3(a). At zero temperature the free-energetic cost of such an excitation

and its energetic cost, $2J$, coincide. For a periodic system there must exist an even number of such domain walls, hence the observation of $4J$ in figure 2.2 is consistent with this interpretation. As the temperature is increased slightly, entropic fluctuations of the kind displayed in figure 2.3(b) occur, reducing the free-energetic cost of the excitation. At some point then the cost will become zero and a phase transition will occur. Indeed, this is perhaps the best way of defining a phase transition in our formalism: $\Delta F(T_c) \equiv 0$. This is easily understood from an Ising model perspective. The transition in that case is defined by the change from the low-temperature ordered phase, where there is a single divergent cluster of one spin orientation, to a high-temperature disordered phase. As the transition temperature is approached the energetic cost of the domain wall is increasingly compensated for by entropic gain, until both exactly balance and macroscopic domain walls are permitted. As we are effectively calculating a correlation length, these observations make sense.

Let us attempt to model this, however simply, to provide some credulity to these assertions. A natural start may be a partition sum, say

$$\Delta F = -2T \log \left(e^{-2\beta J} + 2e^{-4\beta J} \right) . \tag{2.6}$$

The first exponential describes the situation in figure 2.3(a), while the second describes 2.3(b). The latter is multiplied by two to indicate the fluctuation

can occur either upwards into the region of + spins or downwards into the region of − spins. The entire expression is also multiplied by two to take into account that there must be a pair of domain walls in a periodic system, though we assume these are sufficiently separated as to be independent. A more thorough analysis would of course suggest the function

$$\Delta F = -2T \log \left(e^{-2\beta J} + 2 \sum_{n=2}^{m} e^{-2n\beta J} \right) . \tag{2.7}$$

Both the basic model and this ever-so-slightly more sophisticated one are plotted in figure 2.4, with the latter having the sum performed up to $m =$ 3, 4, 5, ∞, and compared to the extrapolation of figure 2.2. We see that all functions perform well at low temperature, though the more basic functions fade away as temperature is increased. This is to be expected. As the temperature increases longer range fluctuations become increasingly relevant until, at the transition temperature, fluctuations of all ranges are equally important and macroscopic domain walls form. A foible of the Ising model, namely its integrability, requires that the function (2.7) is indeed the exact answer when the infinite sum is calculated. This plot then should be considered both a successful testing of our interpretation of these eigenvalues as well as a confirmation of the extrapolations accuracy.

A physical interpretation of the high-temperature results is also possible. As

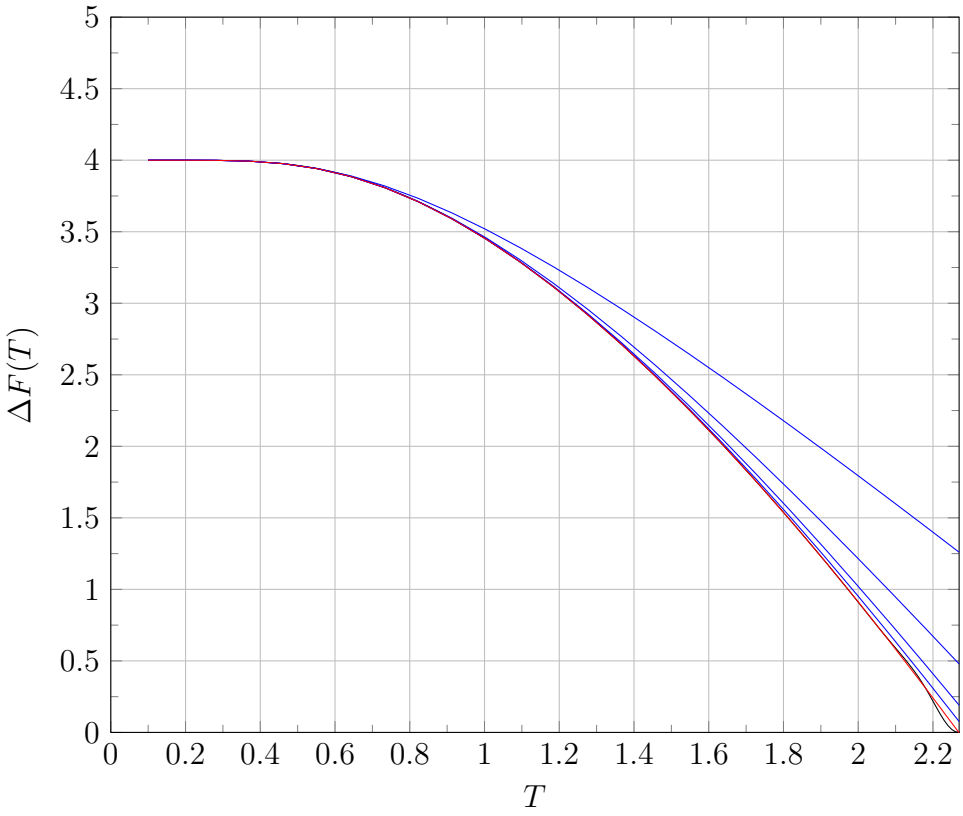


Figure 2.4: Comparison of the extrapolated result for the free-energy gap calculated via exact diagonalisation with various modelling results described in the text. The blue lines describe the formula (2.7) for the cases $m = 2$, 3, 4, 5. The black line indicates the same formula for $m = \infty$, which is the exact result.

the Ising model exhibits Kramers-Wannier duality, displayed in equation (1.17), one can exactly map calculations from low to high temperature and any associated interpretations must also be so mapped. Mathematically, instead of having a single spin-flip operator in the free-energy operator we would instead have a neighbouring pair of such operators.

This then concludes our remarks on the general numerical approach and interpretation. To reiterate, we exactly diagonalise a set of one-dimensional spin

Hamiltonians, extrapolate the results for each temperature point, then interpret that extrapolation as being the free-energy cost of a topological excitation as a function of temperature. It will turn out that for all models we consider in this work the dichotomy in interpretation of domain walls at low-temperature and spin-waves at high-temperature will remain true, though this is may not be truly generic.

2.2 Rotated lattice

As promised, our next example begins with equation (2.1) where again we use Ising spins. The difference to the previous section will lie in the choice of coupling constants a_n. We shall choose many different sets of these coupling constants, each of which will correspond to what we shall term a *rotated lattice*. These geometries, as the name suggests, are square lattices which have been rotated with respect to the direction of transfer by some angle θ. One may think that this ought to have no effect as nothing fundamentally has changed and indeed for thermodynamic quantities this view is entirely correct. The magnetisation, phase transition, and any other number of quantities one wishes to list will remain untouched. However, the rotation does change the direction in which the transfer matrix performs its transfer and thus does affect its excitations and the free-energy gap. If our low-temperature domain wall

interpretation withstands this change then our level of confidence in it may increase.

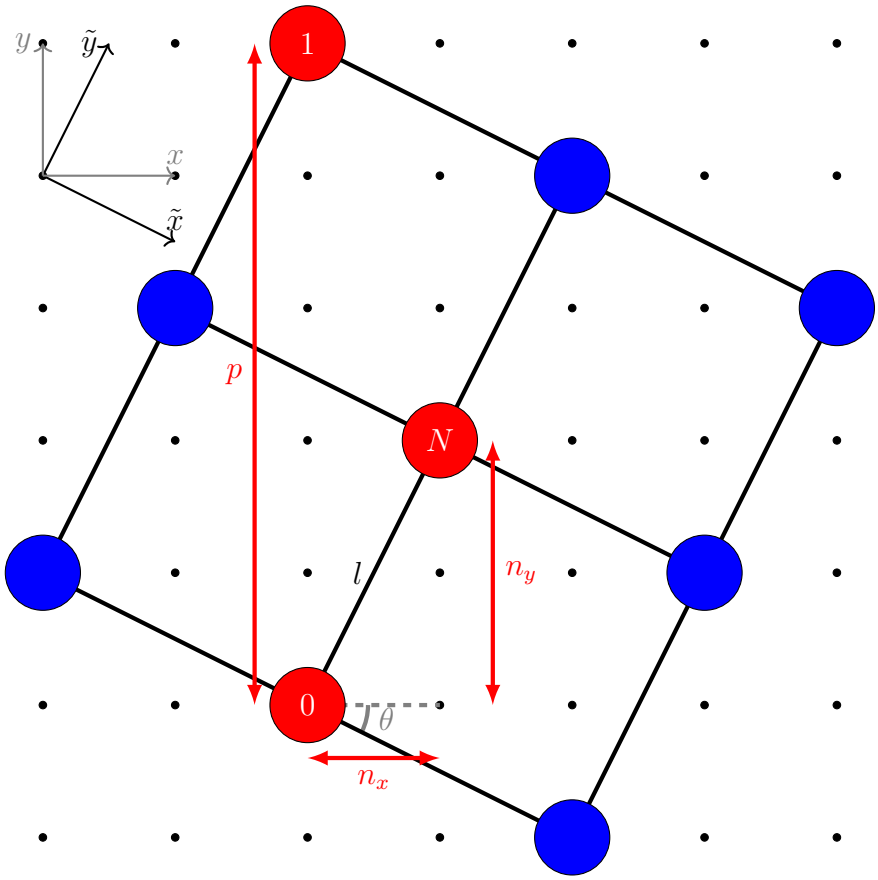


Figure 2.5: Depiction of a rotation. We call the black dots the underlying helical lattice while the blue circles represent the actual position of the spins. The submatrix $\hat{t}$ acts to ratchet the chain of spins highlighted in red along one. That is, spin zero transfers to spin one, spin one to spin two, and so on until spin N transfers to spin $N+1$. The mismatch parameter α is then given by the fraction n_y/p.

Figure 2.5 depicts a rotation and labels some useful variables. To define a certain rotation one may use the variable θ, the angle of this rotation. However, it is more convenient to recognise that this θ must be the arctangent of a rational number, if the rotation is to be commensurate with the underlying

square lattice. Hence the variables n_x and n_y are introduced through the equation

$$\theta = \tan^{-1}\left(\frac{n_x}{n_y}\right) . \tag{2.8}$$

Of course, many pairs of n_x and n_y will give the same θ but it is numerically sensible to use the smallest pair, which results in the smallest transfer matrices. The variable p, which measures the distance between states in a given column and whose utility shall be shortly introduced, can then be calculated via

$$p^2 = (n_x l)^2 + (n_y l)^2 = l^4 \quad \Longrightarrow \quad p = n_x^2 + n_y^2 \,, \tag{2.9}$$

where l is the distance between neighbouring spins.

Let us give an example for clarity. The choice of coupling constants

$$a_N = a_{2N-1} = J \,, \tag{2.10}$$

with all others zero, gives a rotation of angle $\theta = \tan^{-1}(1/2) \approx 27°$. We should note that the size of the transfer matrix scales with the largest index of these bonds, so that in this case it will be of size $2^{2N-1} \times 2^{2N-1}$. There will thus be roughly half as many systems for us to study and use in any extrapolation for this rotation compared to the previous section. Other rotations will be even more severely restricted.

Adapting equations (2.3) and (2.4), we construct a transfer matrix

$$\hat{T} = \hat{t}^{N-\alpha} \equiv e^{-\beta \hat{F}/l}\,, \tag{2.11}$$

where α is given by the ratio n_y/p. As the submatrix $\hat{t}$ is designed to transfer between spins, say between those labelled 0 and 1 in figure 2.5, while the transfer matrix $\hat{t}$ ought to transfer from one underlying lattice to the next along the x direction, the mismatch parameter α is a required alteration. Similarly, the factor of l dividing the free-energy operator is key to calculating quantities on a per-unit-length basis, allowing sensible comparisons between different rotations. The free-energy gap, adapting the definition (2.5), is given by

$$\Delta F_m = \frac{1}{\beta} l(N-\alpha) \log \frac{t_0}{t_m}\,. \tag{2.12}$$

Let us discuss now what our topological excitation picture would predict, at least for low temperature. Duality, of course, remains for this model so this is not overly restrictive. The topological excitations would still be expected to be the domain walls discussed in the last section, but now the direction in which they travel is slightly less obvious. Overall they must move in the direction of transfer, denoted as the x direction in figure 2.5. However, the question remains whether they prefer the $\tilde{x}$ or the $\tilde{y}$ directions. Depending on the choice of angle θ one of these will be *easy* direction, in which fewer bonds

are broken per unit length in the x direction, while the other will be a *hard* direction. The energy cost per unit length of a domain wall propagating in the $\tilde{x}$ and $\tilde{y}$ directions is given by

$$E_{\tilde{x}} = \frac{2J}{\cos\theta}\,, \qquad \text{and} \qquad E_{\tilde{y}} = \frac{2J}{\sin\theta}\,, \tag{2.13}$$

respectively. Clearly both are greater than the unrotated cost of $2J$ and so one would expect the free-energy gap to be larger for a rotated lattice than for the unrotated case at zero temperature and to increase as the lattice is further rotated. Next, for sufficiently low temperature one would expect the domain wall to move in only the low energy direction, giving a relatively constant free-energetic cost similar to that in the unrotated case as seen in figure 2.2. However, as temperature is increased we would expect fluctuations, arising form the domain wall occasionally travelling in the high energy direction, to be encouraged in order to gain entropically. Hence the free-energy cost should reduce. Crucially, however, it should reduce independently from the previously discussed reduction due to fluctuations of the groundstate. As such, one would expect it to reduce sooner in the rotated case, with the temperature at which it does so being determined by the rotation. Everything just described is exhibited in figure 2.6(c), strengthening the case of our interpretation.

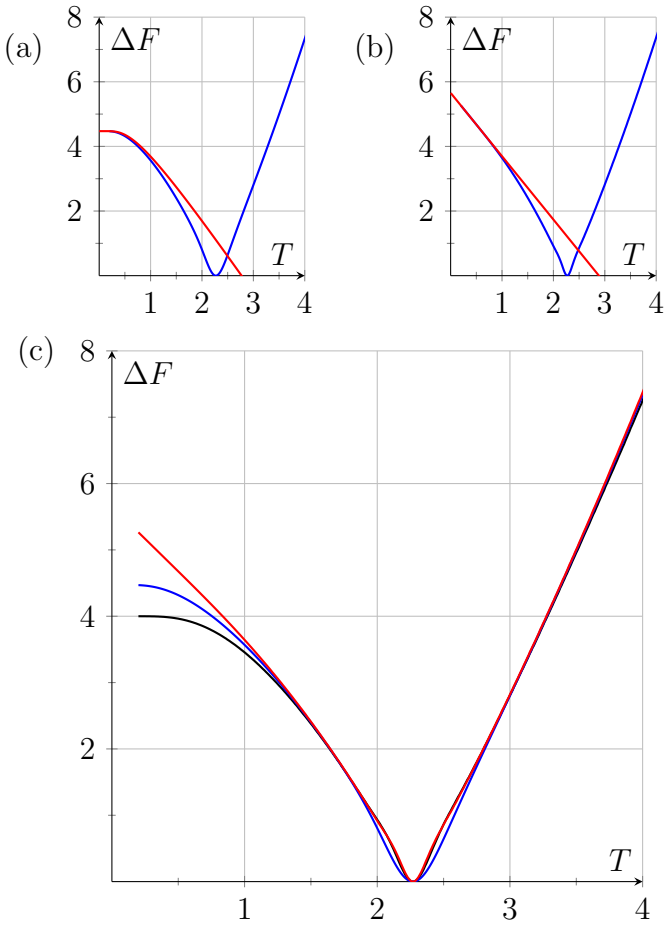


Figure 2.6: The accuracy of the domain wall description for low temperature for (a) the rotation $\tan\theta = 1/2$ (b) the rotation $\theta = \pi/4$. (c) Free-energy gap between the largest two eigenvalues in the symmetric subspace of the transfer matrix. Each line represents a different rotation of a square lattice. Note how larger rotations lead to a more immediate entropic effect against the constant energetic cost at zero to low temperatures.

We ought also to discuss modelling, similar to that of the previous section. The expected free-energy cost, accounting for the two choices of direction but no other fluctuations, can be written as

$$\Delta F = -2\lambda T \log\left[e^{-\frac{1}{\lambda T}E_{\tilde{x}}} + e^{-\frac{1}{\lambda T}E_{\tilde{y}}}\right], \tag{2.14}$$

where

$$\lambda \equiv \frac{1}{2}\left(\frac{1}{\cos\theta} + \frac{1}{\sin\theta}\right) , \tag{2.15}$$

scales the results on a per-unit-length basis. This modelling is compared to the numerically obtained answer for the angle $\theta = \tan^{-1}(1/2)$ in figure 2.6(a) and shows good agreement.

There is also a special case highlighted in figure 2.6. This is when $\theta = \pi/4$ and $E_{\tilde{x}} = E_{\tilde{y}}$. The zero-temperature degeneracy leads to the domain wall having an arbitrary choice of proceeding in the $\tilde{x}$ or $\tilde{y}$ directions, which manifests itself in the free-energy gap having non-zero gradient at zero temperature. We would expect, and indeed find, the slope at that point to be $2\sqrt{2}\log 2$; the $\log 2$ reflects the choice of two equal directions, the outer factor of two the pair of domain walls, and the $\sqrt{2}$ is the length scaling we previously discussed. The modelling of equation (2.14) also works remarkably well for this rotation, as shown in figure 2.6(b).

This section has been one of success, albeit limited. We have made the most minor change to our first system and shown that our interpretation is robust to this change. However, we have not even changed the fundamental Hamiltonian, instead just altered the direction in which the transfer matrix operates. We ought not, then, be too proud of ourselves; much is left to be done if we are to be convinced.

2.3 Bilayer lattice

This next section aims to be more persuasive. Not only will our topological excitation picture survive, it will thrive as we find excitations described by objects which are not merely the simple domain walls of the past sections. This shall still be done within an Ising framework, but now we shall change the interactions in the Hamiltonian to those of a bilayer system. This section is taken from reference [1], though the aim of that paper was different to the current goal. There we were attempting to argue for some technical alterations to a particular scaling law involving multiple-layered systems, specifically treating each layer as an effective magnetic field for its neighbours and using this to find the critical exponent for magnetic susceptibility, γ [42–61]. Such technicalities do not concern us here, though analysis of the accuracy of our numerical technique in finding the transition temperature of a model may be of interest if one is both of a numerical temperament and sceptical of our use of polynomial extrapolation of exact diagonalisation results. For the purpose of this section, however, we shall instead be leaning on just the section entitled *Modelling the topological excitations*.

First, however, let us begin with our now familiar starting point equation (2.1),

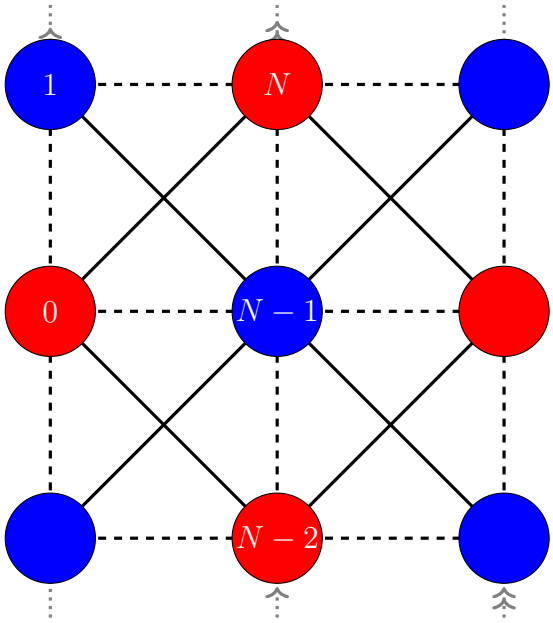


Figure 2.7: Depiction of helical boundary conditions as described in the text. In the thermodynamic limit this leads to the square lattice with nearest neighbour (dashed lines) and second-nearest neighbour (solid lines) interactions, as described by equation (2.18). This may be thought of as a two-layered system, coupled by the dashed bonds, though otherwise the red and blue sites are identical.

again using Ising spins. To form our desired lattice we shall choose

$$a_1 = a_{N-1} \equiv J_1 \,, \tag{2.16}$$

$$a_{N-2} = a_N \equiv J_2 \,, \tag{2.17}$$

with all others zero. This is depicted in figure 2.7 and leads to the Hamiltonian of the square lattice with nearest and second-nearest neighbour interactions

$$H = -J_1 \sum_{\langle jj' \rangle_1} \sigma_j \sigma_{j'} - J_2 \sum_{\langle jj' \rangle_2} \sigma_j \sigma_{j'} \,. \tag{2.18}$$

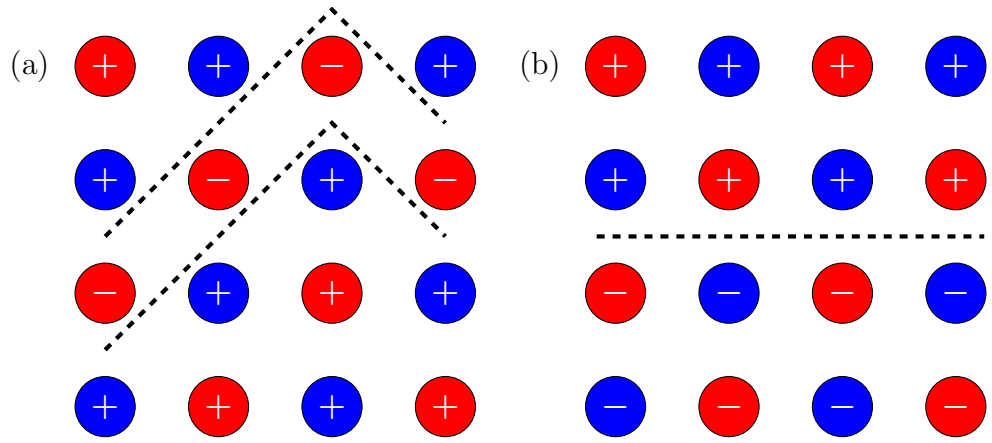


Figure 2.8: Depiction of two styles of topological excitations. The red and blue sites indicate different layers, while + and − indicates the value of the Ising spin at each site. a) A localized excitation, with a line of flipped spins in one layer. b) A domain wall running through both layers, separating the system into a region of + spins and a region of − spins. For a periodic system these must appear in pairs, though they may be well separated and so be thought of as independent.

We again construct a transfer matrix from equations (2.3) and (2.4), writing

$$\hat{T} = \hat{t}^{N-1} \equiv e^{-\beta \hat{F}}\,, \tag{2.19}$$

and similarly defining a free-energy gap as

$$\Delta F_m = \frac{N-1}{\beta} \log \frac{t_0}{t_m}\,. \tag{2.20}$$

Figure 2.8 depicts two natural styles of topological excitations for this model. These were obtained by observation of the numerical data and careful thought. It may be harder to systematically find all such excitations for a more sophisticated model. The first is a line of flipped spins in just one of the layers.

Every horizontal step breaks two intra- and four inter-layer bonds, resulting in a energy cost of $8J_1 + 4J_2$ compared to the groundstate. The second is a domain wall which cuts through both layers, leaving a region of up spins on one side and down spins on the other. As this is a periodic system, such domain walls must come in pairs and hence the energy cost of such an excitation is $2(2J_1 + 4J_2)$. Clearly for low $\lambda \equiv J_1/J_2$ the localised excitation is cheaper and so at low temperature would be preferred. However, this model is in the 2D Ising universality class [1] and hence the phase transition is controlled by excitations of the second type, as discussed in our first example of the unrotated square lattice. As such there must be a crossover between the two as temperature is increased, caused by the aforementioned fluctuations. In essence we expect thermal fluctuations to cause the localised excitation to spread out over some range, whose average is determined by the temperature, though the edges remain bound. This costs an energy proportional to λ multiplied by the range and thus cannot expand indefinitely. Once sufficiently spread out, it is preferable to instead flip some of the spins in the other layer to match those of the excitation. Such an unbound excitation is then indistinguishable from the independent pair of domain walls, albeit dressed with thermal fluctuations. It is this prediction that we will use to test the validity of our assertion.

We can again use a simple partition sum argument to model each style of topological excitation. For the localised excitation the free-energy cost may be

written as

$$\Delta F_{\text{local}} = -T \log \left(2e^{-\beta(8J_1+4J_2)} + e^{-\beta(4J_1+8J_2)} \right) . \tag{2.21}$$

The three exponents come from the three directions the excitation can choose to go in a given step: 45° up, 45° down, or horizontally across. The first two of course have the same energetic cost as we previously discussed, while the final may be thought of as the cheapest fluctuation. Similarly for the Ising domain wall excitation we may write

$$\Delta F_{\text{Ising}} = -2T \log \left(e^{-\beta(2J_1+4J_2)} + 2e^{-\beta(4J_1+4J_2)} \right) . \tag{2.22}$$

This is equivalent to equation (2.6).

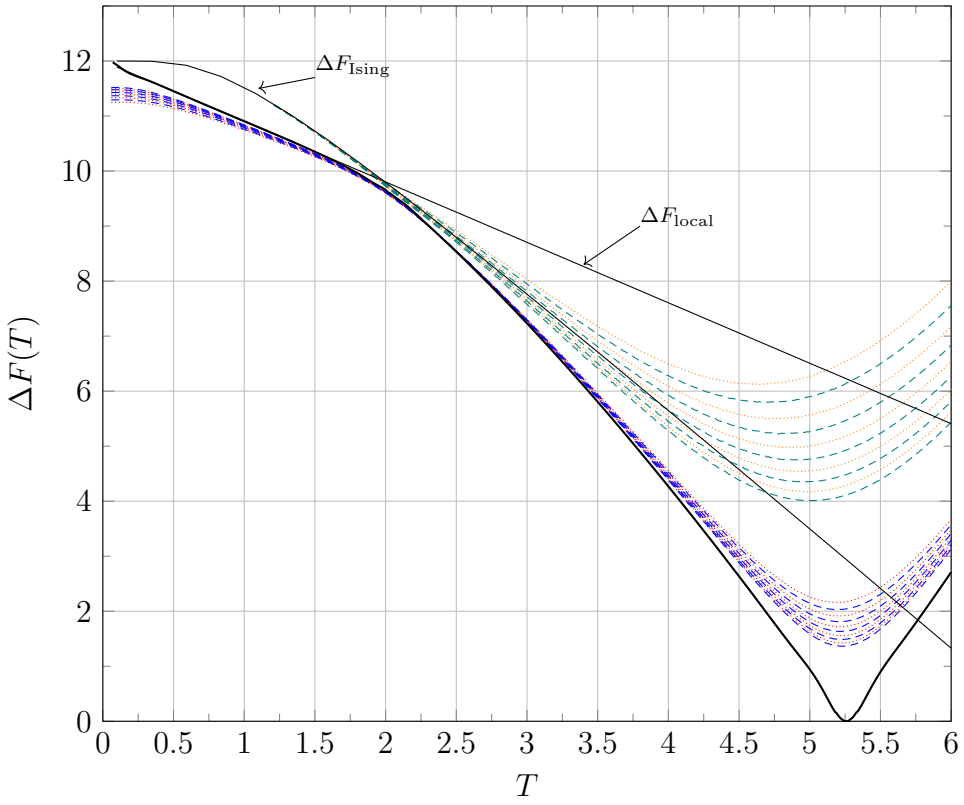


Figure 2.9: The free-energy difference for the first and second excitations for the J_1-J_2 model at $\lambda \equiv J_1/J_2 = 1$, where $N = 16$-25. Odd data is dashed while even data is dotted. The solid lines are fits to models of topological excitations described in the text along with an extrapolation to infinity from the odd curves for the first excited state.

In figure 2.9 we plot the first two excitations, that is ΔF_1 and ΔF_2 in the language of equation (2.20), for the J_1-J_2 system for $N = 16$-21 and $\lambda \equiv J_1/J_2 = 1$. We overlay each of the two approximate models given above. As expected, for very low temperatures the localised excitation model (2.21) fits remarkably well to the first excitation and the Ising domain wall description (2.22) to the second. The two models then cross around $T \approx 2$ where indeed we see an avoided crossing in the exact data, indicating a change of behaviour. Beyond

this point neither line fits well as both are in fact low-temperature expansions and low-temperature pictures. To improve the fit more expensive fluctuations would need to be included. We have tested this modelling for various values of λ and find similar results to that just described. Changing the bond ratio λ alters where this crossover takes place. When λ is small the bonds between layers are weak compared to those within each layer, meaning that the localised excitation can spread out further within one layer before it changes behaviour. This pushes the crossover closer to the transition temperature.

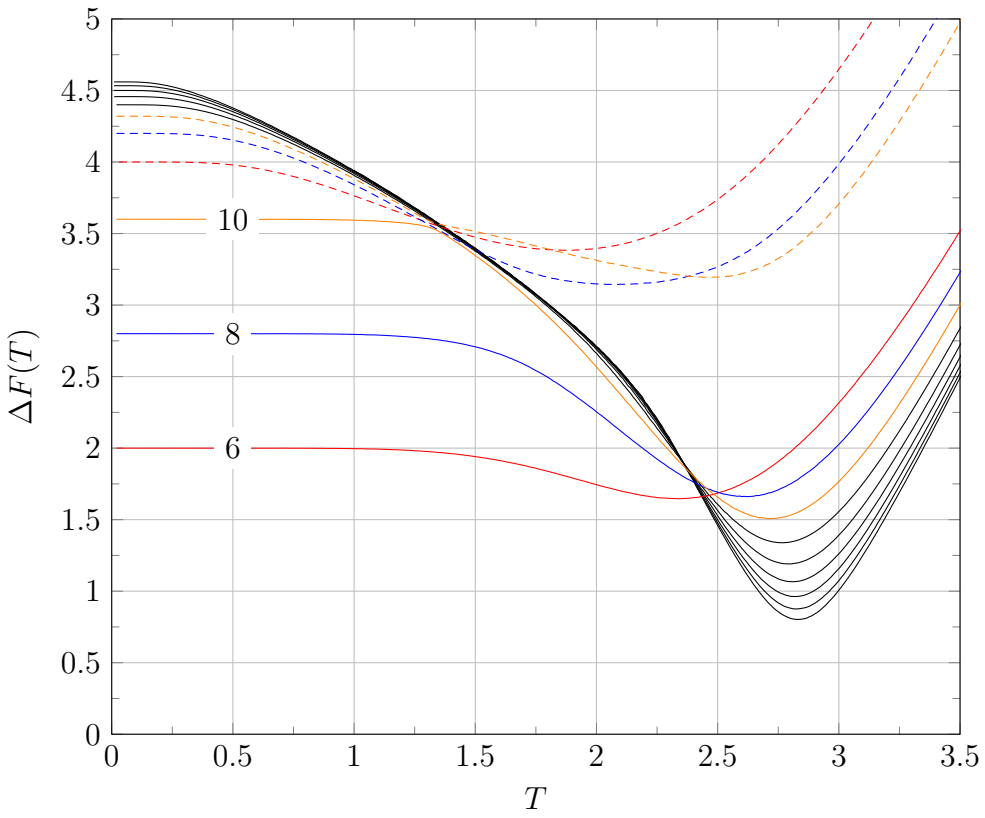


Figure 2.10: The free-energy difference for the J_1-J_2 model for $\lambda \equiv J_1/J_2 = 0.1$, for even systems with N = 6-22. The solid lines are first excitations and the dashed lines second excitations for N = 6, 8, 10.

Finally, there is a style of excitation only present in some even sized systems. These excitations are when one layer is of opposite spin to the other, incurring an energy cost of $4(N-1)J_1$. This can be seen in Fig. 2.10 for $\lambda \equiv J_1/J_2 = 0.1$, where systems of size 6, 8, and 10 all exhibit this phenomenon. Initially the free-energy cost of these excitations is constant, as there is no natural way for these to gain entropically, before there is an avoided crossing similar to what was previously described. After this point the lines behave as usual. These superfluous, from the point of view of the thermodynamic limit, excitations impact any extrapolation. Odd systems, of course, have Möbius boundary conditions, meaning there is only one lattice.

In this example, then, we have seen that the simple picture of topological excitations provide clear physical interpretations in a variety of situations. It predicts crossings of eigenvalues from an expected change of behaviour of excitations and even helps one to understand seemingly irrelevant excitations in small systems. Thus it provides a helpful intuition for otherwise mysterious numerical results. This model is in many ways the most useful of our examples; it is certainly the most convincing.

2.4 Summary

To summarise, this section has developed a numerical basis for our future work. We have been able to see the eigenvalues of transfer matrices in a physical light, namely as the free-energetic cost of a topological excitation. These excitations, which can also be thought of through a quasi-particle lens, can be described through a simple picture inferred from the zero-temperature limit then extended beyond this point. We use the word topological to indicate that each excitation is split by an infinite energy barrier from another, and that the number of domain walls, in the case of the ordinary square lattice Ising model, acts as a topological number. This terminology may not be appropriate in the case of more sophisticated models.

Several distinct pieces of modelling over differing geometries has lent credence to this view, extending a zero-temperature picture into low temperature. Of course, this is still restricted to low temperature until better modelling is performed.

Let us review the examples of this section with this terminology in mind. First, the ordinary square lattice Ising model. Here we saw that the excitations of the transfer matrix corresponded, at zero temperature, to pairs of domain walls of infinite extent. Through modelling we could extend that picture away from zero temperature by including fluctuations. To be clear, we define a topological

excitation by its zero temperature behaviour then try to extend the interpretation using modelling. There is still, even with these fluctuations, an infinite energy gap between this excited state and the groundstate. As we know that the quasi-particles for the square lattice Ising model are independent, these pairs of domain walls can be arbitrarily close to one another; the separation in the domain wall picture maps to momenta in the quasi-particle picture.

Next, let us turn to the bilayer Ising example. In this case we saw a interesting change of behaviour from one style of topological excitation to another. Again, each excitation is defined by its zero-temperature behaviour and modelling extends the interpretation to low temperature. The first style of excitation, which dominates the low temperature behaviour for bilayer models with sufficiently weak between-layer bonds, is a bound pair of domain walls in one of the layers. If fluctuations could destroy these domain walls then the excitation would not be distinct from the groundstate and hence the plot in figure 2.9, for example, would hit zero immediately. This does not occur and so this bound pair of domain walls must be protected. Indeed, as eigenvectors of a Hermitian matrix are forced to retain their integrity through the requirement that eigenvectors with different eigenvalues must be orthogonal, mathematically these excitations cannot appear and disappear. In reality, what happens as fluctuations become more relevant is that the bound pair become less strongly bound and they spread out over one layer, until it is free-energetically favourable to turn

into the independent pair of domain walls cutting through both layers we see closer to the transition temperature.

We have also seen the usefulness of transfer matrices as a tool for examining different square lattice models. It is this usefulness which we shall attempt to build upon for the remainder of this work.

CHAPTER 3

THE BAKER-CAMPBELL-HAUSDORFF FORMULA

The major mathematical result of this thesis is contained within this chapter. This has two unfortunate consequences. Firstly, it means that a paper has already been written about this result, namely [2], and as such this chapter will be largely drawn from that work. The tone and mathematical level may therefore not be perfectly suited to a thesis, but we have attempted to alleviate this issue by including several more examples and explanations. If one, quite sensibly, prefers a terser treatment then [2] is perhaps a more enjoyable read. Secondly, a physically inclined reader may find the absence of any real physics,

beyond a passing mention in this introduction, off-putting. Such a reader, then, is invited to skip vast swathes of this chapter. One does not need to know how to prove something in order to use it, after all. They could join us again in section 3.3, wherein we give some examples of our result, then continue on to the following chapter where physics shall once again return.

With the administrative preamble being over, let us now turn to the actual content of this chapter. The title, *The Baker-Campbell-Hausdorff formula*, may elicit nods of understanding or perhaps a murmur of expectation from a reader who is familiar with this theorem [5,62–71]. We ended the last chapter, after all, with a formula

$$\hat{T} = \hat{T}_{\perp}^{\frac{1}{2}} \, T_{/\!/} \hat{T}_{\perp}^{\frac{1}{2}} \, , \tag{3.1}$$

or more pertinently

$$e^{-\beta \hat{F}} = e^{-\beta \hat{F}_{/\!/}} \, e^{-\beta \hat{F}_{\perp}} \, e^{-\beta \hat{F}_{/\!/}} \, . \tag{3.2}$$

In this latter formulation we essentially know both $\hat{F}_{/\!/}$ and $\hat{F}_{\perp}$ for a given problem but wish to find the overall free-energy operator $\hat{F}$. As such a symmetric form of the Baker-Campbell-Hausdorff (BCH) formula

$$\mathcal{S}(A, \, B) = \log \left(e^{A} \, e^{2B} \, e^{A} \right) \, , \tag{3.3}$$

for two matrices A and B, appears painfully relevant. The more standard BCH formula, which can be easily transformed into the above symmetric form, is

usually given as

$$Z(X,\,Y) \equiv \log\left(e^X e^Y\right) = X+Y+\frac{1}{2}[X,\,Y]+\frac{1}{12}\left([X,\,[X,\,Y]]+[Y,\,[Y,\,X]]\right)+\cdots\,, \tag{3.4}$$

and tends to be followed by a statement saying that extending this series is truly difficult and no mere exercise in pattern recognition. The work of Campbell [72, 73], Baker [74–76], Hausdorff [77], and others [5] whose names were omitted for fear the theorems title was becoming too long (Poincaré [78, 79] contemporaneously and more recently Dynkin [80] come to mind), was to show that under certain conditions this series must converge and that every order can be written in terms of commutators. This latter point is important in the theory of Lie algebras, where the commutator is thought of as a Lie bracket and the theorem is really about the relationship between Lie groups and Lie algebras. Luckily for us, however, such considerations are largely irrelevant and we are solely concerned with the practical use of this formula in finding a Z given an X and a Y.

Dynkin [80] found this formula explicitly in terms of commutators for every order, where order means combined powers of X and Y. Unfortunately, this means that if a truncation of the series is to give a good approximation to the full expansion both X and Y must be sufficiently close to zero. Instead, however, there is an alternative representation which contains all orders of X

but is linear in Y. Letting $L_X Y \equiv [X\,, Y]$ denote commutator operators, it is given by

$$Z(X,\, Y) = X + \frac{\frac{1}{2}L_X}{\sinh\left(\frac{1}{2}L_X\right)}\left(e^{\frac{1}{2}L_X}Y\right) + \mathcal{O}(Y^2)\,. \tag{3.5}$$

The aim of this chapter will be to extending this representation to all powers of Y. That is, express $Z(X,\, Y)$ as

$$Z(X\,, Y) = X + \sum_{n=1}^{\infty} G_n\left(e^{\frac{1}{2}L_X}Y\right)^n\,, \tag{3.6}$$

finding explicitly the operators G_n, which will depending non-trivially on commutator operators L_X. This series may be truncated and give a good approximation to the full expansion if only Y is small, as opposed to both X and Y in the previous. The radius of convergence of this new series is not currently known. One would imagine that it is at least as large as that of the regular Baker-Campbell-Hausdorff series, but we have not proved this.

The chapter is structured as follows. Section 3.1 contains the derivation of the main result, that is calculating the operators G_n. It will be split into various parts. Some of the demarcations are superficial, intended to help the reader parse the result more easily, but others do indeed mark a shift of thought, strategy, or ideology. Next, section 3.2 argues, based upon a conjecture, that the result remains a sum of commutators, as would be expected. As we mentioned previously, these sections can be safely ignored by any reader who wishes to

avoid mathematical detail. Instead they may prefer to proceed to section 3.3, where finite examples are given and discussed which provides immediately usable formulae for the G_n. Section 3.4 proves an alternative representation for the operators G_n, which is perhaps more practical as it deals with some apparent singularities which shall be encountered. Finally in section 3.5 we argue that this result is particularly useful in the basis where the non-perturbative matrix is diagonal. In this case the operators become merely functions of real numbers and so it is elementary to perform calculations with them. This will be true throughout the rest of the thesis.

3.1 Derivation of the main result

We start with the symmetric version of the BCH formula we previously mentioned,

$$\mathcal{S}(A,\, B) = \log\left(e^{A}\, e^{2B}\, e^{A}\right)\,. \tag{3.7}$$

While this formulation is more natural to work with than (3.4), we may transform each into the other and so they are equivalent. We shall employ the notation for commutators which shall be used throughout this article, $LB \equiv [A,\, B]$ and $L^n B \equiv [A,\, [A,\, \ldots, [A,\, B]\, \ldots]]$, and thus write the Hadamard formula as

$$e^{A} B e^{-A} = e^{L} B\,. \tag{3.8}$$

This then implies

$$e^A e^B e^{-A} = e^{e^L B}\,, \tag{3.9}$$

from which we may see easily that $\mathcal{S}(A,\, B) = Z(2A,\, 2\exp(-L)B) = Z(2\exp(L)B, 2A)$ and additionally $Z(X,\, Y) = \mathcal{S}(X/2, \exp(L_X/2)Y/2)$. The factors of two are introduced here in order to simplify the final representation.

Our task ahead is to expand equation (3.7). The matrix B will be our focus, with the aim being to write the expansion as a power series in this matrix. Once we achieve this, we will examine the coefficients of the power series in depth and obtain closed form expressions.

We will employ the identity

$$\begin{aligned}\log M &= \log(1-(1-M)) \\ &= -\sum_{l=1}^{\infty}\frac{1}{l}(1-M)^l \\ &= -\sum_{l=1}^{\infty}\frac{1}{l}\sum_{m=0}^{l}(-1)^m\frac{l!}{m!(l-m)!}M^m\,,\end{aligned} \tag{3.10}$$

setting $M = \exp(A)\exp(2B)\exp(A)$. Note, this is expected to converge if M is in a neighbourhood of the identity. We will find that M^m separates into the sum of several parts. Each of these parts will take the form $f_i \exp(2mA) g_i$, for m-independent quantities f_i and g_i. The f_i and g_i may then each be pulled out of the above sums, leaving $\exp(2mA)$ in place of M^m. We will then use the

identity in reverse to obtain $\log(M) = \sum_i f_i \, 2A \, g_i$. This then constitutes the fundamental mathematical approach and is hopefully a sufficient indication of what is to come.

3.1.1 Expanding M^m in powers of B

We will now focus on calculating M^m. We may use the Hadamard formula (3.8) to symmetrically move exponentials of A to the edges. For example, we may write

$$M^2 = e^A e^{2B} e^A e^A e^{2B} e^A \tag{3.11}$$

$$= e^{2A} \left(e^{-A} e^{2B} e^A\right) \left(e^A e^{2B} e^{-A}\right) e^{2A} \tag{3.12}$$

$$= e^{2A} e^{2e^{-L}B} e^{2e^L B} e^{2A} \,, \tag{3.13}$$

and

$$M^3 = e^A e^{2B} e^A e^A e^{2B} e^A e^A e^{2B} e^A \tag{3.14}$$

$$= e^{3A} \left(e^{-2A} e^{2B} e^{2A}\right) e^{2B} \left(e^{2A} e^{2B} e^{-2A}\right) e^{3A} \tag{3.15}$$

$$= e^{3A} e^{2e^{-2L}B} e^{2e^{0L}B} e^{2e^{2L}B} e^{3A} \,. \tag{3.16}$$

Continuing this idea we find

$$M^m = e^{mA}\left[\prod_{n=-\frac{m-1}{2}}^{\frac{m-1}{2}} \exp\left(2e^{2nL}B\right)\right]e^{mA}\,, \tag{3.17}$$

where we must take the product in the correct order, namely increasing n. We may then Taylor expand the exponentials involving B

$$M^m = e^{mA}\left[\prod_{n=-\frac{m-1}{2}}^{\frac{m-1}{2}} \sum_{k_n=0}^{\infty} \frac{1}{k_n!}\left(2e^{2nL}B\right)^{k_n}\right]e^{mA}\,, \tag{3.18}$$

and gather terms in orders of B,

$$M^m = e^{mA}\left[1 + 2\left(\sum_{-\frac{m-1}{2}\leq n_1\leq\frac{m-1}{2}} e^{2n_1L}B\right)\right.$$
$$\left.+2^2\left(\sum_{-\frac{m-1}{2}\leq n_1<n_2\leq\frac{m-1}{2}} e^{2n_1L}Be^{2n_2L}B + \frac{1}{2!}\sum_{-\frac{m-1}{2}\leq n_1\leq\frac{m-1}{2}} e^{2n_1L}Be^{2n_1L}B\right) + \cdots\right]e^{mA}\,.$$

This will be described in more detail throughout this section.

In the above expression, we must think of each term $\exp(2n_iL)B$ as one object - that particular commutator operator L is acting on that particular matrix B and so the two are intrinsically linked. It is helpful to formalise this link, labelling the pair with an index. Then we understand that the operator L_i acts on only the matrix B_i, and no other. We can then label each such pair. This allows the commutation of operators and matrices with different labels,

enabling us to pull out all matrices B out of each sum in the above expression. Explicitly,

$$M^m = e^{mA}\left[1 + 2\left(\sum_{-\frac{m-1}{2}\le n_1\le\frac{m-1}{2}} e^{2n_1L_1}\right)B_1 + 2^2\left(\sum_{-\frac{m-1}{2}\le n_1<n_2\le\frac{m-1}{2}} e^{2n_1L_1}e^{2n_2L_2} + \frac{1}{2!}\sum_{-\frac{m-1}{2}\le n_1\le\frac{m-1}{2}} e^{2n_1(L_1+L_2)}\right)B_1B_2 + \cdots\right]e^{mA} \tag{3.19}$$

$$\equiv e^{mA}\left[F_0 + F_1(L_1)\,B_1 + F_2(L_1,\,L_2)\,B_1B_2 + F_3(L_1,\,L_2,\,L_3)\,B_1B_2B_3 + \cdots\right]e^{mA}\,. \tag{3.20}$$

3.1.2 Rewriting F_N in terms of fundamental sums S_N

Our first aim has thus been achieved; we have expanded the formula (3.7) with a power series in the matrix B. The next is to find closed form expressions for the coefficients F_N. First we define the sum S_N as

$$S_N(L_1,\,L_2,\,\ldots,\,L_N) \equiv 2^N \sum_{-\frac{m-1}{2}\le n_1<n_2<\cdots<n_N\le\frac{m-1}{2}} e^{2n_1L_1}e^{2n_2L_2}\cdots e^{2n_NL_N}\,, \tag{3.21}$$

then the first few of the coefficients F_N are given by

$$\begin{aligned} F_0 &= 1\,, \\ F_1(L_1) &= S_1(L_1)\,, \\ F_2(L_1,\,L_2) &= S_2(L_1,\,L_2) + \frac{2}{2!}S_1(L_1+L_2)\,, \\ F_3(L_1,\,L_2,\,L_3) &= S_3(L_1,\,L_2,\,L_3) + \frac{2}{2!}\Big(S_2(L_1+L_2,\,L_3) + S_2(L_1,\,L_2+L_3)\Big) + \frac{2^2}{3!}S_1(L_1+L_2+L_3)\,. \end{aligned} \tag{3.22}$$

Writing the coefficients F_N for an arbitrary order N is a problem in partitioning. As seen in the above examples, the string $L_1 + L_2 + \cdots + L_N$ is split in all possible ways. The resultant substrings are then used as arguments for the sums S_n. However, each sum is also divided by factorials. These factorials are determined by the length of the substrings used as arguments. For example, the string $L_1 + L_2 + L_3$ may be split in the following ways giving the following factorials:

$$\begin{aligned} L_1 + L_2 + L_3 \quad &\longrightarrow \quad 3!\,, \\ L_1 + L_2 \;,\; L_3 \quad &\longrightarrow \quad 2!\,1!\,, \\ L_1 \;,\; L_2 + L_3 \quad &\longrightarrow \quad 1!\,2!\,, \\ L_1 \;,\; L_2 \;,\; L_3 \quad &\longrightarrow \quad 1!\,1!\,1!\,, \end{aligned} \tag{3.23}$$

demonstrating how F_3 was constructed in equation (3.22).

There are then two major hurdles to finding closed form expressions for each

coefficient of the power series. The first is to calculate the explicit sum S_N. As we may think of the sum S_N as N finite geometric series, it may be expected to have 2^N terms. However, it may be split into $N+1$ parts, each of which is a collection of infinite geometric series. This lifting of the constraint is crucial and will be discussed shortly. The second hurdle is then to perform the partition sum, that is to calculate F_N given the functions S_r.

3.1.3 Calculating S_N

It is useful at this point to deal with a concrete example. Consider the sum

$$S_2(L_1,\, L_2) \equiv \sum_{-\frac{m-1}{2} \leq n_1 < n_2 \leq \frac{m-1}{2}} 2^2 e^{n_1 L_1} e^{n_2 L_2}\,. \tag{3.24}$$

The summation variables, n_1 and n_2, are constrained from both above and below. We may think of these constraints as forming a triangle, as depicted in figure 3.1. The sum may then be thought of as the combination of three semi-constrained sums, constructed by taking a given vertex of the triangle

and extending the constraining lines to form infinite sectors. Explicitly,

$$\sum_{-\frac{m-1}{2}\leq n_1<n_2\leq\frac{m-1}{2}} 2^2 e^{2n_1L_1}e^{2n_2L_2} = 2^2\left(\sum_{n_1<n_2\leq\frac{m-1}{2}} e^{2n_1L_1}e^{2n_2L_2}\right) - 2^2\left(\sum_{n_1<-\frac{m-1}{2}} e^{2n_1L_1}\sum_{n_2\leq\frac{m-1}{2}} e^{2n_2L_2}\right) + 2^2\left(\sum_{n_2\leq n_1<-\frac{m-1}{2}} e^{2n_1L_1}e^{2n_2L_2}\right), \tag{3.25}$$

or, using the labels for regions shown in figure 3.1,

$$\textcircled{1} \;=\; \Big(\textcircled{1}+\textcircled{2}\Big) \;-\; \Big(\textcircled{2}+\textcircled{3}\Big) \;+\; \textcircled{3}\,. \tag{3.26}$$

We may then evaluate the sums on the right-hand side to obtain

$$S_2(L_1,\,L_2) = \frac{\coth(L_1)-1}{\sinh(L_1+L_2)}e^{m(L_1+L_2)} + \frac{1}{\sinh(-L_1)}\frac{1}{\sinh(L_2)}e^{m(-L_1+L_2)} + \frac{\coth(-L_2)-1}{\sinh(-L_1-L_2)}e^{m(-L_1-L_2)}\,.$$

It is here that we shall see our first substantial divergence from the paper [2]; the generalised calculation and proof thereof was, in that work, relegated to an appendix. As this is a thesis and thus intended to be ever-so-slightly more pedagogical instead we shall perform this calculation now. Generalising this idea to the sum S_N involves $N+1$ vertices of an N-dimensional tetrahedron. The constraining lines are extended, creating $N+1$ sums similar to those in

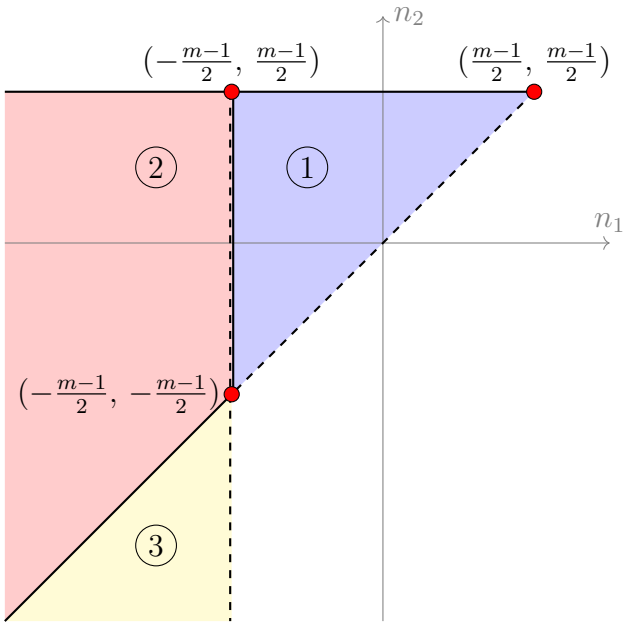


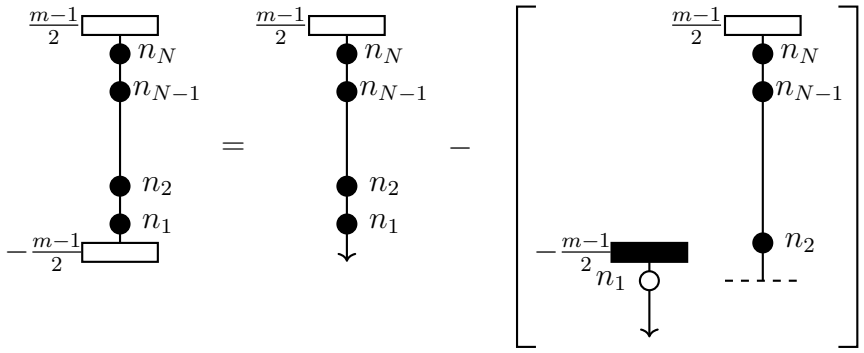
Figure 3.1: A depiction of the parameter space of n_1 and n_2 in equation (3.25). Solid lines imply inclusiveness of that line in a given sum, while dashed imply the line of parameters is excluded. The variables of the original sum are constrained to the triangle formed from the vertices marked with a red circle.

equation (3.25).

If we wish to calculate the sum S_N, as defined in (3.21), first we ought to note that

$$\sum_{-\frac{m-1}{2}\leq n_1<n_2<\cdots<n_N\leq\frac{m-1}{2}} = \sum_{n_1<n_2<\cdots<n_N\leq\frac{m-1}{2}} - \sum_{-\frac{m-1}{2}>n_1<n_2<\cdots<n_N\leq\frac{m-1}{2}}, \quad (3.27)$$

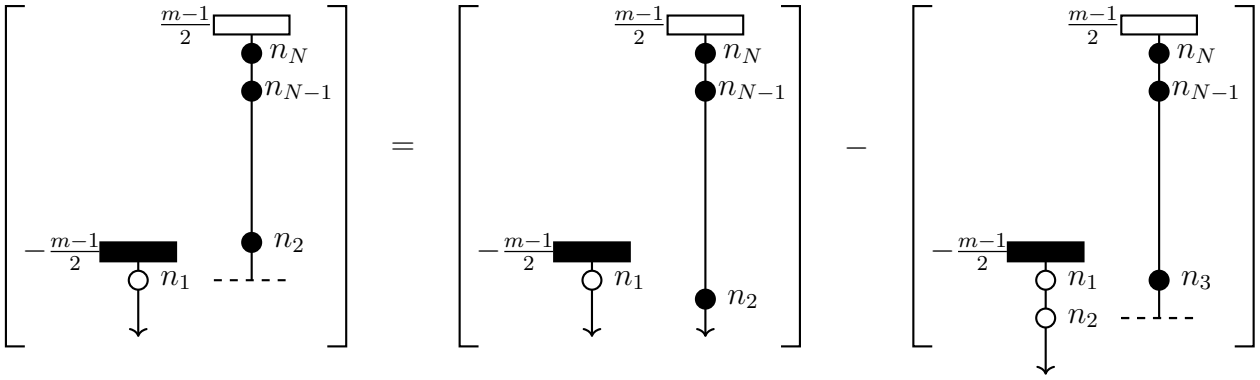
which is demonstrated by the diagram below. In this, circles represent the variables of the sum and their position along the line indicates the value said variables take, while rectangles represent the bounds of the sums. Open rectangles and circles allow equality, while filled do not.



This identity has transformed the constrained sum on the left into two sums. One of these is semi-constrained, as we were targeting, while the other has one semi-constrained and $N-1$ constrained variables. Applying this idea again gives

$$\sum_{-\frac{m-1}{2}>n_1<n_2<\cdots<n_N\leq\frac{m-1}{2}} = \sum_{\substack{n_1<-\frac{m-1}{2}\\ n_2<\cdots<n_N\leq\frac{m-1}{2}}} - \sum_{-\frac{m-1}{2}>n_1\geq n_2<\cdots<n_N\leq\frac{m-1}{2}}, \qquad (3.28)$$

or pictorially,



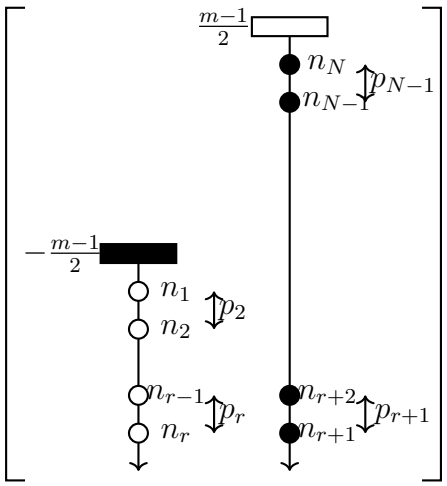
The number of constrained variables on the right hand side is now reduced to $N-2$. We can continue this until there are no such variables remaining,

resulting in an identity relating a sum with N constrained variables to $N+1$ sums with only semi-constrained variables.

A generic term in this identity for the particular sum (3.21) is given by

$$2^N \sum_{\substack{n_r \le \cdots \le n_1 < -\frac{m-1}{2} \\ n_{r+1} < \cdots < n_N \le \frac{m-1}{2}}} e^{2n_1 L_1} e^{2n_2 L_2} \cdots e^{2n_N L_N}$$
$$= \left[(-1)^r 2^r \sum_{n_r \le \cdots \le n_1 < -\frac{m-1}{2}} e^{2n_1 L_1} \cdots e^{2n_r L_r} \right] \left[2^{N-r} \sum_{n_{r+1} < \cdots < n_N \le \frac{m-1}{2}} e^{2n_{r+1} L_{r+1}} \cdots e^{2n_N L_N} \right],$$

that is,



A simple change of variables, indicated on the picture above, gives

$$\left[(-1)^r 2^r \sum_{n_1=-\infty}^{-\frac{m-1}{2}-1} e^{2n_1(L_1+\cdots+L_r)} \sum_{p_2=-\infty}^{0} e^{2p_2(L_2+\cdots+L_r)} \cdots \sum_{p_r=-\infty}^{0} e^{2p_r L_r} \right]$$
$$\times \left[2^{N-r} \sum_{p_{r+1}=-\infty}^{-1} e^{2p_{r+1}L_{r+1}} \cdots \sum_{p_{N-1}=-\infty}^{-1} e^{2p_N - 1(L_{r+1}+\cdots+L_{N-1})} \sum_{n_N=-\infty}^{\frac{m-1}{2}} e^{2n_N(L_{r+1}+\cdots+L_N)} \right],$$

which may be trivially calculated. Using the identities

$$-2\sum_{n=-\infty}^{-\frac{m-1}{2}-1} e^{2nx} = \frac{-2e^{-(m+1)x}}{1-e^{-2x}} = \frac{e^{-mx}}{\sinh(-x)}\,, \qquad -2\sum_{n=-\infty}^{0} e^{2nx} = \frac{-2}{1-e^{-2x}} = \coth(-x)-1\,, \tag{3.29}$$

and

$$2\sum_{n=-\infty}^{\frac{m-1}{2}} e^{2nx} = \frac{2e^{(m-1)x}}{1-e^{-2x}} = \frac{e^{mx}}{\sinh(x)}\,, \qquad 2\sum_{n=-\infty}^{-1} e^{2nx} = \frac{2e^{-2x}}{1-e^{-2x}} = \coth(x)-1\,, \tag{3.30}$$

we find that the final result is

$$S_N(L_1,\,\ldots,\,L_N) = \sum_{r=0}^{N} \tilde{S}_r(-L_r,\,-L_{r-1}\,,\ldots,\,-L_1)\tilde{S}_{N-r}(L_{r+1},\,L_{r+2},\,\ldots,\,L_N) \times e^{m(-L_1-\cdots-L_r+L_{r+1}+\cdots+L_N)}\,, \tag{3.31}$$

where $\tilde{S}_0 \equiv 1$ and

$$\tilde{S}_r(x_1,\,\ldots,\,x_r) = \frac{s_{r-1}(x_1,\,\ldots,\,x_{r-1})}{\sinh\left(x_1+x_2+\cdots+x_r\right)} \qquad \text{for } r \in \mathbb{Z}^+\,, \tag{3.32}$$

where similarly $s_0 \equiv 1$ and

$$s_{r-1}(x_1,\,\ldots,\,x_{r-1}) = \prod_{j=1}^{r-1}\left[\coth(x_1+x_2+\cdots+x_j)-1\right] \qquad \text{for } (r-1) \in \mathbb{Z}^+\,. \tag{3.33}$$

There are several things we ought to note from this result. Firstly, the $N+1$

different forms that the exponential above may take clearly correspond to the vertices of the N-dimensional tetrahedron discussed previously. As we mentioned earlier, this exponential, containing all the m-dependence, is crucial in reversing the identity (3.10). Next, note the splitting of each term into $\tilde{S}_r$ and $\tilde{S}_{N-r}$ functions. This structure remains for the coefficients F_N, as we shall see shortly, and appears fundamental to the problem. Furthermore, the representation of the result in hyperbolic functions is perhaps not unexpected; previous results in equation (3.5) showed that the order B term is best written with a sinh function. Finally, the arguments of the hyperbolic functions only ever contain sums of the commutator operators L_i. As such it is mathematically sensible to think of the active variables not as these commutator operators L_1, L_2, L_3 etc., but rather as strings of such operators, for example L_1, $L_1 + L_2$, $L_1 + L_2 + L_3$ etc. We will discuss such strings in much more detail in later sections, in particular section 3.5.

3.1.4 Rewriting F_N as a partition sum in terms of f_r

Our next task is to perform the partition sum, or in other words calculate F_N given the now known S_r. Once again it is useful to turn to an example. Using

the above results, we can simply read off that

$$F_3(L_1,\, L_2,\, L_3) \equiv S_3(L_1,\, L_2,\, L_3) + \frac{2}{2!}\Big[S_2(L_1,\, L_2+L_3) + S_2(L_1+L_2,\, L_3)\Big] + \frac{2^2}{3!}S_1(L_1+L_2+L_3) \tag{3.34}$$

$$= C_0 e^{m(L_1+L_2+L_3)} + C_1 e^{m(-L_1+L_2+L_3)} + C_2 e^{m(-L_1-L_2+L_3)} + C_3 e^{m(-L_1-L_2-L_3)} \tag{3.35}$$

where

$$C_0 = \frac{\big(\coth(L_1)-1\big)\big(\coth(L_1+L_2)-1\big) + \frac{2}{2!}\Big[\big(\coth(L_1)-1\big) + \big(\coth(L_1+L_2)-1\big)\Big] + \frac{2^2}{3!}}{\sinh\left(L_1+L_2+L_3\right)},$$

$$C_1 = \left[\frac{1}{\sinh\left(-L_1\right)}\right]\left[\frac{\big(\coth(L_2)-1\big)+\frac{2}{2!}}{\sinh\left(L_2+L_3\right)}\right],$$

$$C_2 = \left[\frac{\big(\coth(-L_2)-1\big)+\frac{2}{2!}}{\sinh\left(-L_1-L_2\right)}\right]\left[\frac{1}{\sinh\left(L_3\right)}\right],$$

$$C_3 = \frac{\big(\coth(-L_3)-1\big)\big(\coth(-L_2-L_3)-1\big) + \frac{2}{2!}\Big[\big(\coth(-L_3)-1\big) + \big(\coth(-L_2-L_3)-1\big)\Big] + \frac{2^2}{3!}}{\sinh\left(-L_1-L_2-L_3\right)}$$

Within this example many of the previous themes are exposed. As in the sums S_N, the result splits into $N+1$ terms. Each of these terms likewise separate into an m-dependent exponential and an m-independent function (the C_i above). The final similarity is the factorisation of these functions, shown clearly in C_1 and C_2. More generally, this factorisation arises from partitioning. Any sums which contribute to the coefficient of a given exponential with argument

$m(-L_1 - \cdots - L_r + L_{r+1} + \cdots + L_N)$ must contain a partition between L_r and L_{r+1}. Any other partitioning which occurs to the left of the split affects a given sums contribution to the term independently of any partitioning to the right. More concretely, in the example above the function C_1, associated with the exponential with argument $m(-L_1 + L_2 + L_3)$, is contributed to by any sums in equation (3.34) with a partition between L_1 and L_2. These are $S_3(L_1,\, L_2,\, L_3)$ and $S_2(L_1,\, L_2 + L_3)$. In the former there is another partition between L_2 and L_3, giving rise to the coth term in the right factor of C_1, while in the latter there is no such extra partition.

These arguments necessitate the partition sum to take the form

$$F_N(L_1,\, L_2,\, \ldots,\, L_N) = \sum_{r=0}^{N} \tilde{F}_r(-L_r, -L_{r-1}\,, \ldots,\, -L_1)\tilde{F}_{N-r}(L_{r+1},\, L_{r+2},\, \ldots,\, L_N) \times e^{m(-L_1-\cdots-L_r+L_{r+1}+\cdots+L_N)}\,, \quad (3.36)$$

where $\tilde{F}_0 \equiv 1$ and

$$\tilde{F}_r(x_1,\, \ldots,\, x_r) = \frac{f_{r-1}(x_1,\, \ldots,\, x_{r-1})}{\sinh\,(x_1 + x_2 + \cdots + x_r)} \qquad \text{for } r \in \mathbb{Z}^+\,. \quad (3.37)$$

3.1.5 A partition formula for f_r

The function $f_{r-1}(x_1,\, x_2,\, \ldots,\, x_{r-1})$ will be a partition sum of the functions s_n which are given from (3.33). For example, we find

$$\begin{aligned} f_2(x_1,\, x_2) &= s_2(x_1,\, x_2) + \frac{2}{2!}\Big(s_1(x_1) + s_1(x_2)\Big) + \frac{2^2}{3!} \\ &= \big[\,\coth(x_1) - 1\big]\big[\,\coth(x_1 + x_2) - 1\big] + \frac{2}{2!}\Big(\big[\,\coth(x_1) - 1\big] + \big[\,\coth(x_1 + x_2) - 1\big]\Big) + \frac{2^2}{3!}\,, \end{aligned}$$

in both C_0 and C_3 above. In general, $f_{r-1}(x_1,\, x_2,\, \ldots,\, x_{r-1})$ is a sum of terms, each involving a product of coth functions minus one. As we have just seen for for f_2, in each of these terms there will be a number of these functions missed out. In a term where m such functions in a row have been missed out, the coefficient will be $a_{m+1} \equiv 2^m/(m+1)!$. This then implies that

$$\begin{aligned} f_{r-1} = a_r + \sum_{n=1}^{r-1} \sum_{p_1=1}^{\infty} \cdots \sum_{p_{n+1}=1}^{\infty} a_{p_1} a_{p_2} \cdots a_{p_{n+1}} \\ \times \big[\,\coth(x_1 + x_2 + \cdots + x_{p_1}) - 1\big] \cdots \big[\,\coth(x_1 + x_2 + \cdots + x_{p_1 + \cdots + p_n}) - 1\big]\, \delta_{r,\, p_1 + \cdots + p_{n+1}}\,, \end{aligned} \tag{3.38}$$

where $\delta_{i,\,j}$ is the Kronecker delta. We must now understand the combinatorial aspect of partitioning expressed in this sum.

While superficially complicated, this sum is actually very simple. In essence, the sum index n counts how many coth functions have not been missed out

and the numbers p_i give the positions of these. Alternatively, the numbers $p_i - 1$ can be interpreted as counting how many functions have been missed out in a row. As an example, one of the terms in the functions f_4 which has two coth functions missing (so $n = 2$ remain) is

$$\underbrace{\big[\coth(x_1) - 1\big]}_{p_1 - 1 = 1}\big[\coth(x_1 + x_2) - 1\big]\underbrace{\big[\coth(x_1 + x_2 + x_3) - 1\big]}_{p_2 - 1 = 0}\underbrace{\big[\coth(x_1 + \cdots + x_4) - 1\big]}_{p_3 - 1 = 2}$$

In the function f_{r-1} there are $r - 1$ different coth functions; for example, $f_2(x_1, x_2)$ has $\coth(x_1)$ and $\coth(x_1 + x_2)$. The sum index n indicates the number of coth functions in a given term. If there are only n such functions in a term, that means $(r - 1) - n$ are missing. These missing coth functions determine the numerical coefficient of the term, given by the numbers a_{m+1}. However, how each function was missed out is important - if m in a row are missed out then they are replaced with a_{m+1}. The indices of the second sum, p_i, are designed to convey this information. For example, if p_2 is 1 then there has been nothing missed out between the first coth and the second. If, however, it took any other value then $p_2 - 1$ possible coth functions must have been missed out between these two functions. Continuing this logic gives all terms in the above sum.

3.1.6 Resumming the partition formula

We are now going to obtain a simpler form of this function. The brackets in the sum may be expanded, putting the function into the form

$$f_{r-1} = t_r + \sum_{n=1}^{r-1} \sum_{p_1=1}^{\infty} \cdots \sum_{p_{n+1}=1}^{\infty} t_{p_1} t_{p_2} \cdots t_{p_{n+1}}$$
$$\times \coth(x_1 + x_2 + \cdots + x_{p_1}) \cdots \coth(x_1 + x_2 + \cdots + x_{p_1+\cdots+p_n})\, \delta_{r,\, p_1+\cdots+p_{n+1}} \,. \tag{3.39}$$

The coefficient $t_{p_1} t_{p_2} \cdots t_{p_{n+1}}$ is of course still a product of equivalent numbers t_{p_i} as the same partitioning arguments apply. In other words, the numbers p_i still label the size of gaps in the product of coth functions and each provide a number t_{p_i} which depends only upon this size, independent of the location of the gap. Comparing the constant term, that is when all coth functions have been missed out, of (3.38) with that of (3.39) gives

$$t_r = \sum_{n=0}^{r-1} (-1)^n \sum_{p_1=1}^{\infty} \cdots \sum_{p_{n+1}=1}^{\infty} a_{p_1} \cdots a_{p_{n+1}}\, \delta_{r,\, p_1+\cdots+p_{n+1}} \,. \tag{3.40}$$

This sum, once computed for an arbitrary index, will give all numbers t_{p_i} which appear in equation (3.39). The key to computation is to lift the constraint imposed by the Kronecker delta, and as such we may employ generating

functions. First multiply both sides by x^r, and sum over r:

$$\sum_{r=1}^{\infty} t_r x^r = \sum_{r=1}^{\infty} \sum_{n=0}^{r-1} (-1)^n \sum_{p_1=1}^{\infty} \cdots \sum_{p_{n+1}=1}^{\infty} a_{p_1} a_{p_2} \cdots a_{p_{n+1}} \, \delta_{r,\, p_1+\cdots+p_{n+1}} x^r \tag{3.41}$$

$$= \sum_{n=0}^{\infty} (-1)^n \left[\sum_{p_1=1}^{\infty} a_{p_1} x^{p_1} \right] \cdots \left[\sum_{p_{n+1}=1}^{\infty} a_{p_{n+1}} x^{p_{n+1}} \right] . \tag{3.42}$$

Now each of the sums over p_i can be done freely, resulting in

$$\left(1 - \sum_{k=1}^{\infty} t_k x^k \right) \left(1 + \sum_{k=1}^{\infty} a_k x^k \right) = 1 \, . \tag{3.43}$$

The above is an expression of a kind of 'partition duality'. It is true for any sequence $\{a_k\}$ and defines a dual sequence $\{t_k\}$ which satisfies equation (3.40). This also implies that equation (3.40) is invertible, that is one can exchange a_k and $-t_k$ and the equation will still hold. Of course, what has been done here is to replace the $(\coth(x) - 1)$ of equation (3.38) with $(\coth(x) - 0)$ in equation (3.39). One could instead replace it with a more general $(\coth(x) - \lambda)$, with the analysis being analogous to that which has been performed, though currently $\lambda = -1$, 0, 1 are the only useful cases as they have been used in this chapter.

In the present case recall $a_k \equiv 2^{k-1}/k!$ and hence it is simple to calculate that

$$\sum_{k=0}^{\infty} t_k x^k = \tanh(x) = x - \frac{1}{3} x^3 + \frac{2}{15} x^5 - \frac{17}{315} x^7 + \cdots \, , \tag{3.44}$$

demonstrating the numbers t_k are generated by tanh. When combined with equation (3.39), this then gives us a clean formula for f_{r-1} and thus F_N. That is, f_{r-1} is a sum of products of coth functions. In each term of this sum, some even number of these functions in a row will be missed out and replaced with the numbers t_k which come from the Taylor expansion of $\tanh(x)$. We will give finite examples of this concept for clarity in section 3.3.

3.1.7 Revisiting M^m and implementing the fundamental mathematical approach

Our focus will now turn to the exponential in equation (3.36). It is here that we will reverse the identity (3.10). We may now rewrite equation (3.20) as

$$M^m = e^{2mA} + \sum_{N=1}^{\infty} e^{mA} \left(\sum_{r=0}^{N} \tilde{F}_r \tilde{F}_{N-r} e^{m(-L_1 - \cdots - L_r + L_{r+1} + \cdots + L_N)} \right) B_1 \cdots B_N e^{mA} , \tag{3.45}$$

where we have suppressed the arguments of the functions for brevity. Upon repeated application of the Hadamard formula (3.8) we can see this as

$$M^m = e^{2mA} + \sum_{N=1}^{\infty} \sum_{r=0}^{N} \tilde{F}_r \tilde{F}_{N-r} \, B_1 \cdots B_r \, e^{2mA} \, B_{r+1} \cdots B_N \, . \tag{3.46}$$

We may then employ the identity (3.10) in reverse, obtaining

$$\log M = 2A + \sum_{N=1}^{\infty} \sum_{r=0}^{N} \tilde{F}_r \tilde{F}_{N-r}\, B_1 \cdots B_r\, 2A\, B_{r+1} \cdots B_N \,. \tag{3.47}$$

Using the commutator operators L_i, the matrix A in the above expression may be moved to either side of the matrices B, via

$$B_1 \cdots B_r\, A\, B_{r+1} \cdots B_N = (-L_1 - L_2 - \cdots - L_r)\, B_1 \cdots B_N + A\, B_1 \cdots B_N \tag{3.48}$$

$$= (L_r + L_{r+1} + \cdots + L_N)\, B_1 \cdots B_N + B_1 \cdots B_N\, A \,. \tag{3.49}$$

The case $m = 0$ in equation (3.46) gives

$$1 = 1 + \sum_{N=1}^{\infty} \sum_{r=0}^{N} \tilde{F}_r \tilde{F}_{N-r} B_1 \cdots B_N \,, \tag{3.50}$$

which implies that for all $N > 0$,

$$\sum_{r=0}^{N} \tilde{F}_r \tilde{F}_{N-r} = 0 \,. \tag{3.51}$$

This identity is extremely useful and will appear again later in this work and has been checked for up to 10th order. For now it allows the extraneous final terms in equations (3.48) and (3.49) to be dropped and hence $\log M$ to be

written in the form

$$\log M = 2A + \sum_{N=1}^{\infty} \left[\sum_{r=0}^{N} \tilde{F}_r \tilde{F}_{N-r} \left(-L_1 - \cdots - L_r + L_{r+1} + \cdots + L_N \right) \right] B_1 \cdots B_N \,. \tag{3.52}$$

This then gives the promised expansion in powers of the matrix B. To be clear, the string $B_1 B_2 \cdots B_N$ does not imply a matrix product, as the operators L_i are to be acted upon it, interspersing this string with commutators operators and then turning it into a simple matrix product involving both B and A matrices.

3.1.8 Final form

To summarise, we have found that

$$\log \left(e^A e^{2B} e^A \right) = 2A + \sum_{N=1}^{\infty} G_N \, B_1 \cdots B_N \,, \tag{3.53}$$

where

$$G_N = \sum_{r=0}^{N} \tilde{F}_r(-L_r, -L_{r-1}, \ldots, -L_1) \tilde{F}_{N-r}(L_{r+1}, L_{r+2}, \ldots, L_N) \times \left(-L_1 - \cdots - L_r + L_{r+1} + \cdots + L_N \right), \tag{3.54}$$

$$\tilde{F}_r(x_1, \ldots, x_r) = \frac{f_{r-1}(x_1, \ldots, x_{r-1})}{\sinh(x_1 + x_2 + \cdots + x_r)} \qquad \text{for } r \in \mathbb{Z}^+ , \qquad \tilde{F}_0 \equiv 1 , \tag{3.55}$$

and

$$f_{r-1} = t_r + \sum_{n=1}^{r-1} \sum_{p_1=1}^{\infty} \cdots \sum_{p_{n+1}=1}^{\infty} t_{p_1} t_{p_2} \cdots t_{p_{n+1}}$$
$$\times \coth(x_1 + x_2 + \cdots + x_{p_1}) \cdots \coth(x_1 + x_2 + \cdots + x_{p_1+\cdots+p_n}) \, \delta_{r,\, p_1+\cdots+p_{n+1}} . \tag{3.56}$$

Here the numbers t_{p_i} are given from the Taylor expansion of $\tanh(x)$. Finite examples are given in section 3.3.

3.2 Representation as a sum of commutators

It is well known that, beyond the initial terms, the Baker-Campbell-Hausdorff formula may be written as the sum of commutators. Unfortunately, for our new representation (3.53) this is not immediately evident. Of course, the commutator operators L_i contained within G_N will be applied to each matrix B_i to form commutators. However, this would naturally lead to products of commutators when, say, a term like $L_i L_j$ is applied to $B_i B_j$. In this section we will give a representation for which each term is a single commutator. This representation will rely on unproved identities of the function G_N, which have

been demonstrated for up to $N = 10$ and thus may be true.

The first identity involves picking one argument of G_N, say L_1, then changing its position while preserving the order of the other arguments. Explicitly for G_4, the following identity is true:

$$G_4(L_1,\, L_2,\, L_3,\, L_4)+G_4(L_2,\, L_1,\, L_3,\, L_4)+G_4(L_2,\, L_3,\, L_1,\, L_4)+G_4(L_2,\, L_3,\, L_4,\, L_1) = 0\,. \tag{3.57}$$

The next identity involves picking two arguments, say L_1 and L_2. This time the position of both arguments is allowed to change, preserving both their own order and the order of the remaining arguments. Explicitly for G_4,

$$\begin{aligned} &G_4(L_1,\, L_2,\, L_3,\, L_4) + G_4(L_1,\, L_3,\, L_2,\, L_4) + G_4(L_1,\, L_3,\, L_4,\, L_2) \\ &\quad + G_4(L_3,\, L_1,\, L_2,\, L_4) + G_4(L_3,\, L_1,\, L_4,\, L_2) + G_4(L_3,\, L_4,\, L_1,\, L_2) = 0 \end{aligned}$$

In general we conjecture that identities hold where $n < N$ arguments of G_N are picked and are dealt with in an analogous way to above. Again, it should be noted that this has been tested successfully up to $N = 10$ and there is no reason to believe this should fail at any higher order.

From these identities it follows, as we shall shortly prove, that

$$G_N\, B_1 \cdots B_N = \frac{1}{N} G_N\, [[\ldots[[B_1,\, B_2],\, B_3],\, \ldots],\, B_N]\,, \tag{3.58}$$

and hence that

$$\log\left(e^A e^{2B} e^A\right) = 2A + G_1 B_1 + \sum_{N=2}^{\infty} \frac{1}{N} G_N \left[[\ldots[[B_1,\, B_2],\, B_3],\, \ldots],\, B_N\right]. \quad (3.59)$$

To begin the proof, subject to the identities we have just discussed, note that the commutator on the right-hand-side of the above equation may be written in terms of permutations of the string $B_1 \cdots B_N$. That is,

$$\begin{aligned} &\frac{1}{N} G_N \left[[\ldots[[B_1,\, B_2],\, B_3],\, \ldots],\, B_N\right] \\ &= \frac{1}{N} G_N \big\{()_B - (1NN-1\ldots 2)_B\big\} \cdots \big\{()_B - (132)_B\big\}\big\{()_B - (12)_B\big\} B_1 \cdots B_N\,, \end{aligned} \quad (3.60)$$

where $(n_1 n_2 \ldots n_N)_B$ represents a permutation of the string $B_1 \cdots B_N$.

Next we may relabel the indices in each term, keeping the order $B_1 \cdots B_N$ and instead permuting the arguments of the function $G_N(L_1, \ldots, L_N)$. For example,

$$\begin{aligned} (132)_B\, G_4(L_1,\, L_2,\, L_3,\, L_4)\, B_1 B_2 B_3 B_4 &= G_4(L_1,\, L_2,\, L_3,\, L_4)\, B_3 B_1 B_2 B_4 \\ &= G_4(L_2,\, L_3,\, L_1,\, L_4)\, B_1 B_2 B_3 B_4 \\ &= (123)_G\, G_4(L_1,\, L_2,\, L_3,\, L_4)\, B_1 B_2 B_3 B_4\,. \end{aligned}$$

In general, for any permutation P,

$$P_B\, G_N(L_1,\, \ldots,\, L_N)\, B_1 \cdots B_N = P_G^{-1}\, G_N(L_1,\, \ldots,\, L_N)\, B_1 \cdots B_N\,, \qquad (3.61)$$

and so we may rewrite equation (3.60) in terms of permutations on G_N as

$$\begin{aligned}
&\frac{1}{N} G_N\,[[\ldots[[B_1,\, B_2],\, B_3],\, \ldots],\, B_N] \\
&\quad = \frac{1}{N}\big\{()_G - (12)_G\big\}\big\{()_G - (123)_G\big\} \cdots \big\{()_G - (1\ldots N)_G\big\} G_N B_1 \cdots B_N \\
&\quad = \frac{1}{N}\left[()_G + \sum_{m=1}^{N} (-1)^m \sum_{1<n_m<\cdots<n_1\le N} (1\ldots n_m)_G \cdots (1\ldots n_1)_G\right] G_N B_1 \cdots B_N
\end{aligned} \qquad (3.62)$$

We may also write the identities of the function G_N in this permutation style. Most relevantly, choosing the arguments $L_m,\, L_{m-1},\, \ldots,\, L_1$ and changing their position with respect to the remaining arguments $L_{m+1},\, L_{m+2},\, \ldots,\, L_N$ while keeping the two sets internally ordered may be written as

$$\sum_{1\le n_m<\cdots<n_1\le N} (1\ldots n_m)_G \cdots (1\ldots n_1)_G\, G_N(L_1,\, L_2,\, \ldots,\, L_N) = 0\,. \qquad (3.63)$$

As an example, for $m = 2$ and $N = 4$ the above reads

$$\begin{aligned}
&G_4(L_2,\, L_1,\, L_3,\, L_4) + G_4(L_2,\, L_3,\, L_1,\, L_4) + G_4(L_2,\, L_3,\, L_4,\, L_1) \\
&+ G_4(L_3,\, L_2,\, L_1,\, L_4) + G_4(L_3,\, L_2,\, L_4,\, L_1) + G_4(L_3,\, L_4,\, L_2,\, L_1) = 0\,.
\end{aligned}$$

We may split the sum in the identity in to two cases: $n_m = 1$ and $n_m \neq 1$. This gives

$$(-1) \sum_{1<n_m<\cdots<n_1\leq N} (1\ldots n_m)_G \cdots (1\ldots n_1)_G G_N$$
$$= \sum_{1<n_{m-1}<\cdots<n_1\leq N} (1\ldots n_{m-1})_G \cdots (1\ldots n_1)_G G_N\,, \quad (3.64)$$

where we have suppressed the argument to the function G_N for brevity. This leads naturally to recursion. Using lower order identities, that is starting at $m-1$ not m and so on, we can see that

$$(-1)^m \sum_{1<n_m<\cdots<n_1\leq N} (1\ldots n_m)_G \cdots (1\ldots n_1)_G G_N = G_N\,. \quad (3.65)$$

The left hand side of the above is exactly what is obtained when expanding equation (3.62), collecting all terms involving m permutations multiplied together. Exactly N copies of this occur, thus proving equation (3.58).

3.3 Finite examples

While we have derived the general formula in the preceding sections, it may be helpful to examine several low-order terms explicitly. This section will begin with the functions f_r, for $r = 0\ldots 5$, highlighting the patterns we have previously discussed. From these we may immediately write down the operators

G_N, the targets of this work, and indeed we will do so for $N = 1 \ldots 5$.

To begin, consider the functions f_r. The first few of these functions are given by

$$
\begin{aligned}
f_0 &\equiv 1\,, \\
f_1 &= c_1\,, \\
f_2 &= c_1 c_{12} - \frac{1}{3}\,, \\
f_3 &= c_1 c_{12} c_{123} - \frac{1}{3}\left(c_1 + c_{123}\right), \\
f_4 &= c_1 c_{12} c_{123} c_{1234} - \frac{1}{3}\left(c_1 c_{12} + c_1 c_{1234} + c_{123} c_{1234}\right) + \frac{2}{15}\,,
\end{aligned}
$$

where we have used compact notation ($c_{123} = \coth(x_1 + x_2 + x_3)$, for example). Here the structure we previously discussed becomes apparent. In equation (3.56) the term in the sum where $n = r - 1$ forces each p_i to be equal to one, giving the full product of coth functions with none missing. This is the leading term in each of the examples above. To generate the rest of the terms, we replace neighbouring pairs of coth functions in this term with $-1/3$, neighbouring quadruplets with $2/15$, and so on. All possible such replacements appear in the above functions, where the replacing numbers are given

from

$$\tanh x = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 - \frac{17}{315}x^7 + \cdots \tag{3.66}$$

$$= \sum_{n=1}^{\infty} \frac{2^{2n}(2^{2n}-1)B_2 n x^{2n-1}}{(2n)!}\,, \tag{3.67}$$

where B_n is the nth Bernoulli number.

We will now examine the targets of this work, namely the operators G_N. We previously mentioned that the leading term G_1 is already well known and while this was calculated for the regular Baker-Campbell-Hausdorff formula $Z(X, Y)$, it is of course trivial to map it to the symmetric version $\mathcal{S}(A, B)$. Using the general formulae of the preceding section, it would be natural to write

$$G_1 = \left[\frac{1}{\sinh(L_1)}\right] L_1 + \left[\frac{1}{\sinh(-L_1)}\right] (-L_1)\,. \tag{3.68}$$

Of course, as both x and $\sinh(x)$ are odd functions, the minus signs are irrelevant and there is only really one term.

Next, at second order and third order we find that

$$G_2 = \left[\frac{\coth(L_1)}{\sinh(L_1+L_2)}\right](L_1+L_2) + \left[\frac{1}{\sinh(-L_1)}\right]\left[\frac{1}{\sinh(L_2)}\right](-L_1+L_2) + \left[\frac{\coth(-L_2)}{\sinh(-L_1-L_2)}\right](-L_1-L_2)\,,$$

and

$$\begin{aligned} G_3 = &\left[\frac{\coth(L_1)\coth(L_1+L_2)-\frac{1}{3}}{\sinh(L_1+L_2+L_3)}\right](L_1+L_2+L_3) \\ &+\left[\frac{1}{\sinh(-L_1)}\right]\left[\frac{\coth(L_2)}{\sinh(L_2+L_3)}\right](-L_1+L_2+L_3) \\ &+\left[\frac{\coth(-L_2)}{\sinh(-L_1-L_2)}\right]\left[\frac{1}{\sinh(L_3)}\right](-L_1-L_2+L_3) \\ &+\left[\frac{\coth(-L_3)\coth(-L_2-L_3)-\frac{1}{3}}{\sinh(-L_1-L_2-L_3)}\right](-L_1-L_2-L_3)\,. \end{aligned}$$

With these, some general themes begin to emerge. We can immediately see that each term factorises into two parts, written above with square brackets. In a given term, all commutator operators with a plus sign gather into one of these parts while those with a minus sign gather into the other. The only question that remains is how to determine the arguments to each coth function.

Let us consider, for example, the term involving $-L_1-L_2-L_3+L_4+L_5+L_6+L_7$ in G_7. Pictorially, the arguments for each function can be found from the diagram

$$-L_1 - L_2 - L_3 + L_4 + L_5 + L_6 + L_7\,.$$

Here, the top red lines highlight the arguments of each sinh function, while the blue lines show the arguments to the coth functions. Combined with the previous discussion on how to write down these coth functions to form the

numerators, this says how to write G_N for any order N. Of course equation (3.54) already provides such a formula, but perhaps observing these patterns for finite results may provide a more intuitive understanding.

For reference, the next two orders in the expansion are given by

$$\begin{aligned}
G_4 = &\left[\frac{c_1 c_{12} c_{123} - \frac{1}{3}\left(c_1 + c_{123}\right)}{s_{1234}}\right]\left(L_1 + L_2 + L_3 + L_4\right) \\
&+ \left[\frac{1}{s_{\bar{1}}}\right]\left[\frac{c_2 c_{23} - \frac{1}{3}}{s_{234}}\right]\left(-L_1 + L_2 + L_3 + L_4\right) \\
&+ \left[\frac{c_{\bar{2}}}{s_{\overline{12}}}\right]\left[\frac{c_3}{s_{34}}\right]\left(-L_1 - L_2 + L_3 + L_4\right) \\
&+ \left[\frac{c_{\bar{3}} c_{\overline{23}} - \frac{1}{3}}{s_{\overline{123}}}\right]\left[\frac{1}{s_4}\right]\left(-L_1 - L_2 - L_3 + L_4\right) \\
&+ \left[\frac{c_{\bar{4}} c_{\overline{34}} c_{\overline{234}} - \frac{1}{3}\left(c_{\bar{4}} + c_{\overline{234}}\right)}{s_{\overline{1234}}}\right]\left(-L_1 - L_2 - L_3 - L_4\right),
\end{aligned}$$

and

$$\begin{aligned}
G_5 = &\left[\frac{c_1 c_{12} c_{123} c_{1234} - \frac{1}{3}\left(c_1 c_{12} + c_1 c_{1234} + c_{123} c_{1234}\right) + \frac{2}{15}}{s_{12345}}\right]\left(L_1 + L_2 + L_3 + L_4 + L_5\right) \\
&+ \left[\frac{1}{s_{\bar{1}}}\right]\left[\frac{c_2 c_{23} c_{234} - \frac{1}{3}\left(c_2 + c_{234}\right)}{s_{2345}}\right]\left(-L_1 + L_2 + L_3 + L_4 + L_5\right) \\
&+ \left[\frac{c_{\bar{2}}}{s_{\overline{12}}}\right]\left[\frac{c_3 c_{34} - \frac{1}{3}}{s_{345}}\right]\left(-L_1 - L_2 + L_3 + L_4 + L_5\right) \\
&+ \left[\frac{c_{\bar{3}} c_{\overline{23}} - \frac{1}{3}}{s_{\overline{123}}}\right]\left[\frac{c_4}{s_{45}}\right]\left(-L_1 - L_2 - L_3 + L_4 + L_5\right) \\
&+ \left[\frac{c_{\bar{4}} c_{\overline{34}} c_{\overline{234}} - \frac{1}{3}\left(c_{\bar{4}} + c_{\overline{234}}\right)}{s_{\overline{1234}}}\right]\left[\frac{1}{s_5}\right]\left(-L_1 - L_2 - L_3 - L_4 + L_5\right) \\
&+ \left[\frac{c_{\bar{5}} c_{\overline{45}} c_{\overline{345}} c_{\overline{2345}} - \frac{1}{3}\left(c_{\bar{5}} c_{\overline{45}} + c_{\bar{5}} c_{\overline{2345}} + c_{\overline{345}} c_{\overline{2345}}\right) + \frac{2}{15}}{s_{\overline{12345}}}\right]\left(-L_1 - L_2 - L_3 - L_4 - L_5\right).
\end{aligned}$$

Here we have made the notation more compact by writing, for example, $s_1 = \sinh(L_1)$ and $c_{\overline{23}} = \coth(-L_2 - L_3)$.

3.4 Apparent singularities and an alternative representation

One may, upon reading section 3.3 and the examples therein, be concerned that the operators G_N appear divergent. Both $\coth(x)$ and $1/\sinh(x)$ have simple poles when their argument is zero. In this section, however, we will provide the framework for removing these apparent singularities at will. While this can be done using the operators as given in the preceding section, we have found it is better to rewrite and potentially simplify using hyperbolic identities, creating alternative representations. In this section we shall discuss one such alternative and use it as the basis for an algorithmic approach to removing singularities, which is performed in detail in appendix A. Also in appendix A is an exhaustive list of possible singularities in the operators G_1, G_2, G_3, and G_4, and the result of removing them.

The starting point for obtaining this alternative representation is the $m = 0$

identity (3.51),

$$\sum_{r=0}^{N} \tilde{F}_r(-L_r, -L_{r-1}, \ldots, -L_1)\tilde{F}_{N-r}(L_{r+1}, L_{r+2}, \ldots, L_N) = 0\,. \tag{3.69}$$

Recalling the definition

$$\tilde{F}_r(x_1, \ldots, x_r) = \frac{f_{r-1}(x_1, \ldots, x_{r-1})}{\sinh(x_1 + x_2 + \cdots + x_r)}\,, \tag{3.70}$$

and that $\tilde{F}_0 \equiv 0$, we may extract the two outer terms, that is $r = 0$, N, to give

$$\frac{f_{N-1}(L_1, \ldots, L_{N-1}) - f_{N-1}(-L_N, \ldots, -L_2)}{\sinh(L_1 + \cdots + L_N)} = -\sum_{r=1}^{N-1} \frac{f_{r-1}(-L_r, \ldots, -L_2)}{\sinh(-L_1 - \cdots - L_r)} \frac{f_{N-r-1}(L_{r+1}, \ldots, L_{N-1})}{\sinh(L_{r+1} + \cdots + L_N)}\,.$$

The key to this representation is to eliminate all sinh functions. To that end, we may use the hyperbolic identity

$$\frac{1}{\sinh(-L_1 - \cdots - L_r)} \frac{1}{\sinh(L_{r+1} + \cdots + L_N)} = \frac{\coth(-L_1 - \cdots - L_r) - \coth(L_{r+1} + \cdots + L_N)}{\sinh(L_1 + \cdots + L_N)} \tag{3.71}$$

to rewrite the right-hand-side of the above equation, and thus eliminate the sinh function. At this point there is a clear divide, with half of the terms containing the variable L_1 but not L_N and the other half containing L_N but

not L_1. We may reorganise the equation to separate each half by the equals sign which, along with linear independence of the functions involving L_1 and L_N, implies that each half separately must be equal to some constant. That is, for the L_1 dependent half,

$$f_{N-1}(L_1, \ldots, L_{N-1}) + \sum_{r=1}^{N-1} \coth(-L_1 - \cdots - L_r) f_{r-1}(-L_r, \ldots, -L_2) f_{N-r-1}(L_{r+1}, \ldots, L_{N-1}) = \text{const} \equiv a_N^{\text{odd}} . \quad (3.72)$$

One of the striking features of this representation is the factorisation structure which has been ubiquitous in this work, that is the split into f_{r-1} and f_{n-r-1}. Its presence here gives reassurance that this formula is natural. Secondly, outside of the f_{N-1}, all dependence on the variable L_1 appears only in the outer coth terms. We may think of these terms as linearly independent basis functions, with the $f_{r-1}f_{N-r-1}$ terms cast as coefficients. This then gives a more controlled way of dealing with these formulae. We will use this equation to rewrite the overall operator G_N, but first we must find the constant.

We may find this constant term by taking the limit $L_1 \to \pm\infty$ in f_{N-1}, which has the effect of setting each coth to one or minus one. Adapting an equation for f_{N-1}, namely equation (3.38), we find

$$\lim_{L_1 \to \pm\infty} f_{N-1} = (\pm 1)^N a_N \equiv (\pm 1)^N \frac{2^{N-1}}{N!} , \quad (3.73)$$

and so taking the same limits on equation (3.72) gives

$$(\pm 1)^{N-1} a_N \pm \sum_{r=1}^{N-1} f_{r-1} f_{N-r-1} = a_N^{\text{odd}} \,. \tag{3.74}$$

We may then sum the two equations contained in (3.74) to find the constant

$$a_N^{\text{odd}} = \begin{cases} 0, & N \text{ even}, \\ a_N, & N \text{ odd}, \end{cases} \tag{3.75}$$

with generating function

$$\sum_{N=1}^{\infty} a_N^{\text{odd}} x^N = \cosh(x) \sinh(x) = x + \frac{2}{3} x^3 + \frac{2}{15} x^5 + \cdots \,. \tag{3.76}$$

Similarly, we may subtract the equations (3.74), giving a set of identities which will prove useful when dealing with apparent singularities,

$$\sum_{r=1}^{N-1} f_{r-1} f_{N-r-1} = \begin{cases} a_N, & N \text{ even}, \\ 0, & N \text{ odd}, \end{cases} \equiv a_N^{\text{even}} \tag{3.77}$$

where

$$\sum_{n} a_n^{\text{even}} x^n = \sinh^2(x) = x^2 + \frac{1}{3} x^4 + \frac{2}{45} x^6 + \cdots \,. \tag{3.78}$$

Returning to the alternative representation, we may now rewrite the overall operators G_N. Using the hyperbolic identity (3.71) to combine all sinh func-

tions and the recursion relation (3.72) to eliminate both instances of f_{N-1}, we can see that

$$G_N = 2s(L_1 + \cdots + L_N)\, g_N \tag{3.79}$$

where

$$s(x) = \frac{x}{\sinh(x)}\,, \tag{3.80}$$

$$g_N = a_N^{\text{odd}} + \sum_{r=1}^{N-1} E(L_1 + \cdots + L_r,\, L_{r+1} + \cdots + L_N) \times f_{r-1}(-L_r,\, \ldots,\, -L_2) f_{N-r-1}(L_{r+1},\, \ldots,\, L_{N-1})\,, \tag{3.81}$$

and

$$E(x, y) = \frac{x\coth(x) - y\coth(y)}{x + y}\,. \tag{3.82}$$

It is fairly clear that both $s(x)$ and $E(x, y)$ are regular and infinitely differentiable and as such any apparent singularity involving L_1 or L_N is automatically safe in this representation.

Note that this representation only deals with the case where one of the arguments is zero. Demonstrating that any other limits are safe involves the identities just introduced in equation (3.77). This is done carefully for a particular example in appendix A, but is also done exhaustively in that same appendix for G_1, G_2, G_3, and G_4. We should note that in applications it is

usual, rather than unusual, that such singularities are relevant. As such, the representation presented in this section should be considered as the starting point for practical use of the new formula. Indeed, we shall treat it as such in future chapters.

3.5 Choice of basis

In this section we shall discuss the sums of commutator operators, that is strings like $L_1 + L_2 + \cdots + L_r$. We previously suggested that these were mathematically natural to use as arguments to various functions. It turns out that in the basis where the matrix A is diagonal, if such a basis exists, these sums result in the difference between two eigenvalues of A. As we shall see, this drastically reduces the complexity of using the new representation and is really the true use of this formula in this thesis.

First let us consider the matrix elements of $LB \equiv [A,\, B]$:

$$\big[LB\big]_{n_1 n_2} = A_{n_1 n'} B_{n' n_2} - B_{n_1 n'} A_{n' n_2}\,, \tag{3.83}$$

where we assume summation over repeated indices. If A is a diagonal matrix then its matrix elements are given in terms of its eigenvalues as $A_{nm} = a_n \delta_{nm}$, where δ_{nm} is the Kronecker delta. Hence in the basis A is diagonal the above

is given by

$$\big[LB\big]_{n_1 n_2} = (a_{n_1} - a_{n_2}) B_{n_1 n_2} \,. \tag{3.84}$$

More generally, for any Taylor expandable function f, it can be seen that

$$\Big[f(L_i{+}L_{i+1}{+}\cdots{+}L_{i+r})B_1 B_2 \cdots B_N\Big]_{n_1 n_{N+1}} = f(a_{n_i}{-}a_{n_{i+r+1}})B_{n_1 n_2} B_{n_2 n_3} \cdots B_{n_N n_{N+1}} \,. \tag{3.85}$$

This is a simple yet powerful result. If the function f is replaced with sinh or coth functions, then we may determine G_N without difficulty. This would allow calculations to be done numerically with relative ease as all the strings of commutator operators are replaced by real numbers. It is this choice of basis, then, which gives the results of this chapter a practical raison d'être.

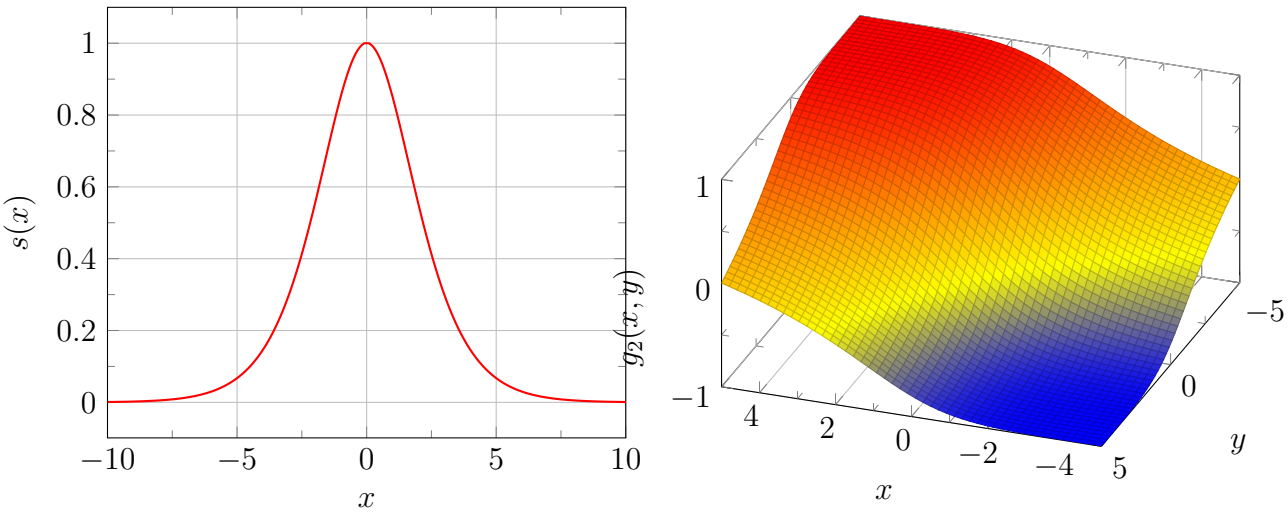


Figure 3.2: (a) The Boltzmann suppression factor as a function of the difference of eigenvalues. (b) The function g_2. Note it appears to limit to ± 1 as the eigenvalue difference tends to infinity, and is bounded.

We ought to say a few words about the full expansion in this basis. Using

the notation of section 3.4, note that we may extract an overall factor of $s(L_1 + L_2 + \cdots + L_N)$ as, at each order, the argument of this function is the same difference of eigenvalues. That is,

$$\begin{aligned}[\mathcal{S}(A,\, B)]_{nn'} &= \left[2A + \sum_{N=1}^{\infty} G_N B_1 \cdots B_N\right]_{nn'} \\ &= 2a_n\delta_{n,\,n'} + 2s(a_n - a_{n'})\Big(B_{nn'} + g_2(a_n - a_{n_1}, a_{n_1} - a_{n'})\, B_{nn_1}B_{n_1n'} \\ &\qquad + g_3(a_n - a_{n_1}, a_{n_1} - a_{n_2}, a_{n_2} - a_{n'})\, B_{nn_1}B_{n_1n_2}B_{n_2n'} + \cdots\Big)\,,\end{aligned} \tag{3.86}$$

where again we assume summation over repeated indices. We may interpret the function s as a Boltzmann suppression factor, appropriately named when one considers potential applications in quantum and statistical mechanics, which reduces the weight of any matrix element for whom the difference in eigenvalues $a_n - a_{n'}$ is sufficiently large. It is plotted in figure 3.2a, for reference.

The reduced functions g_N, then, are what remains. Figure 3.2b displays g_2, which has several features generic to these functions. Firstly, it appears to be bounded and its extrema occur as its arguments diverge. It can be proved, though we shall not do so here, that under such limits the relatively complicated g_N and the comparatively simple f_{N-1} coincide. Then finding these extrema is elementary as each constituent coth function within f_{N-1} takes values ± 1. That these limits do in fact correspond to the extrema of g_N is

not proved, but we have numerically verified this up to g_8 and the results are displayed in table 3.1. Furthermore, we may obtain a generating function for the outermost of these bounds which is given by

$$b(x) = \frac{1 - e^{-2x}}{1 + e^{-4x}} = x + x^2 - \frac{4}{3}x^3 - \frac{5}{3}x^4 + \frac{122}{45}x^5 + \cdots \tag{3.87}$$

This series is absolutely convergent when $|x| < \pi/4$, though this is not proved here. It is entirely possible for the series to converge outside of this region, however. This provides reassurance that the series (3.86) converges for sufficiently small B.

Function	Lower bound	Upper bound
g_2	-1	1
g_3	$-4/3$	$2/3$
g_4	$-5/3$	$5/3$
g_5	$-6/5$	$32/15$
g_6	$-122/45$	$122/45$
g_7	$-1088/315$	$676/315$
g_8	$-227/63$	$227/63$

Table 3.1: Bounds of the functions g_N, obtained using the procedure outlined in the text and numerically verified

3.6 Summary

We have found and proved a new representation for the Baker-Campbell-Hausdorff formula. This representation is a perturbative expansion in just one of two matrices, as opposed to both in the original representation, though

its radius of convergence has not been calculated. The series may then be truncated and give a good approximation to the full expansion for situations where only this second object is small, as we shall describe in more detail in subsequent chapters. For physical problems this then would give access to a much larger parameter space than is currently available. Additionally, new problems for which the original representation was unusable may now be tackled. Transfer matrices in statistical mechanics is an example of one such problem, which is indeed what we shall spend the rest of this work dealing with.

CHAPTER 4

PERTURBATION THEORY

The previous chapter left us with a mathematical result which we may use to find the free-energy operator $\hat{F}$ for a given model. However, in statistical mechanics one is rarely interested in this operator. Instead we care more about its eigenvalues, in particular the groundstate which is the free-energy. Chapter 2 also argued that we ought also to care about the first excited state. Perhaps then we should concentrate on calculating these numbers, rather than the operator itself. Fortunately there is a technique all physicists and mathematicians are familiar with which does just that: perturbation theory.

In this section we shall introduce the particular form of perturbation theory we shall need to use, then apply it to our new Baker-Cambell-Hausdorff formula. We will be able to calculate the subsequent formula explicitly up to sixth

order and will offer some suggestion as to what we expect later results to be. However, we do not currently have a full, rigorous answer.

4.1 General perturbation result

We start with the eigenvalue equation for some matrix H acting on its ground-state eigenvector

$$(H - \varepsilon_0)\,|\psi\rangle = 0\,. \tag{4.1}$$

We shall introduce the perturbation theory by setting

$$H = H^{(0)} + H^{(1)} + H^{(2)} + \cdots\,, \tag{4.2}$$

$$\varepsilon_0 = \varepsilon_0^{(0)} + \varepsilon_0^{(1)} + \varepsilon_0^{(2)} + \cdots\,, \tag{4.3}$$

$$|\psi\rangle = |0\rangle + \sum_{n\neq 0}(\varphi_n^{(1)} + \varphi_n^{(2)} + \cdots)\,|n\rangle\,, \tag{4.4}$$

where the set of vectors $|0\rangle\,,\ |1\rangle\,,\ |2\rangle\,,\ \ldots$ are an orthonormal basis for the matrix $H^{(0)}$. One should think of quantities $x^{(n)}$ as being perturbative, where x can be any of H, ε_0, or φ_n. That is, $x^{(1)} \ll x^{(2)} \ll x^{(3)} \ll \cdots$, but $x^{(n_1)}x^{(n_2)}\cdots x^{(n_m)} \sim x^{(n_1+n_2+\cdots+n_m)}$. Alternatively, think of $x^{(n)} = \lambda^n \tilde{x}^{(n)}$ where λ is the small formal expansion parameter and $\tilde{x}^{(n)}$ is of order one. This λ is consistent across variables, meaning $H^{(N)} \sim \varepsilon_0^{(N)} \sim \varphi_n^{(N)}$. Whichever way you formulate this, the task is to write down a formula at each order in

the perturbation theory and find the ground-state eigenvalue $\varepsilon_0^{(N)}$. This will involve knowing the coefficients $\varphi_n^{(1)}$, $\varphi_n^{(2)}$, ..., $\varphi_n^{(N-1)}$ which we shall shortly calculate. Note we are assuming that there are no degeneracies.

Momentarily going back to the Baker-Campbell-Hausdorff expansion, we will think of the perturbative matrices as having matrix elements

$$\langle n| H^{(0)} |n\rangle = 2A_{nn} \,, \tag{4.5}$$

$$\langle n| H^{(1)} |m\rangle = G_1(a_n - a_m)B_{nm} \,, \tag{4.6}$$

$$\langle n| H^{(2)} |m\rangle = \sum_{n_1} G_2(a_n - a_{n_1},\, a_{n_1} - a_m)B_{nn_1}B_{n_1m} \,, \tag{4.7}$$

$$\langle n| H^{(3)} |m\rangle = \sum_{n_1n_2} G_3(a_n - a_{n_1},\, a_{n_1} - a_{n_2},\, a_{n_2} - a_m)B_{nn_1}B_{n_1n_2}B_{n_2m} \,, \tag{4.8}$$

and so on. We are of course working in the basis for which A is diagonal.

For now, however, we shall deal with this perturbation theory in the abstract. First, the zeroth order equation simply states

$$(H^{(0)} - \varepsilon_0^{(0)}) |0\rangle = 0 \qquad \Longrightarrow \qquad \varepsilon_0^{(0)} = \langle 0| H^{(0)} |0\rangle \,. \tag{4.9}$$

Recall $H^{(0)}$ is diagonal and

$$\varepsilon_n^{(0)} \equiv \langle n| H^{(0)} |n\rangle \,.$$

Next, replacing variables in equation (4.1) with their perturbative equivalents, at N-th order we have the equation

$$(H^{(N)} - \varepsilon_0^{(N)})\,|0\rangle + \sum_{r=1}^{N}\sum_{n\neq 0}(H^{(N-r)} - \varepsilon_0^{(N-r)})\varphi_n^{(r)}\,|n\rangle = 0\,. \tag{4.10}$$

As we wish to find the groundstate eigenvalue ε_0 we should first target $\varepsilon_0^{(N)}$. To do so we may multiply the above equation from the left by the eigenvector $\langle 0|$, obtaining

$$\varepsilon_0^{(N)} = \langle 0|\,H^{(N)}\,|0\rangle + \sum_{r=1}^{N-1}\sum_{n\neq 0}\langle 0|\,H^{(N-r)}\,|n\rangle\,\varphi_n^{(r)}\,. \tag{4.11}$$

As this equation depends on the variables $\varphi_n^{(r)}$ we ought to also target $\varphi_N^{(n)}$ in our formula at N-th order. This can be done by multiplying said formula from the left by the vector $\langle n|$, which gives

$$\langle n|\,H^{(N)}\,|0\rangle + \sum_{r=1}^{N-1}\left(\sum_{m\neq 0}\langle n|\,H^{(N-r)}\,|m\rangle\,\varphi_m^{(r)} - \varepsilon_0^{(N-r)}\varphi_m^{(r)}\right) + \sum_{m\neq 0}\langle n|\,H^{(0)}\,|m\rangle\,\varphi_m^{(N)} - \varepsilon_0^{(0)}\varphi_n^{(N)} = 0\,, \tag{4.12}$$

where we have separated the case $r = N$ and also changed the dummy variable. Noting that $\langle n|\,H^{(0)}\,|m\rangle = \delta_{nm}\varepsilon_n^{(0)}$, we may write

$$\varphi_n^{(N)} = \frac{1}{\varepsilon_0^{(0)} - \varepsilon_n^{(0)}}\left[\langle n|\,H^{(N)}\,|0\rangle + \sum_{r=1}^{N-1}\left(\sum_{m\neq 0}\langle n|\,H^{(N-r)}\,|m\rangle\,\varphi_m^{(r)} - \varepsilon_0^{(N-r)}\varphi_n^{(r)}\right)\right]. \tag{4.13}$$

We may substitute our formula for $\varepsilon_N^{(0)}$ into the above to obtain a recursion relation for the variables $\varphi_n^{(N)}$, though writing this out now is perhaps not necessary.

To summarise, we have a recursion relation for the variables $\varphi_n^{(M)}$, from which we can find said variables in terms of the matrix elements of the perturbative matrices. We can then find $\varepsilon_0^{(N)}$ using these variables, with $M < N$, which gives us a perturbative formula for the groundstate ε_0 of our original matrix H.

We shall now give a few examples, finding the perturbative groundstate up to third order. To do so we need to calculate both $\varphi_n^{(1)}$ and $\varphi_n^{(2)}$ using the recursion relation, finding

$$\varphi_n^{(1)} = \frac{1}{\varepsilon_0^{(0)} - \varepsilon_n^{(0)}} \langle n| H^{(1)} |0\rangle \ , \tag{4.14}$$

and the slightly more complicated

$$\varphi_n^{(2)} = \frac{1}{\varepsilon_0^{(0)} - \varepsilon_n^{(0)}} \langle n| H^{(2)} |0\rangle + \sum_{m\neq 0} \frac{1}{(\varepsilon_0^{(0)} - \varepsilon_n^{(0)})(\varepsilon_0^{(0)} - \varepsilon_m^{(0)})} \langle n| H^{(1)} |m\rangle \langle m| H^{(1)} |0\rangle - \frac{1}{(\varepsilon_0^{(0)} - \varepsilon_n^{(0)})^2} \langle 0| H^{(1)} |n\rangle \langle n| H^{(1)} |0\rangle \ . \tag{4.15}$$

These can then be used to find the variables in which we are interested. We

already know

$$\varepsilon_0^{(0)} = \langle 0| H^{(0)} |0\rangle \ , \tag{4.16}$$

and the first order result is similarly trivial:

$$\varepsilon_1^{(0)} = \langle 0| H^{(1)} |0\rangle \ . \tag{4.17}$$

At second order, however, we must now use our formula for $\varphi_n^{(1)}$ to find

$$\varepsilon_0^{(2)} = \langle 0| H^{(2)} |0\rangle + \sum_{n\neq 0} \frac{1}{\varepsilon_0^{(0)} - \varepsilon_n^{(0)}} \langle 0| H^{(1)} |n\rangle \langle n| H^{(1)} |0\rangle \ . \tag{4.18}$$

Third order is even more complicated as we must use both $\varphi_n^{(1)}$ and $\varphi_n^{(2)}$ to find

$$\begin{aligned}
\varepsilon_0^{(3)} = {} & \langle 0| H^{(3)} |0\rangle \\
& + \sum_{n\neq 0} \frac{1}{\varepsilon_0^{(0)} - \varepsilon_n^{(0)}} \langle 0| H^{(2)} |n\rangle \langle n| H^{(1)} |0\rangle \\
& + \sum_{n\neq 0} \frac{1}{\varepsilon_0^{(0)} - \varepsilon_n^{(0)}} \langle 0| H^{(1)} |n\rangle \langle n| H^{(2)} |0\rangle \\
& \quad + \sum_{n,\,m\neq 0} \frac{1}{(\varepsilon_0^{(0)} - \varepsilon_n^{(0)})(\varepsilon_0^{(0)} - \varepsilon_m^{(0)})} \langle 0| H^{(1)} |n\rangle \langle n| H^{(1)} |m\rangle \langle m| H^{(1)} |0\rangle \\
& \quad - \sum_{n\neq 0} \frac{1}{(\varepsilon_0^{(0)} - \varepsilon_n^{(0)})^2} \langle 0| H^{(1)} |n\rangle \langle n| H^{(1)} |0\rangle \langle 0| H^{(1)} |0\rangle \ .
\end{aligned} \tag{4.19}$$

At this point patterns become clear, with essentially the order of the matrix $H^{(3)}$ being split in all possible ways, down to $H^{(2)}H^{(1)}$, $H^{(1)}H^{(2)}$, and

$H^{(1)}H^{(1)}H^{(1)}$. Post-split these are then multiplied by the inverse of the difference between the groundstate eigenvalue and some exited eigenvalue, both of the matrix $H^{(0)}$. The way this is written is also fairly clear. One can use these patterns to write down higher order perturbative results, then check against a machine generated result coming from the recursion relations. It would be prohibitively complex to proceed to much higher order in the way we have presented the first few orders.

4.2 Applying perturbation theory to the Baker-Campbell-Hausdorff series

We will now consider how to apply this general perturbation result to the Baker-Campbell-Hausdorff series calculated in the preceding chapter. As mentioned earlier, we shall set

$$\langle n|\, H^{(0)}\, |n\rangle = 2a_n\,, \tag{4.20}$$

$$\langle n|\, H^{(1)}\, |m\rangle = G_1(a_n - a_m)B_{nm}\,, \tag{4.21}$$

$$\langle n|\, H^{(2)}\, |m\rangle = \sum_{n_1} G_2(a_n - a_{n_1},\, a_{n_1} - a_m)B_{nn_1}B_{n_1m}\,, \tag{4.22}$$

$$\langle n|\, H^{(3)}\, |m\rangle = \sum_{n_1n_2} G_3(a_n - a_{n_1},\, a_{n_1} - a_{n_2},\, a_{n_2} - a_m)B_{nn_1}B_{n_1n_2}B_{n_2m}\,, \tag{4.23}$$

and so on. We are of course working in the basis for which A is diagonal, setting $A_{nn} \equiv a_n$.

We shall begin by calculating the examples discussed in the previous section, namely $\varepsilon_0^{(0)}$, $\varepsilon_1^{(0)}$, $\varepsilon_2^{(0)}$, and $\varepsilon_3^{(0)}$. Both zeroth and first order are trivial, giving

$$\varepsilon_0^{(0)} = 2a_0\,, \tag{4.24}$$

and

$$\varepsilon_0^{(1)} = G_1(0)B_{00} = 2B_{00}, \tag{4.25}$$

respectively. At second order we have a non-trivial, yet still simple, result,

$$\varepsilon_0^{(2)} = \sum_{n_1} G_2(a_0{-}a_{n_1},\, a_{n_1}{-}a_0)\, B_{0n_1}B_{n_10} + \sum_{n_1\neq 0} \frac{1}{2(a_0 - a_{n_1})} G_1(a_0{-}a_{n_1})B_{0n_1}G_1(a_{n_1}{-}a_0)B_{n_10}\,. \tag{4.26}$$

The first sum may be split into two cases, $n_1 \neq 0$ and $n_1 = 0$. The former gives

$$\sum_{n_1\neq 0} \left[G_2(a_0 - a_{n_1},\, a_{n_1} - a_0) + \frac{1}{2(a_0 - a_{n_1})} G_1(a_0 - a_{n_1})G_1(a_{n_1} - a_0) \right] B_{0n_1}B_{n_10}\,, \tag{4.27}$$

which can be simplified by noting

$$G_2(a_0 - a_{n_1},\, a_{n_1} - a_0) + \frac{1}{a_0 - a_{n_1}} G_1(a_0 - a_{n_1})G_1(a_{n_1} - a_0) = 2f_1(a_{n_1} - a_0)\,. \tag{4.28}$$

We shall discuss this simplification in more detail and generality shortly, so it is not essential to verify it here. Instead we shall move to the latter of the two cases, namely $n_1 = 0$, which gives

$$G_2(0,\,0)B_{00}B_{00}\,, \tag{4.29}$$

which equals zero as $G_2(0,\,0) = 0$. In general $G_N(0,\,0\,,\ldots,\,0) = 0$ for $N > 1$ due to identities discussed in the previous chapter. This leaves the result

$$\varepsilon_0^{(2)} = 2\sum_{n_1} B_{0n_1}B_{n_10}\,f_1(a_0 - a_{n_1})\,. \tag{4.30}$$

Finally at third order we have

$$\begin{aligned}
\varepsilon_0^{(3)} = &\sum_{n_1,\,n_2} G_3(a_0 - a_{n_1}, a_{n_1} - a_{n_2}, a_{n_2} - a_0)B_{0n_1}B_{n_1n_2}B_{n_20} \\
&+ \sum_{n_1,\,n_2\neq 0} \frac{1}{2(a_0 - a_{n_2})} G_2(a_0 - a_{n_1},\, a_{n_1} - a_{n_2})B_{0n_1}B_{n_1n_2}G_1(a_{n_2} - a_0)B_{n_20} \\
&+ \sum_{n_1\neq 0,\,n_2} \frac{1}{2(a_0 - a_{n_1})} G_1(a_0 - a_{n_1})B_{0n_1}G_2(a_{n_1} - a_{n_2},\, a_{n_2} - a_0)B_{n_1n_2}B_{n_20} \\
&+ \sum_{n_1\neq 0,\,n_2\neq 0} \frac{1}{4(a_0 - a_{n_1})(a_0 - a_{n_2})} G_1(a_0 - a_{n_1})B_{0n_1}G_1(a_{n_1} - a_{n_2})B_{n_1n_2}G_1(a_{n_2} - a_0)B_{n_20} \\
&- \sum_{n_1\neq 0} \frac{1}{4(a_0 - a_{n_1})^2} G_1(a_0 - a_{n_1})B_{0n_1}G_1(a_{n_1} - a_0)B_{n_10}G_1(0)B_{00}\,.
\end{aligned} \tag{4.31}$$

Again, we need to think about when the n_1 or n_2 dummy variables may be

zero or are disallowed from being so. If $n_1 \neq 0$ and $n_2 \neq 0$ then we have

$$\sum_{n_1\neq 0\,,n_2\neq 0}\Bigg[G_3(a_0-a_{n_1},a_{n_1}-a_{n_2},a_{n_2}-a_0) + \frac{1}{2(a_0-a_{n_2})}G_2(a_0-a_{n_1},\,a_{n_1}-a_{n_2})G_1(a_{n_2}-a_0) + \frac{1}{2(a_0-a_{n_1})}G_1(a_0-a_{n_1})G_2(a_{n_1}-a_{n_2},\,a_{n_2}-a_0) + \frac{1}{4(a_0-a_{n_1})(a_0-a_{n_2})}G_1(a_0-a_{n_1})G_1(a_{n_1}-a_{n_2})G_1(a_{n_2}-a_0)\Bigg]B_{0n_1}B_{n_1n_2}B_{n_20}\,. \tag{4.32}$$

If one is zero and the other not then we have

$$\sum_{n_1\neq 0}\Bigg[G_3(a_0-a_{n_1},a_{n_1}-a_0,0)+\frac{1}{2(a_0-a_{n_1})}G_1(a_0-a_{n_1})G_2(a_{n_1}-a_0,\,0) + G_3(0,a_0-a_{n_1},a_{n_1}-a_0)+\frac{1}{2(a_0-a_{n_1})}G_2(0,\,a_0-a_{n_1})G_1(a_{n_1}-a_0) - \frac{1}{4(a_0-a_{n_1})^2}G_1(a_0-a_{n_1})G_1(a_{n_1}-a_0)G_1(0)\Bigg]B_{00}B_{0n_1}B_{n_10}\,. \tag{4.33}$$

Both n_1 and n_2 equal to zero of course leads to zero, as before. These expressions can both be simplified, leading to the result

$$\varepsilon_0^{(3)} = 2\sum_{n_1\neq 0\,,n_2\neq 0} B_{0n_1}B_{n_1n_2}B_{n_20}\,f_2(a_0-a_{n_1},\,a_{n_1}-a_{n_2}) - \sum_{n_1\neq 0} B_{00}B_{0n_1}B_{n_10}\,f_2(a_0-a_{n_1},\,0)\,. \tag{4.34}$$

Without calculation we shall give the next two results, $\varepsilon_0^{(4)}$ and $\varepsilon_0^{(5)}$, though

these will be largely meaningless until the following section where we shall explain how to write these down using simple patterns. We ought to note that we have also calculated $\varepsilon_0^{(6)}$ but it would be unreasonable to present that here. Instead, an inordinate amount of space shall be dedicated to

$$\begin{aligned}\varepsilon_0^{(4)} = 2 \sum_{n_1\neq 0\,,n_2\neq 0\,,n_3\neq 0} & B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_30}\, f_3(a_0-a_{n_1},\, a_{n_1}-a_{n_2},\, a_{n_2}-a_{n_3}) \\ -2 \sum_{n_1\neq 0\,,n_2\neq 0} & B_{0n_1}B_{n_10}B_{0n_2}B_{n_20}\Big[f_3(a_0-a_{n_1},\, a_0-a_{n_2},\, a_{n_1}-a_0) \\ & \qquad + f_3(a_0-a_{n_1},\, a_0-a_{n_2},\, a_{n_2}-a_0)\Big] \\ - \sum_{n_1\neq 0\,,n_2\neq 0} & B_{00}B_{0n_1}B_{n_1n_2}B_{n_20}\Big[f_3(a_0-a_{n_1},\, 0,\, a_{n_1}-a_{n_2}) \\ & \qquad + f_3(a_0-a_{n_1},\, a_{n_1}-a_{n_2},\, 0)\Big] \\ + \sum_{n_1\neq 0} & B_{00}B_{00}B_{0n_1}B_{n_10}\, f_3(a_0-a_{n_1},\, 0,\, 0)\,, \end{aligned} \tag{4.35}$$

and

$$
\begin{aligned}
\varepsilon_0^{(5)} = 2 \sum_{n_1\neq 0\,,n_2\neq 0\,,n_3\neq 0\,,n_4\neq 0} & B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_3n_4}B_{n_40}\, f_4(a_0-a_{n_1},\, a_{n_1}-a_{n_2},\, a_{n_2}-a_{n_3},\, a_{n_3}-a_{n_4}) \\
-2 \sum_{n_1\neq 0\,,n_2\neq 0\,,n_3\neq 0} & B_{0n_1}B_{n_10}B_{0n_2}B_{n_2n_3}B_{n_30} \Big[f_4(a_0-a_{n_1},\, a_0-a_{n_2},\, a_{n_1}-a_0,\, a_{n_2}-a_{n_3}) \\
& + f_4(a_0-a_{n_1},\, a_0-a_{n_2},\, a_{n_2}-a_{n_3},\, a_0-a_{n_1}) \\
& + f_4(a_0-a_{n_1},\, a_0-a_{n_2},\, a_{n_2}-a_{n_3},\, a_0-a_{n_3}) \Big] \\
- \sum_{n_1\neq 0\,,n_2\neq 0\,,n_3\neq 0} & B_{00}B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_30} \Big[f_4(a_0-a_{n_1},\, 0,\, a_{n_1}-a_{n_2},\, a_{n_2}-a_{n_3}) \\
& + f_4(a_0-a_{n_1},\, a_{n_1}-a_{n_2},\, 0,\, a_{n_2}-a_{n_3}) \\
& + f_4(a_0-a_{n_1},\, a_{n_1}-a_{n_2},\, a_{n_2}-a_{n_3},\, 0) \Big] \\
+ \sum_{n_1\neq 0\,,n_2\neq 0} & B_{00}B_{0n_1}B_{n_10}B_{0n_2}B_{n_20} \Big[f_4(a_0-a_{n_1},\, 0,\, a_0-a_{n_2},\, a_{n_1}-a_0) \\
& + f_4(a_0-a_{n_1},\, 0,\, a_0-a_{n_2},\, a_{n_2}-a_0) \\
& + 2f_4(a_0-a_{n_1},\, a_0-a_{n_2},\, 0,\, a_{n_1}-a_0) \\
& + 2f_4(a_0-a_{n_1},\, a_0-a_{n_2},\, 0,\, a_{n_2}-a_0) \\
& + f_4(a_0-a_{n_1},\, a_0-a_{n_2},\, a_{n_1}-a_0,\, 0) \\
& + f_4(a_0-a_{n_1},\, a_0-a_{n_2},\, a_{n_2}-a_0,\, 0) \Big] \\
+ \sum_{n_1\neq 0\,,n_2\neq 0} & B_{00}B_{00}B_{0n_1}B_{n_1n_2}B_{n_20} \Big[f_4(a_0-a_{n_1},\, 0,\, 0,\, a_{n_1}-a_{n_2}) \\
& + f_4(a_0-a_{n_1},\, 0,\, a_{n_1}-a_{n_2},\, 0) \\
& + f_4(a_0-a_{n_1},\, a_{n_1}-a_{n_2},\, 0,\, 0) \Big] \\
- \sum_{n_1\neq 0} & B_{00}B_{00}B_{00}B_{0n_1}B_{n_10}\, f_4(a_0-a_{n_1},\, 0,\, 0,\, 0)\,. \qquad (4.36)
\end{aligned}
$$

4.3 Understanding the perturbation result

We left the previous section having presented a collection of long and mysterious formulae. This section will attempt to remove at least part of the mystery, though unfortunately not much can be done about the length. It will be necessarily vague, pattern-based, and non-rigorous. This is by design; this formula is not completely under control and to pretend otherwise would be dishonest. The section after this will deal with what can and cannot be proven, but we will give some indication throughout as to what ground is solid and what is more unstable.

Firstly, let us begin by understanding the pattern for the B matrices. To begin writing down $\varepsilon_0^{(N)}$ we first start with the zero-zero-th matrix element of the product of N of the B matrices. That is, for example at fifth order,

$$\sum_{n_1,n_2,n_3,n_4} B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_3n_4}B_{n_40}\,. \tag{4.37}$$

If none of these indices are allowed to be zero then we have the first term in equation (4.36)

$$\sum_{n_1\neq 0,n_2\neq 0,n_3\neq 0,n_4\neq 0} B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_3n_4}B_{n_40}\,. \tag{4.38}$$

If instead either n_2 or n_3 equals zero then, after relabelling the dummy indices,

we have the second term

$$\sum_{n_1\neq 0\,,n_2\neq 0\,,n_3\neq 0} B_{0n_1}B_{n_10}B_{0n_2}B_{n_2n_3}B_{n_30}\,, \tag{4.39}$$

while if instead we allowed either n_1 or n_4 to be zero we would have the third term

$$\sum_{n_1\neq 0\,,n_2\neq 0\,,n_3\neq 0} B_{00}B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_30}\,. \tag{4.40}$$

This logic then continues, giving the remaining terms in equation equation (4.36).

Next we need to think about the weights which are multiplying each of these paths through products. That is, for $\varepsilon_0^{(5)}$, the functions f_4. Of course, in general for $\varepsilon_0^{(N)}$ we would expect these functions to be f_{N-1}. Let us think again about the first term in equation (4.36). It has matrix elements $B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_3n_4}B_{n_40}$ with weight $f_4(a_0-a_{n_1},\,a_{n_1}-a_{n_2},\,a_{n_2}-a_{n_3},\,a_{n_3}-a_{n_4})$. This is perhaps best visualised as

$$\begin{array}{ccccc} B_{0n_1} & B_{n_1n_2} & B_{n_2n_3} & B_{n_3n_4} & B_{n_40} \\ \downarrow & \downarrow & \downarrow & \downarrow & \end{array}$$
$$f_4(a_0-a_{n_1},\,a_{n_1}-a_{n_2},\,a_{n_2}-a_{n_3},\,a_{n_3}-a_{n_4})\,, \tag{4.41}$$

which makes it fairly obvious how to transcribe the arguments from the B's.

Next recall that the function f_4 is made up of coth functions whose arguments are the sum of arguments of the f_4, which in this case are $a_0 - a_{n_1}$, $a_0 - a_{n_2}$, $a_0 - a_{n_3}$, and $a_0 - a_{n_4}$. If the final $B_{n_4 0}$ were allowed to contribute an argument to an f, then the constituent coth would diverge, indicating why there is no such arrow in the above diagram.

This is also seen in our next example, where the matrix elements are $B_{0n_1} B_{n_1 0} B_{0n_2} B_{n_2 n_3} B_{n_3 0}$. As written, this would be transcribed to an f_4 as

$$\begin{array}{ccccc} B_{0n_1} & B_{n_1 0} & B_{0n_2} & B_{n_2 n_3} & B_{n_3 0} \\ \downarrow & \downarrow & \downarrow & \downarrow & \end{array}$$

$$f_4(a_0 - a_{n_1},\, a_{n_1} - a_0,\, a_0 - a_{n_2},\, a_{n_2} - a_{n_3})\,, \tag{4.42}$$

but the arguments to the coth functions would be $a_0 - a_{n_1}$, 0, $a_0 - a_{n_2}$, and $a_0 - a_{n_3}$. The second of these clearly leads to a divergence and so this ordering is prohibited. Instead, we should reorder these matrix elements to give an

allowed weight. Sensible options are

$$B_{0n_1}B_{0n_2}B_{n_10}B_{n_2n_3}B_{n_30} \to f_4(a_0 - a_{n_1},\, a_0 - a_{n_2},\, a_{n_1} - a_0,\, a_{n_2} - a_{n_3})\,, \tag{4.43}$$

$$B_{0n_1}B_{0n_2}B_{n_2n_3}B_{n_10}B_{n_30} \to f_4(a_0 - a_{n_1},\, a_0 - a_{n_2},\, a_{n_2} - a_{n_3},\, a_{n_1} - a_0)\,, \tag{4.44}$$

$$B_{0n_1}B_{0n_2}B_{n_2n_3}B_{n_30}B_{n_10} \to f_4(a_0 - a_{n_1},\, a_0 - a_{n_2},\, a_{n_2} - a_{n_3},\, a_{n_3} - a_0)\,, \tag{4.45}$$

which are exactly those found in equation (4.36). A few notes on ordering. Think of $B_{n_in_j}$ as an opening element for n_j and a closing one for n_i, then any order of B's must only close something which has already been opened. Next, from symmetry we are permitted to always begin our sequence with B_{0n_1}, i.e. the opening variable for n_1.

Terms involving B_{00} elements follow identical logic to that just discussed. Take, for example, the term involving $B_{00}B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_30}$. Clearly we cannot start our string with B_{00} as that would suggest zero is the first argument to f_4 which in turn would give zero as an argument to a coth function. Likewise, it cannot appear at the end of the string as then the sum of the arguments would equal zero. Instead it must appear in each of the other three available

spots, giving

$$B_{0n_1}B_{00}B_{n_1n_2}B_{n_2n_3}B_{n_30} \to f_4(a_0 - a_{n_1},\, 0,\, a_{n_1} - a_{n_2},\, a_{n_2} - a_{n_3})\,, \tag{4.46}$$

$$B_{0n_1}B_{n_1n_2}B_{00}B_{n_2n_3}B_{n_30} \to f_4(a_0 - a_{n_1},\, a_{n_1} - a_{n_2},\, 0,\, a_{n_2} - a_{n_3})\,, \tag{4.47}$$

$$B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{00}B_{n_30} \to f_4(a_0 - a_{n_1},\, a_{n_1} - a_{n_2},\, a_{n_2} - a_{n_3},\, 0)\,, \tag{4.48}$$

which again are the functions given in equation (4.36).

We shall now turn to the overall sign of these terms. In equation (4.36) some terms appeared with a plus sign, namely $B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_3n_4}B_{n_40}$, $B_{00}B_{0n_1}B_{n_10}B_{0n_2}B_{n_20}$, and $B_{00}B_{00}B_{0n_1}B_{n_1n_2}B_{n_20}$, while others appear with a minus sign, namely $B_{0n_1}B_{n_10}B_{0n_2}B_{n_2n_3}B_{n_30}$, $B_{00}B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_30}$, and $B_{00}B_{00}B_{00}B_{0n_1}B_{n_10}$. Deciding on this sign is actually very straightforward and involves partitioning. Simply put, split the term into all independent parts and each partition gives a minus sign. For example, $B_{00}|B_{0n_1}B_{n_10}|B_{0n_2}B_{n_20}$ is split into three independent parts, giving two partitions, and so has an overall plus sign. To give a second example, $B_{0n_1}B_{n_10}|B_{0n_2}B_{n_2n_3}B_{n_30}$ splits into two independent parts, giving one partition, and so has an overall minus sign.

Finally, the keen-eyed observer will notice some factors of two in the fourth term of equation (4.36). These numerical factors are somewhat mysterious and come from several sources in higher order eigenvalues $\varepsilon_0^{(N)}$. At this level, however, they are understood. The factors of two appear on the terms with B

ordering

$$B_{0n_1}B_{0n_2}B_{00}B_{n_20}B_{n_10} \qquad \text{and} \qquad B_{0n_1}B_{0n_2}B_{00}B_{n_10}B_{n_20}\,. \tag{4.49}$$

If one thinks of a B_{00} as somehow 'pausing' in the sequence, then when it is in this position we can choose to 'pause' either the open n_1 or the open n_2. This choice gives a factor of two. In any other sequences in this term there is no such choice and so no numerical factors appear. In later terms, we could have for example the sequence

$$B_{0n_1}B_{0n_2}B_{0n_3}B_{00}\cdots\,, \tag{4.50}$$

where the dots indicate that any permitted sequence could follow. This term would have a numerical factor of three, thanks to the choice of three open variables in this position, coming from that particular B_{00}. Any subsequent B_{00} within the remaining sequence could also contribute numerical factors in a similar fashion, with all such factors being multiplied to give the overall number.

As mentioned previously, other numerical factors also appear and seem to be related to the numbers of open variables, irrespective of the presence of a B_{00}.

For example, the sequence

$$B_{0n_1}B_{0n_2}B_{0n_3}B_{n_10}B_{n_20}B_{n_30}\,, \tag{4.51}$$

has a factor of two, which appears to come from the fact that at some point there were three open variables. As another example, the sequence

$$B_{0n_1}B_{0n_2}B_{0n_3}B_{0n_4}B_{n_10}B_{n_20}B_{n_30}B_{n_40}\,, \tag{4.52}$$

has a factor of six, which may be a factor of two from the point there were three open variables and a factor of three from the point where there were four open variables. As a final example, and to illustrate the mystery, the sequence

$$B_{0n_1}B_{0n_2}B_{0n_3}B_{n_3n_4}B_{n_10}B_{n_20}B_{n_40}\,, \tag{4.53}$$

has a factor of six. It has a point at which there are three open variables and then this is maintained to the next point, which appears to be a consideration when determining this number. Why this is so and the general rule-set for finding these numbers is currently unknown.

To summarise, with these simple but admittedly non-rigorous rules we can easily write down all sequences of B matrices involved in an eigenvalue $\varepsilon_0^{(N)}$, find the corresponding weights f_{N-1} multiplying each sequence, and say what sign

each sequence must have. We hit a stumbling block, however, in determining numerical factors which multiply some of these f_{N-1} functions. We have found up to $\varepsilon_0^{(6)}$ rigorously using the previously discussed perturbation theory and checked this rule-set with success.

4.4 Summary

To summarise, we have applied perturbation theory to our new representation of the Baker-Campbell-Hausdorff formula. This allows us access to the eigenvalues of the resultant operator. While this has been done in an ad-hoc manner without a full rigorous answer, we do have absolutely correct answers up to sixth order. We may now use these formulae to interrogate physical models.

CHAPTER 5

APPLYING PERTURBATION THEORY TO THE POTTS MODEL

With the mathematics of the previous chapters having been completed we may finally turn our attention away from such general concerns and towards practical matters. The q-state Potts model, held in abeyance since the introductory remarks, will be the focus of this chapter. It is not a particularly physical model, with the exception of some low q examples. However, it is a model with a clear mathematical foundation. Recall the basic ethos; spins which align gain energy while spins which do not, do not. This simplicity lends itself well to mathematical interrogation.

Other models can, of course, also be investigated using the techniques displayed here. We certainly do not wish to give the impression that only such symmetric models are accessible. What complications arise from abandoning the simplicity of the Potts model will be highlighted when they appear.

We shall begin this chapter by introducing the transfer matrix for the q-state Potts model and combining it with the previous Baker-Campbell-Hausdorff formula to find a low-temperature expansion for its partition function. We will find a great simplification of the formula, hinted at in the previous section. We shall then argue that duality allows us to immediately write down a high-temperature variant of this expansion. To consolidate this knowledge we will then provide plots for certain values of q, focusing particularly on the Ising case of $q = 2$. Note, we do not discover new physics in this chapter, nor did we set out to do. Instead we provide a new technique which can then contribute to the search for new phenomena.

5.1 Low-temperature expansion

Recall from chapter 1 that the transfer matrix for a square lattice model may be split

$$\hat{T} = \hat{T}_{\perp}^{\frac{1}{2}} \hat{T}_{/\!/} \hat{T}_{\perp}^{\frac{1}{2}} , \tag{5.1}$$

where $\hat{T}_\perp$ represents the transfer matrix for a ring of the cylinder while $\hat{T}_{/\!/}$ the matrix which transfers one ring to another. Explicitly for the q-state Potts model we may write

$$\hat{T}_\perp = e^{\beta J \sum_j \sum_\alpha |\alpha\rangle_j |\alpha\rangle_{j+1} \langle\alpha|_j \langle\alpha|_{j+1}} \,. \tag{5.2}$$

Note, the integer $\alpha \in \mathbb{Z}_q$ denotes the spin at a given site while j denotes the site itself, and $\langle\alpha|_j \, |\beta\rangle_i = \delta_{\alpha\beta}\delta_{ij}$ as one would expect. For the other constituent part of the transfer matrix we have

$$\hat{T}_{/\!/} = \prod_j \left(e^{\beta J} + \sum_{\alpha\beta} |\alpha\rangle_j \, \langle\beta|_j - 1 \right) , \tag{5.3}$$

where identity matrices may be placed where appropriate if one is so inclined. Here the first part ($e^{\beta J}$) concerns the situation where two neighbouring spins are aligned and energy is gained, while the second part describes the result when the spins are not aligned. The minus one allows for the case $\alpha = \beta$ in the sum. This should be compared with the transfer matrix for the Ising case, presented in equation (1.15).

As in the Ising case, we wish to write both of these constituent transfer matrices as exponentials so we can use our Baker-Campbell-Hausdorff formula. This is automatically done in the case of $\hat{T}_\perp$ but requires some effort for $\hat{T}_{/\!/}$. First let

us define $S_j = \sum_{\alpha\beta} |\alpha\rangle_j \langle\beta|_j$, then note

$$S_j^2 = \sum_{\alpha\beta\gamma\delta} |\alpha\rangle_j \langle\beta|_j |\gamma\rangle_j \langle\delta|_j \tag{5.4}$$

$$= q \sum_{\alpha\delta} |\alpha\rangle_j \langle\delta|_j \tag{5.5}$$

$$= qS_j \,. \tag{5.6}$$

Hence

$$e^{\frac{\beta\tilde{J}}{q} S_j} = \sum_n \frac{(\frac{\beta\tilde{J}}{q} S_j)^n}{n!} = 1 + \frac{1}{q} S_j \left(e^{\beta\tilde{J}} - 1 \right) . \tag{5.7}$$

Now then, we wish to recast the parallel transfer matrix as

$$\hat{T}_{/\!/} = e^{b_0 + \frac{\beta\tilde{J}}{q} \sum_j (S_j - 1)} , \tag{5.8}$$

where b_0 and $\tilde{J}$ are to be found. Parenthetically we ought to remark on the choice of variables. At first glance, writing $\beta\tilde{J}/q$ may seem strange as one could quite easily combine this into a single variable. However, we will find that this variable essentially acts as the high-temperature equivalent as βJ. The inclusion of the β then is immediately transparent; what then of the q? The answer to this lies in symmetry. Below the transition temperature spins predominantly point in but one of q directions. Above the transition temperature they are largely free to point in any of these directions. The new bond strength then must be divided by this number to be comparable to the

uni-directional bond strength. This becomes much clearer when one reaches the duality relation, which we shall do momentarily.

Using our newly found identity we may write

$$e^{b_0+\frac{\beta\tilde{J}}{q}\sum_j(S_j-1)} = e^{b_0}\prod_j e^{-\frac{\beta\tilde{J}}{q}}\left[1+\frac{1}{q}S_j\left(e^{\beta\tilde{J}}-1\right)\right] \tag{5.9}$$

$$= e^{b_0}\left[\frac{1}{q}e^{-\frac{\beta\tilde{J}}{q}}\left(e^{\beta\tilde{J}}-1\right)\right]^N\prod_j\left[\frac{q}{e^{\beta\tilde{J}}-1}+S_j\right]. \tag{5.10}$$

Comparison with equation (5.3) implies

$$e^{b_0}\left[\frac{1}{q}e^{-\frac{\beta\tilde{J}}{q}}\left(e^{\beta\tilde{J}}-1\right)\right]^N = 1\,, \tag{5.11}$$

where N denotes the size of the ring, and

$$\frac{q}{e^{\beta\tilde{J}}-1} = e^{\beta J}-1\,. \tag{5.12}$$

The former can be used to find b_0 and the latter b, but the latter should also be arranged into a more proper form

$$q = (e^{\beta J}-1)(e^{\beta\tilde{J}}-1) \tag{5.13}$$

and compared to the famous Kramers-Wannier duality for the Ising model

$$\sinh(2\beta J)\sinh\left(2\beta\tilde{J}\right) = 1. \tag{5.14}$$

Equation (5.13) is a duality relation for the Potts model. This will prove invaluable in mapping a low-temperature expansion to high-temperature, circumventing the need to additional calculation. It also provides the transition temperature, if one sets $J = \tilde{J}$. If these are both set to unity, as we will do in later numerical calculations, then we find that $T_c = 1/\log\left(1+\sqrt{q}\right)$. For now, however, we have succeeded in our aim to write both $\hat{T}_\perp$ and $\hat{T}_{/\!/}$ as exponentials and thus can use the Baker-Campbell-Hausdorff formula.

We now have

$$\hat{T} \equiv e^{-\beta\hat{F}} = e^{\frac{1}{2}\beta J\sum_j\sum_\alpha|\alpha\rangle_j|\alpha\rangle_{j+1}\langle\alpha|_j\langle\alpha|_{j+1}}\, e^{b_0+\frac{\beta\tilde{J}}{q}\sum_j(S_j-1)}\, e^{\frac{1}{2}\beta J\sum_j\sum_\alpha|\alpha\rangle_j|\alpha\rangle_{j+1}\langle\alpha|_j\langle\alpha|_{j+1}} \tag{5.15}$$

with the task being to find the free-energy operator $\hat{F}$. We may rewrite this as

$$e^{-\beta\hat{F}} = e^{b_0}e^{A}e^{2B}e^{A}\,, \tag{5.16}$$

where

$$A \equiv \frac{1}{2}\beta J\sum_j\sum_\alpha|\alpha\rangle_j|\alpha\rangle_{j+1}\langle\alpha|_j\langle\alpha|_{j+1}\,, \tag{5.17}$$

and

$$B \equiv \frac{1}{2}\frac{\beta\tilde{J}}{q}\sum_j (S_j - 1) = \frac{1}{2}\frac{\beta\tilde{J}}{q}\sum_j \left(\sum_{\alpha\beta} |\alpha\rangle_j \langle\beta|_j - 1\right) , \tag{5.18}$$

to put it in a form familiar from the previous chapters. The task now is to find the partition function, Z, via

$$\log Z = b_0 + \varepsilon_0^{(0)} + \varepsilon_0^{(1)} + \varepsilon_0^{(2)} + \cdots , \tag{5.19}$$

where $\varepsilon_0^{(n)}$, which shall be restated shortly, are the perturbative values calculated in the previous chapter.

First, recall that $\varepsilon_0^{(0)} = 2a_0$, the largest eigenvalue of the matrix A, hence

$$\varepsilon_0^{(0)} = N\beta J . \tag{5.20}$$

Next consider $\varepsilon_0^{(1)} = 2B_{00}$. Our matrix for B is entirely off-diagonal, so $B_{00} = 0$ and therefore $\varepsilon_0^{(1)} = 0$. This fact leads to great simplification later in the perturbation series as all parts which contain a B_{00} are automatically zero. This is true in the calculation of any model's partition function but ceases to be true when one wishes to calculate higher excitations.

Now let us turn to the first non-trivial calculation, $\varepsilon_0^{(2)}$. Recall,

$$\varepsilon_0^{(2)} = 2\sum_{n_1} B_{0n_1} B_{n_1 0}\, f_1(a_0 - a_{n_1}) . \tag{5.21}$$

This formula is essentially saying that the matrix B acts on the groundstate 0 to take it to the state n_1, then acts on the state n_1 to take it back to the groundstate. The matrix B acts on a state by taking a spin on a site and changing it; as such, the state n_1 can differ from the groundstate 0 only by a single spin-flip. However, this spin-flip can be on any of the N sites and the spin can be flipped to any of $q-1$ different states. Pictorially,

$$\begin{aligned} \text{groundstate}: &\qquad \cdots 0\,0\,0\,0\,0\cdots \\ n_1: &\qquad \cdots 0\,0\,\alpha\,0\,0\cdots \\ \text{groundstate}: &\qquad \cdots 0\,0\,0\,0\,0\cdots \end{aligned}$$

Here the zeros indicate spins which are in their groundstate alignment, chosen arbitrarily as one of the q spin options, while α indicates a spin which is not in this orientation. The eigenvalue of the matrix A associated with the excited state is

$$a_{n_1} = a_0 - \beta J\,, \tag{5.22}$$

as the flipped spin has broken two of the bonds, hence we have

$$\varepsilon_0^{(2)} = 2(q-1)N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^2 f_1(\beta J)\,. \tag{5.23}$$

We have thus calculated the logarithm of the partition function to second

order in our perturbative series. Later calculations follow very similar, albeit increasingly complicated, logic.

Turning now to calculating $\varepsilon_0^{(3)}$,

$$\varepsilon_0^{(3)} = 2 \sum_{n_1 \neq 0, n_2 \neq 0} B_{0n_1} B_{n_1 n_2} B_{n_2 0} \, f_2(a_0 - a_{n_1},\, a_{n_1} - a_{n_2}) - \sum_{n_1 \neq 0} B_{00} B_{0n_1} B_{n_1 0} \, f_2(a_0 - a_{n_1},\, 0)\,. \tag{5.24}$$

Of course, as mentioned previously $B_{00} = 0$ and so the second part of this equation is irrelevant. The first part essentially tells us that by a single spin-flip we go from the groundstate to n_1, by a single spin-flip from n_1 to n_2, then by a single spin-flip from n_2 back to the groundstate. These restrictions limit what these states may be. The first state, n_1, can be thought of identically as before so the only new thing to consider is n_2. Any amount of thought will lead one to the conclusion that the only possible option is the following diagram

$$\begin{aligned} \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0 \cdots \\ n_1: &\quad \cdots 0\,0\,\alpha\,0\,0 \cdots \\ n_2: &\quad \cdots 0\,0\,\beta\,0\,0 \cdots \\ \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0 \cdots \end{aligned}$$

As before α must not be the same orientation as zero but also β must be

different from both α and zero. Any attempts in the second step to flip a spin other than the one previously flipped will prevent a return to the groundstate. We thus have

$$a_{n_1} = a_0 - \beta J\,, \tag{5.25}$$

as before, and

$$a_{n_2} = a_{n_1} \tag{5.26}$$

as no additional bonds have been broken. As such, we have found

$$\varepsilon_0^{(3)} = 2(q-1)(q-2)N \left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^3 f_2(\beta J,\, 0)\,. \tag{5.27}$$

As a final example, let us now think about $\varepsilon_4^{(0)}$,

$$\begin{aligned}
\varepsilon_0^{(4)} = 2 \sum_{n_1\neq 0\,, n_2\neq 0\,, n_3\neq 0} & B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_30}\, f_3(a_0-a_{n_1},\, a_{n_1}-a_{n_2},\, a_{n_2}-a_{n_3}) \\
-4 \sum_{n_1\neq 0\,, n_2\neq 0} & B_{0n_1}B_{n_10}B_{0n_2}B_{n_20} \Big[f_3(a_0-a_{n_1},\, a_0-a_{n_2},\, a_{n_1}-a_0) \\
& \qquad + f_3(a_0-a_{n_1},\, a_0-a_{n_2},\, a_{n_2}-a_0)\Big] \\
- \sum_{n_1\neq 0\,, n_2\neq 0} & B_{00}B_{0n_1}B_{n_1n_2}B_{n_20} \Big[f_3(a_0-a_{n_1},\, 0,\, a_{n_1}-a_{n_2}) \\
& \qquad + f_3(a_0-a_{n_1},\, a_{n_1}-a_{n_2},\, 0)\Big] \\
+ \sum_{n_1\neq 0} & B_{00}B_{00}B_{0n_1}B_{n_10}\, f_3(a_0-a_{n_1},\, 0,\, 0)\,.
\end{aligned} \tag{5.28}$$

Only the first two terms are non-zero and need to be considered. Take the first term. A possible diagram, following in the footsteps of our previous results, is

$$\begin{aligned}
\text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,0\cdots \\
n_1: &\quad \cdots 0\,0\,\alpha\,0\,0\cdots \\
n_2: &\quad \cdots 0\,0\,\beta\,0\,0\cdots \\
n_3: &\quad \cdots 0\,0\,\gamma\,0\,0\cdots \\
\text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,0\cdots
\end{aligned}$$

Here we require that $\alpha \neq 0$, $\beta \neq 0$ and $\beta \neq \alpha$, and $\gamma \neq 0$ and $\gamma \neq \beta$ but not $\gamma \neq \alpha$. That is, the states n_1 and n_3 could very well be the same. They do not have to be, but they can be without issue. This leads to the contribution

$$(q-1)(q-2)^2 2N \left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\, 0,\, 0)\,. \tag{5.29}$$

This is, of course, not the only permissible diagram. With four spin-flips allowed to us we have the freedom to flip more than one spin and still return

to the groundstate. For example,

$$\begin{aligned} \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,0\cdots \\ n_1: &\quad \cdots 0\,0\,\alpha\,0\,0\cdots \\ n_2: &\quad \cdots 0\,0\,\alpha\,\alpha\,0\cdots \\ n_3: &\quad \cdots 0\,0\,\alpha\,0\,0\cdots \\ \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,0\cdots \end{aligned}$$

In this diagram the second spin to be flipped can be either side of the first one, picking up a factor of two, and either spin could be reversed between n_2 and n_3, picking up another factor of two. The energy difference between states, however, does not change as no additional bonds have been broken; instead a domain wall has merely been moved. This contribution is thus

$$(q-1)8N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\,0,\,0)\,. \tag{5.30}$$

Another acceptable diagram is the same as before but the second spin-state is

different to the first. That is,

$$\begin{aligned}
\text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,0\cdots \\
n_1: &\quad \cdots 0\,0\,\alpha\,0\,0\cdots \\
n_2: &\quad \cdots 0\,0\,\alpha\,\beta\,0\cdots \\
n_3: &\quad \cdots 0\,0\,\alpha\,0\,0\cdots \\
\text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,0\cdots
\end{aligned}$$

Again, the second spin can appear on either side of the first and either of the two spins can disappear as we go from n_2 to n_3. This time, however, the energy of state n_2 differs from that of n_1 and n_3 as an additional bond has been broken. Here

$$a_{n_2} = a_{n_1} - \frac{1}{2}\beta J \tag{5.31}$$

and

$$a_{n_3} = a_{n_2} + \frac{1}{2}\beta J \tag{5.32}$$

hence this diagrams contribution is

$$(q-1)(q-2)8N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\, \frac{1}{2}\beta J,\, -\frac{1}{2}\beta J)\,. \tag{5.33}$$

Finally, consider the diagram

$$\begin{aligned} \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,0\cdots \\ n_1: &\quad \cdots 0\,\alpha\,0\,0\,0\cdots \\ n_2: &\quad \cdots 0\,\alpha\,0\,\beta\,0\cdots \\ n_3: &\quad \cdots 0\,\alpha\,0\,0\,0\cdots \\ \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,0\cdots \end{aligned}$$

It does not matter if $\beta = \alpha$ or not as the energy of the state n_2 would always be

$$a_{n_2} = a_{n_1} - \beta J\,, \tag{5.34}$$

and the energy of the state n_3

$$a_{n_3} = a_{n_2} + \beta J\,, \tag{5.35}$$

Likewise, it does not matter where the second spin is flipped, so long as it does not neighbour the first. If it did neighbour the first it would be one of the diagrams we have already considered. As such, this diagram's contribution is

$$(q-1)^2 4N(N-3)\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\,\beta J,\,-\beta J)\,. \tag{5.36}$$

This expression may give cause for alarm; there is an N^2 component. We are

essentially calculating a free-energy, after all, which is an extensive quantity and thus we expect everything to be of order N. Fear not. We still have the second term in $\varepsilon_4^{(0)}$ to consider and we will conveniently find that its contributions as such that they exactly cancel any N^2 term here. Indeed, in higher order calculations of $\varepsilon_m^{(0)}$ it will always be such that the very first term, that containing $B_{0n_1}B_{n_1n_2}\cdots B_{n_{m-1}n_m}B_{n_m0}$, is the only term with contributions of order N. It will have other parts as well, of order N^2, N^3, etc, but these corrections will be found in the other terms of $\varepsilon_0^{(m)}$ with precisely the opposite sign. Thus if one is being practical one can only calculate the diagrams for the very first term and just disregard anything which is not of order N, safe in the knowledge that it would cancel anyway. To be clear, this is not a rigorous statement. It is, however, exactly what one would expect to happen from physical concerns.

To illustrate this point, let us now consider the second term in $\varepsilon_0^{(4)}$, namely

$$-\sum_{n_1\neq 0\,,n_2\neq 0} B_{0n_1}B_{n_10}B_{0n_2}B_{n_20}\Big[f_3(a_0-a_{n_1},\,a_0-a_{n_2},\,a_{n_1}-a_0) + f_3(a_0-a_{n_1},\,a_0-a_{n_2},\,a_{n_2}-a_0)\Big]\,. \quad (5.37)$$

The cancelling minus sign is already conspicuously present. Look first at the B matrices. They clearly separate into two distinct parts, similar perhaps to the separated spins in the problematic final diagram. It is this separation which

allows one to claim that of course it cannot contribute to the extensive free-energy and that of course some miraculous mathematical cancellation must take place. To see this miracle take place consider the only diagram which is acceptable to this term,

$$\begin{aligned} \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,\cdots \\ n_1: &\quad \cdots 0\,\alpha\,0\,0\,0\,\cdots \\ \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,\cdots \end{aligned}$$

$$\begin{aligned} \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,\cdots \\ n_2: &\quad \cdots 0\,0\,0\,\beta\,0\,\cdots \\ \text{groundstate}: &\quad \cdots 0\,0\,0\,0\,0\,\cdots \end{aligned}$$

We have split this diagram into two parts, to represent the two independent parts $B_{0n_1}B_{n_10}$ and $B_{0n_2}B_{n_20}$ The first spin α can be any of $q-1$ variants on any of N sites. Similarly the second spin β can also be any of $q-1$ variants on any of N sites; it does not care about the previous existence of the first spin. The energies of these states are of course identical,

$$a_{n_1} = a_{n_2} = a_0 - \beta J\,, \tag{5.38}$$

and so careful inputting of this information into the term (5.37) provides

$$-4(q-1)^2N^2\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\,\beta J,\,-\beta J)\,. \tag{5.39}$$

This cancels exactly with the N^2 contribution calculated before.

The final answer is then

$$\begin{aligned}\varepsilon_0^{(4)} = {} & (q-1)(q-2)^2 2N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\,0,\,0)\\ & + (q-1)8N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\,0,\,0)\\ & + (q-1)(q-2)8N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\,\frac{1}{2}\beta J,\,-\frac{1}{2}\beta J)\\ & - (q-1)^2 12N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\,\beta J,\,-\beta J)\,.\end{aligned}$$

These results indicate how to perform this calculation in the most simple way. There are better and worse ways to proceed which may be more generalisable. However, the calculation as presented is intended to be clear and understandable and the underlying physical ideas immediately apparent. In order to aid this clarity we shall now present specific Potts models and plot the results, starting of course with the Ising model.

5.2 Perturbation results for $q = 2$

The 2-state Potts model is the simplest starting point for showing tangible results. It is the Ising model, albeit rescaled slightly. In a typical ferromagnetic Ising set-up spins which align gain energy while spins which do not lose energy; if we just put $q = 2$ into the previous results we would obtain the former but not the latter. We will calculate with the Potts formulation but perform this rescaling when it comes to comparing to exact results in plots.

The physical simplicity of the Ising model leads to mathematical simplicity in the perturbation result. As there are only two spin-states, up or down, we can no longer have the situation where a spin on one site changes from the ground-state configuration to another, then to another again, then back to the ground-state. This eliminates all odd orders in the previous calculation as well as greatly reducing the terms in all even orders. This reduction appears mathematically as anything multiplied by $q - 2$ is zero when $q = 2$. The previous results thus become

$$\log Z = b_0 + N\beta J + 2N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^2 f_1(\beta J) + 8N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\, 0,\, 0) - 12N\left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\, \beta J,\, -\beta J)\,. \quad (5.40)$$

We can go much further than fourth order. One could do this by hand, in the

style of the previous section. However, it is much simpler to cheat and use various observations to our advantage. First, recall the claim that it is only the first term, for example $B_{0n_1}B_{n_1n_2}B_{n_2n_3}B_{n_30}$ at fourth order, which will contribute objects of order N. If we take this to be true then we only need to calculate this term and ignore any results which are of higher order, with the assumption that these extraneous results would be inevitably cancelled with the uncalculated terms.

We can take this one step further and say exactly what we will be ignoring. At fourth order we noticed that there was a term of order N^2, thanks to the two distinct parts that four spin-flips allows. At sixth order then we can predict that six spin-flips will allow there to be a maximum of three independent parts and hence we would expect a term of order N^3. There will also be terms of order N^2 and N, so in general the terms we calculate will be multiplied by a third order polynomial in N. Logically then we would expect that at M-th order the terms in our perturbation theory will be multiplied by a $M/2$-th order polynomial in N.

This observation leads to a method of cheating the calculation, as alluded to earlier. We can ask a computer to calculate all the different possible states resulting from a spin-flip for a finite ring and write down the energies of each of these states. We can then count the number of states with the same energy. If we do this for enough different sized rings then we will have enough data

to use Lagrange interpolation to find exactly the $M/2$-th order polynomial we are seeking. The energies we record can then be used as arguments for the functions f_{M-1} and the result thus immediately written down.

With this computer aided mathematics we can go much further than one could reasonably hope to by hand. Indeed, as we shall see shortly, we have used this technique to get to order ten.

First, however, let us illustrate this point with a concrete example. Take fourth order, which we have already calculated. We have four spin-flips to use and we must begin and end in the groundstate without returning there in-between. Let us begin with a finite ring of size four in its groundstate,

$$0\,0\,0\,0$$

Without loss of generality we can tell the computer to instead begin with the state

$$0\,0\,0\,1$$

effectively putting in the application of B_{0n_1} by hand. This small optimisation effectively divides the polynomial we are calculating by N and thus reduces the number of finite systems we need by one. It is merely translational invariance in action.

Next then we ask the computer to print out all possible states which differ

from this by a single spin-flip and are not the groundstate, and also to print out the energy difference between these new states and the last. It returns to us the following

$$\begin{array}{rcll} & & 0\,0\,1\,1 & \text{Energy difference : } 0 \\ 0\,0\,0\,1 & \rightarrow & 0\,1\,0\,1 & \text{Energy difference : } \beta J \\ & & 1\,0\,0\,1 & \text{Energy difference : } 0 \end{array}$$

We then take the each state in turn and ask it to remove a spin-flip to help us slowly get back to the groundstate.

$$\begin{array}{rcll} 0\,0\,1\,1 & \rightarrow & 0\,0\,0\,1 & \text{Energy difference : } 0 \\ & & 0\,0\,1\,0 & \text{Energy difference : } 0 \end{array}$$

$$\begin{array}{rcll} 0\,1\,0\,1 & \rightarrow & 0\,0\,0\,1 & \text{Energy difference : } -\beta J \\ & & 0\,1\,0\,0 & \text{Energy difference : } -\beta J \end{array}$$

$$\begin{array}{rcll} 1\,0\,0\,1 & \rightarrow & 0\,0\,0\,1 & \text{Energy difference : } 0 \\ & & 1\,0\,0\,0 & \text{Energy difference : } 0 \end{array}$$

We thus have 4 routes with energy differences (βJ, 0, 0) and 2 routes with

energy differences (βJ, βJ, $-\beta J$), where we have included the first difference from the groundstate to the state 0 0 0 1 in the energy differences.

Repeating this process with a ring of five spins provides 4 routes with energy differences (βJ, 0, 0) and 4 routes with energy differences (βJ, βJ, $-\beta J$). As we are expecting a polynomial of up to N^2 but have effectively divided it by N, this data is sufficient to calculate the result. We are expecting the result to be

$$N(c_1 + c_2 N) f_3(\beta J,\, 0,\, 0) + N(d_1 + d_2 N) f_3(\beta J,\, \beta J,\, -\beta J)$$

Reintroducing the factor of N to our data gives the simultaneous equation

$$4(c_1 + 4c_2) = 16 \quad \text{and} \quad 5(c_1 + 5c_2) = 20 \qquad \Longrightarrow\ c_1 = 4\,, c_2 = 0\,, \tag{5.41}$$

$$4(d_1 + 4d_2) = 8 \quad \text{and} \quad 5(d_1 + 5d_2) = 20 \qquad \Longrightarrow\ d_1 = -6\,, d_2 = 2\,. \tag{5.42}$$

As we are ignoring anything other than the order N contribution we thus have the answer

$$\varepsilon_0^{(4)} = 8N \left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\, 0,\, 0) - 12N \left(\frac{1}{2}\frac{\beta\tilde{J}}{q}\right)^4 f_3(\beta J,\, \beta J,\, -\beta J)\,, \tag{5.43}$$

which is precisely what was calculated earlier, after we reintroduce the factors of $2(\beta\tilde{J}/2q)^4$.

This idea can be put into practice with a computer and results up to tenth

order obtained. In principle one could go further but the number of states increases exponentially with each order and so obtaining the twelfth order result is prohibitively expensive in time. Code could be perhaps optimised to somewhat reduce this time, but realistically this Ising calculation is more proof-of-concept rather than something interesting in and of itself. As such, one should not expend too much effort on it.

The reason that the Ising model is a good test calculation, beyond the immediate benefit of a binary system when we perform these computational tricks, is that an exact solution exists. We can thus compare our perturbative results to this exact solution to see how well it performs, then hope that its level of performance can be considered similar in models for which we do not have a solution. It is customary to compare not the partition function but rather an observable. The heat capacity is a sensible choice. As the Ising model undergoes a second order phase transition, it is the second derivative of the free-energy which diverges. It will be interesting to see how our perturbative calculation approaches the singularity at the critical point.

The heat capacity is given by

$$C = \beta^2 \frac{\partial^2}{\partial \beta^2} \log Z \,, \tag{5.44}$$

and we have conveniently calculated a formula for $\log Z$. The exact result is

the rather complicated formula [81]

$$C = \frac{2}{\pi}(\beta J \coth 2\beta J)^2 \left[2K(k) - 2E(k) - (1 - \sqrt{1-k^2})\left(\frac{\pi}{2} + \sqrt{1-k^2}K(k)\right)\right], \tag{5.45}$$

where

$$K(k) = \int_0^{\frac{\pi}{2}} \left[1 - k^2 \sin x\right]^{-\frac{1}{2}} \mathrm{d}x\,,$$
$$E(k) = \int_0^{\frac{\pi}{2}} \left[1 - k^2 \sin x\right]^{\frac{1}{2}} \mathrm{d}x\,,$$

are elliptic integrals and

$$k \equiv \frac{2\sinh 2\beta J}{\cosh^2 2\beta J}\,. \tag{5.46}$$

Despite its somewhat intimidating form this can be easily plotted alongside our perturbative calculation and the results seen in figure 5.1. What is shown in this figure is a succession of perturbative plots, each including one more perturbative order than the last. In other words, the line which performs worst is the perturbative calculation up to second order, the next worst the calculation up to fourth order, and so on until the line which performs the best is the calculation up to tenth order. Beneath this plot are a pair of lines with circles drawn on. The first blue circle on the top line indicates the point at which the second order calculation ceases to agree with the exact answer to machine accuracy, the second blue circle the point at which the fourth order

ceases to agree to machine accuracy, and so on. The second line contains the same information to but within 1% rather than machine accuracy. The bounding red circles indicate zero temperature and the transition temperature, to give an indication of scale. For reference, the transition temperature of the square lattice Ising model is $2/\log\left(1+\sqrt{2}\right) \sim 2.27$ when $J = 1$ as in the figure.

As we can see, the perturbative calculations perform exceptionally well. They agree with the exact answer to machine accuracy for a large temperature range then to within one percent, which is more accurate than one can detect by eye, for much further. Of course, they do not and indeed cannot reach the critical point but this is not unexpected. As this is similar to the well-known high-temperature expansion in methodology this new technique must also suffer from the same impotency in the face of a divergence. This is not to be decried but rather accepted as the cost of using a perturbative technique. There are ways to ameliorate this condition, such as the use of Padè approximants, but this will not be discussed here.

If we were to add more orders to the calculation we would expect that the additional region of agreement between the new order and the exact solution to decrease with each order added. That is, the amount of benefit one gets from each additional order would be smaller and smaller. In principle, one would need to go to infinite order to reach the critical point.

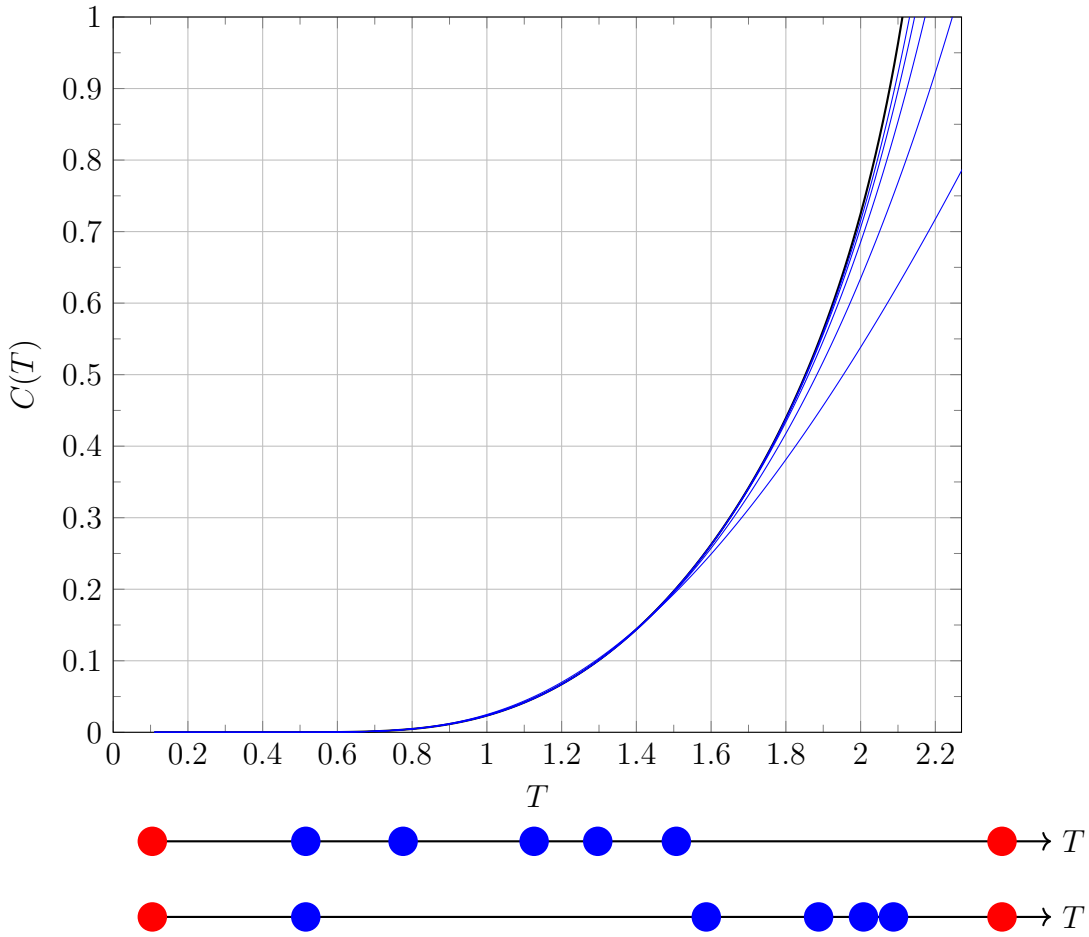


Figure 5.1: The heat capacity of the Ising model. The black line represents the exact results. The blue lines represent the result coming from the perturbative series, including more terms up to tenth order. The blue circles below indicate how accurate the perturbative series is, with red circles setting the scale at $T = 0$ and $T = T_c$. The first blue circle on the top line appears where the lowest order perturbative answer ceases to agree with the exact result to machine accuracy, the second the next lowest order, etc. The blue circles on the bottom line perform the same role but appear when the perturbative answer ceases to agree with the exact result to within 1% accuracy.

The basic physical idea behind this is fairly obvious. At low temperatures fluctuating spins are fairly isolated. At a slightly higher temperature, however, it is more likely that its neighbours will also fluctuate. As the temperature increases so does the probable range of this fluctuating neighbourhood. The perturbation calculation takes into account more and more spins at higher and higher order, reflecting this physical observation.

5.3 Perturbation results for $q \geq 3$

Let us now turn to more complicated Potts models. For $q \geq 3$ there exists no exact solution for the q-state Potts model and so we cannot compare our perturbation results to a known answer, as before. Instead, we must use some other method to determine for what temperature range the result is valid.

One obvious way is to check where different orders of the expansion agree. If up to some temperature the second order expansion is the same as all higher order expansions to machine accuracy, one can assume it is also the same as the exact answer to the same accuracy. This assumption is grounded in the previous results for the Ising model, where this was fact.

At some temperature the second order result will begin to differ. At this point we must say it has become incorrect and no longer consider it accurate. The third order should continue to agree with higher orders, however, for more

of the temperature range before it too disagrees. Following this trend we may find the highest temperature that our second-highest order agrees with our highest order and claim we know we have the correct answer to a given accuracy below that point. Beyond this point it is likely that we still have the correct answer, contained within our highest order, but unfortunately we have no way of verifying it.

Figure 5.2 shows this picture for the $q = 3$ Potts model. We can confirm the veracity of our answer up to fourth order for some of the temperature range, but our fifth order result is somewhat wasted. We should note that while the agreement up to machine accuracy seems to be only at a low temperature, before any significant features have occurred in the heat capacity, this was also true in our Ising calculation. In that model we saw that the data still looked to the eye extremely accurate beyond that point. As such, it is likely that we are seeing the real shape of the heat capacity in this plot.

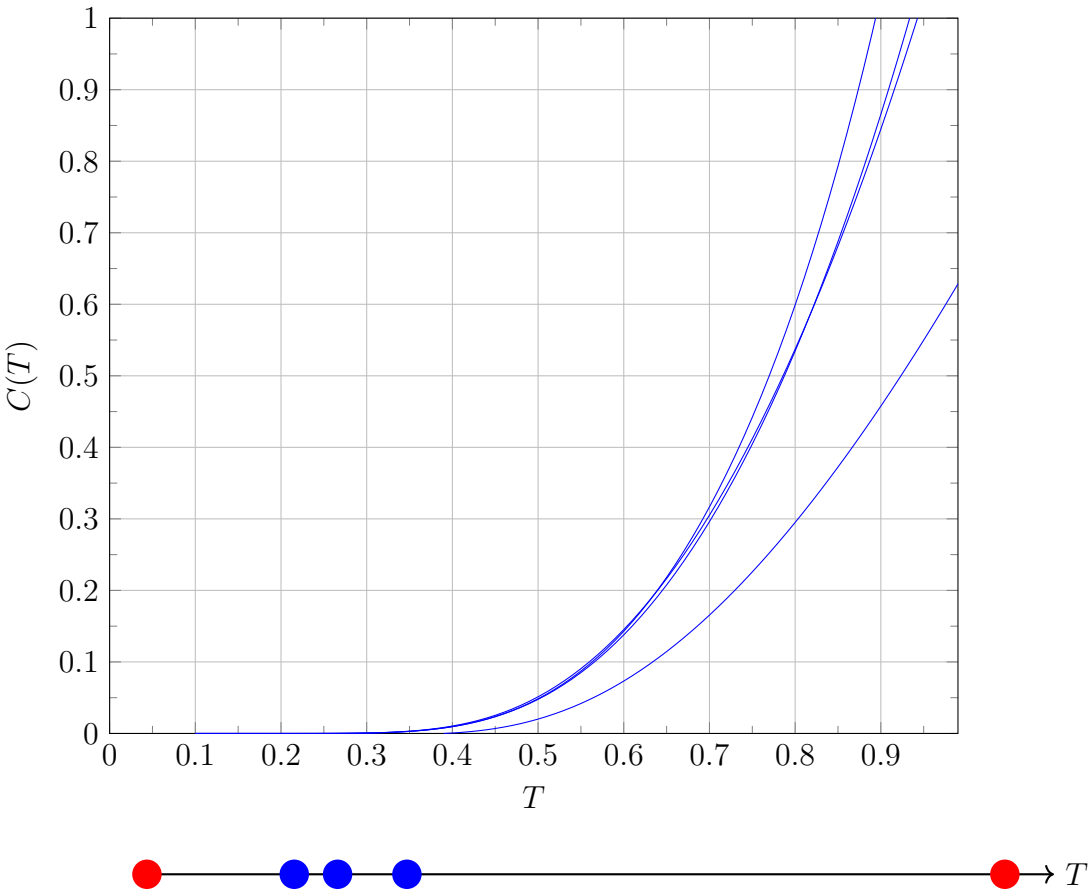


Figure 5.2: The heat capacity of the 3-state Potts model. The lines represent the result coming from the perturbative series, including more terms up to fifth order. The blue circles below indicate how accurate the perturbative series is, with red circles setting the scale at $T = 0$ and $T = T_c$. The first blue circle on the line below appears where the lowest order perturbative answer ceases to agree with the highest order answer to machine accuracy, the second circle when the next lowest order answer ceases to agree with the highest order answer, etc.

Next, figure 5.3 shows the accuracy of our result for a variety of Potts models. The Ising result up to fourth order has also been displayed for reference. Scaled by their differing transition temperatures, the results we get are remarkably similar in each case. However, there does seem to be a slight trend that shows

the region of accuracy for our perturbative series increasing as a function of q. More work would need to be done to understand and indeed confirm this. It is though an early indication that this perturbative formula may hold some interesting hidden information.

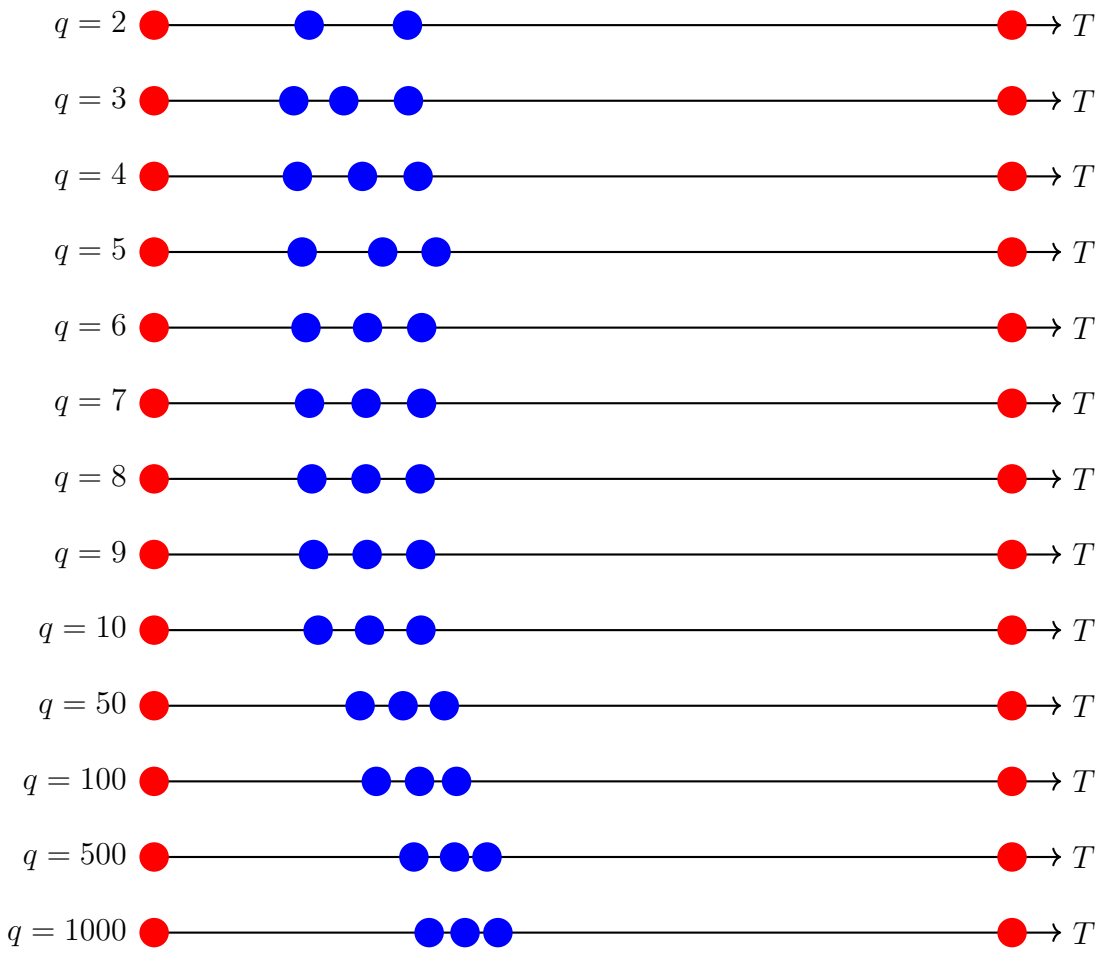


Figure 5.3: Accuracy of the perturbative formula for various Potts models. The red circles set the temperature scale at $T = 0$ and $T = T_c$. The first blue circle on the line below appears where the lowest order perturbative answer ceases to agree with the highest order answer to machine accuracy, the second circle when the next lowest order answer ceases to agree with the highest order answer, etc.

CHAPTER 6

CONCLUSION AND OUTLOOK

This work has explored a small subset of statistical mechanics ideas on the square lattice. We began by employing exact diagonalisation on transfer matrices, recasting the matrix as a free-energy operator and arguing in the process that one ought to care for more than just its groundstate. We have ended by finding just its groundstate for the Potts model. Such is the curse of a project of finite length, among other things. As such, these concluding remarks will be followed by an outlook; a proposal of what future work may look like through the eyes of someone unlikely to do it.

First, however, let us remember what this work has accomplished. Chapter 2,

the motivational chapter which began this work, proposed thinking of transfer matrices as the exponential of a free-energy operator and thinking of that free-energy operator as any self-respecting quantum mechanist would: through its spectrum. It argued that its spectrum described physical phenomena, namely thermalised domain walls propegating in the direction of transfer. It presented copious amounts of numerical evidence to back this up. We were then left asking the question – can we attack this analytically?

Chapter 3 presented a way in which we could. It is really the highlight of this work, solving a problem in mathematics which had been open for over a hundred years. Of course one could argue, with some justification, that no one in the intervening century had considered it a problem, but that does not take away from the solving of it. Nor should it reduce its future significance to theoretical physics.

Chapters 4 and 5 presented just one way in which the new representation of the Baker-Campbell-Hausdorff formula has the potential to be exceptionally valuable. The former essentially simplified the mathematics to the point that any reasonably competent physicists could use it. No longer would they have to worry about the adjoint endomorphisms or complicated operators of the preceding chapter. Instead we just use perturbation theory, that ubiquitous technique that theorists seemingly imbibe in the cradle, along with a curious disdain for pure mathematics. The latter chapter then saw the resulting

perturbative formula used in action, deployed against the Potts model. No one can argue that this model is not important, at least on a mathematical level.

However, this formula has the potential to be used much more widely. The technique we presented in those later chapters can just as easily be used on any square lattice spin model, be it the clock model or something more general. Beyond statistical mechanics as well it has potential, though perhaps it is safer for the purposes of this document not to comment further.

6.1 What is to be done?

This section will contain some suggestions for natural future work on this project. It will be arranged in the order that these issues arose in this document, but this is not the order that they should be dealt with.

6.1.1 Proving the non-singularity of the Baker-Campbell-Hausdorff formula

This proposal concerns the L'Hôpital issues discussed in chapter 3. We could not prove the formula is non-singular which rankles ever so slightly. It is the only non-proved part of that chapter and thus rankles ever so slightly. However,

it is fairly obvious that the formula is not singular so this suggestion is more intended to soothe ones conscience rather than provoke serious mathematical invention.

6.1.2 Making perturbation theory rigorous

This proposal is, however, intended to be taken seriously. Chapter 4 was astonishingly unsatisfying, particularly when following the previous chapter. Admittedly the work contained within it is much more recent and thus much more raw, but it still causes distress. We only calculated the perturbative Baker-Campbell-Hausdorff formula up to some finite order, rather than finding a general formula. The finding and proving of this formula should be high on anyone's agenda.

Just finding it would be an accomplishment, as the numerology at the end of the section attests. One could also consider just proving parts of the formula, as the subsequent section demonstrated that large swathes of the result inevitably cancels.

6.1.3 Proving cancellation

This cancellation, observed at the beginning of chapter 5 for the Potts model, is expected to take place from purely physical concerns. Proving mathematically

that it must do so would be a great boon.

6.1.4 Calculate higher eigenvalues

Chapter 5 presented a way to use the perturbative Baker-Campbell-Hausdorff formula to find the groundstate of the Potts model's associated free-energy operator. One can also use it to find other eigenvalues, thus finally achieving what was intended at the beginning of this work. This is a significantly harder calculation, though still fairly straightforward, but time pressure has led to this remaining undone.

6.1.5 Use the results to say something meaningful

We just calculated the perturbative formula, we did not actually use it to say anything. The Potts model is a very interesting in its own right and one could ask our technique to shine some light on it. Whether this is done by the use of Padè approximants to extrapolate to the critical region, as suggested in the text, or otherwise, it is something that remains undone.

In particular, the Potts model changes the nature of its phase transition as q is changed. It goes from being a second order transition for $q \leq 4$ to a first order transition. While we discussed the idea that one would need to calculate the perturbative formula to infinite order to get accurate results for the Ising

model at the critical point, this may not be true for a first order transition. Instead, as the low-temperature expansion would know nothing of the state it is meant to jump to at the critical point, it may just sail past it oblivious to the transition the model undergoes. Without knowledge of that transition, why should it require an infinite order to reach this point? That is, in principle the calculation of the low-temperature expansion could provide machine accurate results up to and indeed beyond the transition temperature, with any error only beginning beyond this point. The high-temperature version would do the same, with errors only beginning below the transition temperature. The combination of these results then would, in principle, give the partition function to machine accuracy for all temperature. We should stress, however, that this discussion is highly speculative.

6.1.6 Apply the results to more models

Finally, the most obvious future direction. We have presented work here on the Potts model alone. Anyone can see that the technique is not special to this model and exploration into other realms has only just begun.

APPENDIX A

ALGORITHMICALLY REMOVING APPARENT SINGULARITIES

This appendix will provide an algorithmic approach to removing any apparent singularities in the operator $\hat{G}_N$, using the representation and identities provided in section 3.4. For immediate use, the first four operators $\hat{G}_N$ have formulae provided for all possible singularities which are given towards the end of this appendix. However, first the general trends shall be discussed via a single larger example, namely $\hat{G}_6$ when all of L_2, L_3, L_4, and L_5 are simultaneously zero.

In the language of section 3.4, that is equation (3.79), the relevant part of $\hat{G}_6$ without any singularities may be written as

$$\begin{aligned} g_6(L_1,L_2,L_3,L_4,L_5,L_6) = {} & E(L_1,L_2+L_3+L_4+L_5+L_6)\, f_4(L_2,L_3,L_4,L_5) \\ & + E(L_1+L_2,L_3+L_4+L_5+L_6)\, f_1(-L_2) f_3(L_3,L_4,L_5) \\ & + E(L_1+L_2+L_3,L_4+L_5+L_6)\, f_2(-L_3,-L_2) f_2(L_4,L_5) \\ & + E(L_1+L_2+L_3+L_4,L_5+L_6)\, f_3(-L_4,-L_3,-L_2) f_1(L_5) \\ & + E(L_1+L_2+L_3+L_4+L_5,L_6)\, f_4(-L_5,-L_4,-L_3,-L_2)\,. \end{aligned}$$

There are five limits to be taken and the order in which they should be performed is crucial. For the approach which will be laid out in this section, it is best to work from the outside in. That is, it is best to take the limit $L_2+L_3+L_4+L_5 \to 0$ first, followed by $L_3+L_4+L_5 \to 0$, and so on. The reason for this will become apparent shortly. For now, under the first limit, both the first and last lines appear singular while the rest are regular. The identity, associated with a_6^{even} in equation (3.77),

$$\begin{aligned} & f_4(L_2,L_3,L_4,L_5) + f_1(-L_2) f_3(L_3,L_4,L_5) + f_2(-L_3,-L_2) f_2(L_4,L_5) \\ & \qquad + f_3(-L_4,-L_3,-L_2) f_1(L_5) + f_4(-L_5,-L_4,-L_3,-L_2) = \frac{2}{45}\,, \end{aligned}$$

allows one to replace the f_4 in the first line. The singular part then be-

comes

$$[E(L_1+L_2+L_3+L_4+L_5,L_6)-E(L_1,L_2+L_2+L_3+L_4+L_5+L_6)]$$
$$\times f_4(-L_5,-L_4,-L_3,-L_2)$$
$$=E^{(1)}(L_1,L_6)\times(-L_2-L_3-L_4-L_5)f_4(-L_5,-L_4,-L_3,-L_2)+\mathcal{O}((L_2+L_3+L_4+L_5)^2)\,,$$

where

$$E^{(n)}(x,y)\equiv\frac{\frac{\mathrm{d}^\mathrm{n}}{\mathrm{d}(-x)^n}(x\coth x)-\frac{\mathrm{d}^\mathrm{n}}{\mathrm{d}y^n}(y\coth y)}{x+y}\,. \tag{A.1}$$

In the limit $L_2\to-L_3-L_4-L_5$ note, using equation (3.72),

$$(-L_2-L_3-L_4-L_5)f_4(-L_5,-L_4,-L_3,-L_2)\to f_3(-L_5,-L_4,-L_3)\,.$$

This is the first of four direct limits that will be taken during this example and is the most simple; more will be said later of the general form of these expressions. For now, when $L_2\to-L_3-L_4-L_5$, it has been found that

$$\begin{aligned}g_6=&\frac{2}{45}E(L_1,L_6)+E^{(1)}(L_1,L_6)f_3(-L_5,-L_4,-L_3)\\&+[E(L_1-L_3-L_4-L_5,L_3+L_4+L_5+L_6)-E(L_1,L_6)]\,f_1(L_3+L_4+L_5)f_3(L_3,L_4,L_5)\\&+[E(L_1-L_4-L_5,L_4+L_5+L_6)-E(L_1,L_6)]\,f_2(-L_3,L_3+L_4+L_5)f_2(L_4,L_5)\\&+[E(L_1-L_5,L_5+L_6)-E(L_1,L_6)]\,f_3(-L_4,-L_3,L_3+L_4+L_5)f_1(L_5)\,.\end{aligned}$$

The next limit to consider is when $L_3 + L_4 + L_5 \to 0$. In this case both lines one and two appear singular, and again an identity should be used to rewrite one of them. The identity now is associated with a_5^{even} and states

$$f_3(L_3, L_4, L_5)+f_1(-L_3)f_2(L_4, L_5)+f_2(-L_4, -L_3)f_1(L_5)+f_3(-L_5, -L_4, -L_3) = 0\,.$$

This then can be used to replace the f_3 in the first line. One may wonder about the choice of how to use this identity; should the f_3 in the first line or the opposing f_3 in the second line be replaced? The generic answer to this is to replace the function multiplying the highest $E^{(n)}$, in order to form a simple expression to be limited. With this the singular part becomes

$$\left\{[E(L_1 - L_3 - L_4 - L_5, L_3 + L_4 + L_5 + L_6) - E(L_1, L_6)]\, f_1(L_3 + L_4 + L_5) - E^{(1)}(L_1, L_6)\right\} \times f_3(L_3, L_4, L_5)\,.$$

This is the second direct limit which shall be taken in this example and contains features that appear in all that remain. First note that Taylor expansion gives

$$E(x - y, y + z) = E(x, z) + \sum_{r=1}^{\infty} E^{(r)}(x, z)\frac{y^r}{r!}\,, \tag{A.2}$$

as the denominator of E is unchanged. This then is the reason to take the limits from outside to in as described before, as all subsequent expansions will

necessarily be of this form. In general, after n limits have been taken, the singular parts of g_N will take the form

$$[E(x-y,y+z)-E(x,z)]\,f_n(y,0,\ldots,0)-\sum_{r=1}^{n}\frac{1}{r!}E^{(r)}(x,z)f_{n-r}(y,0,\ldots,0)$$
$$=\frac{1}{(n+1)!}E^{(n+1)}(x,z)\,y+\mathcal{O}(y^2)\,. \quad \text{(A.3)}$$

This is proved using a generating function for $f_n(y,\,0,\,\ldots,\,0)$. Using this knowledge for the current example, the second limit $L_3 \to -L_4 - L_5$ may be taken to leave

$$\begin{aligned} g_6 = {} & \frac{2}{45}E(L_1,L_6)+\frac{1}{2}E^{(2)}(L_1,L_6)f_2(-L_5,-L_4) \\ & +\left\{[E(L_1-L_4-L_5,L_4+L_5+L_6)-E(L_1,L_6)]\,f_2(L_4+L_5,0)-E^{(1)}(L_1,L_6)f_1(L_4+L_5)\right\} \\ & \qquad \times f_2(L_4,L_5) \\ & +\left\{[E(L_1-L_5,L_5+L_6)-E(L_1,L_6)]\,f_3(-L_4,L_4+L_5,0)-E^{(1)}(L_1,L_6)f_2(-L_4,L_4+L_5)\right\} \\ & \qquad \times f_1(L_5)\,. \end{aligned}$$

The third limit to take is that when $L_4 + L_5 \to 0$, with lines one and two appearing singular. The approach now is hopefully becoming familiar. First use the identity, associated with a_4^{even},

$$f_2(L_4,L_5)+f_1(-L_4)f_1(L_5)+f_2(-L_5,-L_4)=\frac{1}{3}\,,$$

to replace the f_2 in the first line. The relevant term is then

$$\Big\{ \left[E(L_1 - L_4 - L_5, L_4 + L_5 + L_6) - E(L_1, L_6)\right] f_2(L_4 + L_5, 0)$$
$$-E^{(1)}(L_1, L_6) f_1(L_4+L_5) - \frac{1}{2} E^{(2)}(L_1, L_6) \Big\} f_2(L_4, L_5) \to \frac{1}{3!} E^{(3)}(L_1, L_6) f_1(-L_5)\,,$$

and hence in the limit $L_4 \to -L_5$,

$$g_6 = \frac{2}{45} E(L_1, L_6) + \frac{1}{6} E^{(2)}(L_1, L_6) + \frac{1}{3!} E^{(3)}(L_1, L_6) f_1(-L_5)$$
$$+ \Big\{ \left[E(L_1 - L_5, L_5 + L_6) - E(L_1, L_6)\right] f_3(L_5, 0, 0) - E^{(1)}(L_1, L_6) f_2(L_5, 0) - \frac{1}{2} E^{(2)}(L_1, L_6) f_1(L_5) \Big\}$$
$$\times f_1(L_5)\,.$$

The final limit in this example is when $L_5 \to 0$. Now a rather trivial identity may be used,

$$f_1(L_5) + f_1(-L_5) = 0\,,$$

to replace the f_1 in the first line. The relevant term under this limit then is

$$\Big\{ \left[E(L_1 - L_5, L_5 + L_6) - E(L_1, L_6)\right] f_3(L_5, 0, 0) - E^{(1)}(L_1, L_6) f_2(L_5, 0)$$
$$- \frac{1}{2} E^{(2)}(L_1, L_6) f_1(L_5) - \frac{1}{3!} E^{(3)}(L_1, L_6) \Big\} f_1(L_5) \to \frac{1}{4!} E^{(4)}(L_1, L_6)\,,$$

leaving

$$g_6 = \frac{2}{45}E(L_1, L_6) + \frac{1}{6}E^{(2)}(L_1, L_6) + \frac{1}{4!}E^{(4)}(L_1, L_6)\,.$$

This then concludes the example for this appendix. The lessons to draw from it are as follows. First, take sequential limits from out to in; this allows the expansion (A.2) to be used as the denominator of E is untouched. Second, using this approach all relevant terms under a limit will be of the form defined in equation (A.3). The limit then can be easily taken and the regular formula found.

More complicated situations than those discussed in this appendix can occur, for example if there are gaps in the set of variables tending to zero. Having $L_2,\, L_3 \to 0$ while simultaneously taking the limit $L_5 \to 0$ is one such example, as there is a gap between variables due to $L_4 \neq 0$. These can be dealt with in an analogous fashion to those of this appendix, but it requires more complicated identities and careful handling.

What follows is concrete and usable formulae for the first four operators, in all possible cases. The first two operators are trivial, as

$$\hat{G}_1 = 2s(L_1)\,, \qquad \hat{G}_2(L_1, L_2) = 2s(L_1 + L_2)E(L_1, L_2)\,,$$

are clearly regular. The first non-trivial example then is $\hat{G}_3$. This has six apparent singularities, of which five involve either L_1 or L_3 and thus are already

resolved. What remains then is the limit $L_2 \to 0$. In this case,

$$\hat{G}_3(L_1, 0, L_3) = 2s(L_1 + L_3)\left[\frac{2}{3} + E^{(1)}(L_1, L_3)\right],$$

where $E^{(n)}$ is defined by equation (A.1).

Next, $\hat{G}_4$ has ten apparent singularities with three of these being independent of L_1 or L_4. Explicitly these are when $L_2 \to 0$, $L_3 \to 0$, $L_2 + L_3 \to 0$. There is also a double singularity when two of these are taken simultaneously.

For the first limit, it can be found that

$$\hat{G}_4(L_1, 0, L_3, L_4) = 2s(L_1+L_2+L_3)\Bigg\{\frac{1}{3}E(L_1,\, L_3+L_4)+E^{(1)}(L_1,\, L_3+L_4)f_1(L_3) + \left[E(L_1 + L_3,\, L_4) - E(L_1,\, L_3 + L_4)\right] f_2(-L_3,\, 0)\Bigg\}.$$

The second limit can be easily found using the identity $\hat{G}_4(L_1, L_2, L_3, L_4) = \hat{G}_4(-L_4, -L_3, -L_2, -L_1)$. The third limit yields

$$\hat{G}_4(L_1,\, L_2,\, -L_2,\, L_4) = 2s(L_1 + L_4)\Bigg\{\frac{1}{3}E(L_1 + L_2,\, -L_2 + L_4) + \left[E(L_1 + L_2,\, -L_2 + L_4) - E(L_1,\, L_4)\right] f_1(-L_2)\, f_1(-L_2) - E^{(1)}(L_1,\, L_4)\, f_1(-L_2)\Bigg\}.$$

Finally, the fourth, double singularity, limit gives

$$\hat{G}_4(L_1, 0, 0, L_4) = 2s(L_1 + L_4) \left[\frac{1}{3} E(L_1,\, L_4) + \frac{1}{2} E^{(2)}(L_1,\, L_4) \right] .$$

BIBLIOGRAPHY

[1] Jordan C. Moodie, Manjinder Kainth, Matthew R. Robson, and M.W. Long. Transition temperature scaling in weakly coupled two-dimensional ising models. *Physica A: Statistical Mechanics and its Applications*, 541:123276, 2020.

[2] Jordan Moodie and Martin W. Long. An exact power series representation of the baker-campbell-hausdorff formula. *Journal of Physics A: Mathematical and Theoretical*, 2020.

[3] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations.* Springer International Publishing, 2015.

[4] Wulf Rossmann. *Lie Groups: An Introduction Through Linear Groups.* Oxford University Press, 2002.

[5] R. Achilles and A. Bonfiglioli. The early proofs of the theorem of Campbell, Baker, Hausdorff, and Dynkin. *Arch. Hist. Exact Sci.*, 66(3):295–358, 2012.

[6] Barry M. McCoy. *Advanced Statistical Mechanics.* Oxford University Press, 2010.

[7] Somendra M. Bhattacharjee and Avinash Khare. Fifty years of the exact solution of the two-dimensional ising model by onsager. *Current Science*, 69(10):816–821, 1995.

[8] W. Lenz. Beitrag zum verständnis der magnetischen erscheinungen in festen körpern. *Phys. Z*, 21:613, 1920.

[9] E. Ising. Beitrag zur theorie des ferromagnetismus. *Phys. Z*, 31:253, 1925.

[10] R. Peierls. On ising's model of ferromagnetism. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32(3):477–481, 1936.

[11] Lars Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65:117–149, Feb 1944.

[12] Bruria Kaufman. Crystal statistics. ii. partition function evaluated by spinor analysis. *Phys. Rev.*, 76:1232–1243, Oct 1949.

[13] T. D. Schultz, D. C. Mattis, and E. H. Lieb. Two-dimensional ising model as a soluble problem of many fermions. *Rev. Mod. Phys.*, 36:856–871, Jul 1964.

[14] N.M. Švrakić. Critical temperatures of ising models. *Physics Letters A*, 80(1):43 – 44, 1980.

[15] R.J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Academic Press, 1973.

[16] Tsong-Ming Liaw, Ming-Chang Huang, Yen-Liang Chou, Simon Lin, and Feng-Yin Li. Partition functions and finite-size scalings of ising model on helical tori. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 73:055101, 06 2006.

[17] David Aasen, Roger Mong, and Paul Fendley. Topological defects on the lattice i: The ising model. *Journal of Physics A: Mathematical and Theoretical*, 49, 01 2016.

[18] R. Baxter. The surface and corner free energies of the square lattice ising model. *Journal of Physics A: Mathematical and Theoretical*, 50, 06 2016.

[19] Fred Hucht. The square lattice ising model on the rectangle i: Finite systems. *Journal of Physics A: Mathematical and Theoretical*, 50, 09 2016.

[20] Fred Hucht. The square lattice ising model on the rectangle ii: Finite-size scaling limit. *Journal of Physics A: Mathematical and Theoretical*, 50, 01 2017.

[21] N. Izmailian and Chin-Kun Hu. Finite-size effects for the ising model on helical tori. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 76:041118, 11 2007.

[22] N. Izmailian. The critical ising model on a torus with a defect line. *EPL (Europhysics Letters)*, 111:60010, 01 2015.

[23] Armen Poghosyan, N. Izmailian, and Ralph Kenna. Exact solution of the critical ising model with special toroidal boundary conditions. *Physical Review E*, 96, 10 2016.

[24] W. P. Wolf. The Ising model and real magnetic materials. *Brazilian Journal of Physics*, 30:794 – 810, 12 2000.

[25] Subir Sachdev. *Quantum Phase Transitions*. Cambridge University Press, 2011.

[26] F.Y. Wu. The Potts model. *Rev. Mod. Phys.*, 54:235–268, 1982. [Erratum: Rev.Mod.Phys. 55, 315–315 (1983)].

[27] R. B. Potts. Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 48(1):106–109, 1952.

[28] M.W. Long J.C. Moodie. Statistical mechanics on the square lattice. *Unpublished*.

[29] Jutho Haegeman and Frank Verstraete. Diagonalizing transfer matrices and matrix product operators: A medley of exact and computational methods. *Annual Review of Condensed Matter Physics*, 8(1):355–406, 2017.

[30] H. A. Kramers and G. H. Wannier. Statistics of the two-dimensional ferromagnet. part i. *Phys. Rev.*, 60:252–262, Aug 1941.

[31] H. A. Kramers and G. H. Wannier. Statistics of the two-dimensional ferromagnet. part ii. *Phys. Rev.*, 60:263–276, Aug 1941.

[32] Jürg Fröhlich and Thomas Spencer. The kosterlitz-thouless transition in two-dimensional abelian spin systems and the coulomb gas. *Comm. Math. Phys.*, 81(4):527–602, 1981.

[33] Michael A. Nielsen. The fermionic canonical commutation relations and the jordan-wigner transform. *Unpublished*, 2005.

[34] N. C. Bartelt, T. L. Einstein, and L. D. Roelofs. Transfer-matrix approach to estimating coverage discontinuities and multicritical-point positions in two-dimensional lattice-gas phase diagrams. *Phys. Rev. B*, 34:1616–1623, Aug 1986.

[35] J. C. Xavier, F. C. Alcaraz, D. Penã Lara, and J. A. Plascak. Critical behavior of the spin-$\frac{3}{2}$ blume-capel model in two dimensions. *Phys. Rev. B*, 57:11575–11581, May 1998.

[36] Mehrdad Ghaemi, M. Ghannadi, and Behrouz Mirza. Calculation of the critical temperature for anisotropic two-layer ising model using the transfer matrix method. *The Journal of Physical Chemistry B*, 107, 07 2004.

[37] Tahmasb Mardani, Behrouz Mirza, and Mehrdad Ghaemi. Calculation of the shift exponent for the two-layer three-state potts model using the transfer matrix method. *Phys. Rev. E*, 72:026127, Aug 2005.

[38] Moonjung Jung and Dong-Hee Kim. First-order transitions and thermodynamic properties in the 2d blume-capel model: the transfer-matrix method revisited. *The European Physical Journal B*, 90, 08 2017.

[39] M. P. Nightingale. Scaling theory and finite systems. *Phys. A*, 83(3):561 – 572, 1976.

[40] M. P. Nightingale. Non-universality for Ising-like spin systems. *Phys. Lett. A*, 59(6):486 – 488, 1977.

[41] M. P. Nightingale and H. W. J. Blöte. Finite size scaling and critical point exponents of the Potts model. *Phys. A*, 104(1):352 – 357, 1980.

[42] D. P. Landau. Phase transitions in the Ising square lattice with next-nearest-neighbor interactions. *Phys. Rev. B*, 21:1285–1297, 1980.

[43] R. H. Swendsen. Monte Carlo renormalization-group studies of the $d = 2$ Ising model. *Phys. Rev. B*, 20:2080–2087, 1979.

[44] Robert H. Swendsen and Samuel Krinsky. Monte Carlo renormalization group and Ising models with $n \geq 2$. *Phys. Rev. Lett.*, 43:177–180, 1979.

[45] J. Oitmaa and I. G. Enting. Critical behaviour of a two-layer Ising system. *J. Phys. A*, 8(7):1097, 1975.

[46] T. Weston Capehart and Michael E. Fisher. Susceptibility scaling functions for ferromagnetic Ising films. *Phys. Rev. B*, 13:5021–5038, 1976.

[47] K. Binder and P. C. Hohenberg. Surface effects on magnetic phase transitions. *Phys. Rev. B*, 9:2194–2214, 1974.

[48] L. Angelini, D. Caroppo, M. Pellicoro, and M. Villani. The two-layer Ising film correlation length and critical curve by CLE and CVM methods. *Phys. A*, 219(3):447 – 466, 1995.

[49] James L Monroe. The bilayer Ising model and a generalized Husimi tree approximation. *Phys. A*, 335(3):563 – 576, 2004.

[50] Behrouz Mirza and T Mardani. Phenomenological renormalization group approach to the anisotropic two-layer Ising model. *J. Phys. Condens. Matter*, 34:321–324, 2003.

[51] Kazuhiko Minami and Masuo Suzuki. Non-universal critical behaviour of the two-dimensional Ising model with crossing bonds. *Phys. A*, 192(1):152 – 166, 1993.

[52] C.-Y. Weng, R. B. Griffiths, and M. E. Fisher. Critical temperatures of anisotropic Ising lattices. I. Lower bounds. *Phys. Rev.*, 162:475–479, 1967.

[53] Michael E. Fisher. Critical temperatures of anisotropic Ising lattices. II. General upper bounds. *Phys. Rev.*, 162:480–485, 1967.

[54] E. Müller-Hartmann and J. Zittartz. Interface free energy and transition temperature of the square lattice Ising antiferromagnet at finite magnetic field. *Z. Phys. B*, 27(3):261–266, 1977.

[55] Theodore W. Burkhardt. Interface free energy and critical line for the Ising model with nearest and next-nearest-neighbor interactions. *Z. Phys. B*, 29(2):129–132, 1978.

[56] J. M. J. van Leeuwen. Singularities in the critical surface and universality for Ising-like spin systems. *Phys. Rev. Lett.*, 34:1056–1058, 1975.

[57] Xiaofeng Qian and Henk W. J. Blöte. Triangular Ising model with nearest- and next-nearest-neighbor couplings in a field. *Phys. Rev. E*, 70:036112, 2004.

[58] Masuo Suzuki. Scaling with a parameter in spin systems near the critical point. I. *Prog. Theor. Phys.*, 46(4):1054–1070, 1971.

[59] Adam Lipowski and Masuo Suzuki. Exact critical temperature by mean-field approximation. *J. Phys. Soc. Jpn*, 61(12):4356–4366, 1992.

[60] Adam Lipowski and Masuo Suzuki. The layered Ising model — mean-field and interfacial approximations. *Phys. A*, 198(1):227 – 244, 1993.

[61] A Lipowski. Critical temperature in the two-layered ising model. *Phys. A*, 250(1):373 – 383, 1998.

[62] S. Vajna, K. Klobas, T. Prosen, and A. Polkovnikov. Replica resummation of the Baker-Campbell-Hausdorff series. *Phys. Rev. Lett.*, 120:200607, 2018.

[63] A.J. Silenko. General method of the relativistic Foldy-Wouthuysen transformation and proof of validity of the Foldy-Wouthuysen Hamiltonian. *Phys. Rev. A*, 91:022103, 2015.

[64] A.J. Silenko. General properties of the Foldy-Wouthuysen transformation and applicability of the corrected original Foldy-Wouthuysen method. *Phys. Rev. A*, 93:022108, 2016.

[65] A.J. Silenko. Exact form of the exponential Foldy-Wouthuysen transformation operator for an arbitrary-spin particle. *Phys. Rev. A*, 94:032104, 2016.

[66] A. Van-Brunt and M. Visser. Simplifying the Reinsch algorithm for the Baker-Campbell-Hausdorff series. *J. Math. Phys.*, 57(2):023507, 2016.

[67] A. Van-Brunt and M. Visser. Explicit Baker-Campbell-Hausdorff Expansions. *Mathematics*, 6:135, 2018.

[68] M. Matone. An algorithm for the Baker-Campbell-Hausdorff Formula. *JHEP05*, page 113, 2015.

[69] M. Matone. Closed form of the Baker-Campbell-Hausdorff formula for the generators of semisimple complex Lie algebras. *Eur. Phys. J. C*, 76:610, 2016.

[70] A. Bravetti, A. Garcia-Chung, and D. Tapias. Exact Baker-Campbell-Hausdorff formula for the contact Heisenberg algebra. *J. Phys. A*, 50:105203, 2017.

[71] D. L. Foulis. The algebra of complex 2×2 matrices and a general closed Baker-Campbell-Hausdorff formula. *J. Phys. A*, 50:305204, 2017.

[72] J. E. Campbell. On a law of combination of operators bearing on the theory of continuous transformation groups. *P. Lond. Math. Soc.*, s1-28(1):381–390, 1897.

[73] J. E. Campbell. On a law of combination of operators (second paper)*. *P. Lond. Math. Soc.*, s1-29(1):14–32, 1898.

[74] H. F. Baker. On the exponential theorem for a simply transitive continuous group, and the calculation of the finite equations from the constants of structure. *P. Lond. Math. Soc.*, s1-34(1):91–129, 1901.

[75] H. F. Baker. Further applications of metrix notation to integration problems. *P. Lond. Math. Soc.*, s1-34(1):347–360, 1902.

[76] H. F. Baker. Alternants and continuous groups. *P. Lond. Math. Soc.*, s2-3(1):24–47, 1905.

[77] F. Hausdorff. Die symbolische Exponentialformel in der Gruppentheorie. *Ber. Sächs. Ges. Wiss.*, 58:19–48, 1906.

[78] H. Poincaré. Sur les groupes continus. *Compt. Rend. Acad. Sci. Paris*, 128:1065–1069, 1899.

[79] H. Poincaré. Sur les groupes continus. *Trans. Cambridge Philos. Soc.*, 18:220–255, 1900.

[80] E.B. Dynkin. Calculation of the coefficients in the Campbell-Hausdorff formula. *Dokl. Akad. Nauk SSSR*, 57:323–326, 1947.

[81] Barry M. McCoy and Tai Tsum Wu. *The Two-Dimensional Ising Model*. Harvard University Press, 1973.