Virginia Commonwealth University

VCU Scholars Compass

2017

Application of the Fisher Dimer Model to DNA Condensation

John C. Baker III
VCU

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

Part of the Applied Statistics Commons, Biological and Chemical Physics Commons, and the Other Physics Commons

Downloaded from

https://scholarscompass.vcu.edu/etd/4791

APPLICATION OF THE FISHER DIMER MODEL TO DNA CONDENSATION

A submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

JOHN C. BAKER

B.S. Longwood University, 2009

Director: Marilyn F. Bishop and Tom McMullen,

Department of Physics

Virginia Commonwewalth University

Richmond, Virginia

May, 2017

Acknowledgements

Thank you to my advisers for putting up with me for too many years. To my family for never giving up on me. And to my friends, for helping me relax when things were overwhelming.

TABLE OF CONTENTS

LIST OF FIGURES

Abstract

APPLICATION OF THE FISHER DIMER MODEL TO DNA CONDENSATION

By John C. Baker

A submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2017.

Director: Marilyn F. Bishop and Tom McMullen,

Department of Physics

This paper considers the statistical mechanics occupation of the edge of a single helix of DNA by simple polymers. Using Fisher's exact closed form solution for dimers on a two-dimensional lattice, a one-dimensional lattice is created mathematically that is occupied by dimers, monomers, and holes. The free energy, entropy, average occupation, and total charge on the lattice are found through the usual statistical methods. The results demonstrate the charge inversion required for a DNA helix to undergo DNA condensation.

CHAPTER 1

INTRODUCTION

Investigations have shown that under the appropriate conditions DNA molecules undergo DNA condensation [1–5]. This process, wherein the DNA chain rolls into a tight toroid, can not happen spontaneously because the phosphate groups along the edges of of the helix create an overwhelming net negative charge. However, condensation will occur in a solution of polyvalent cations, which will bind to the edges of the DNA molecule, causing instances of charge inversion along the chain [1]. The charge inversion creates enough positive or neutral sites that the DNA molecule naturally compacts itself into a toroid.

Understanding the statistical mechanics of the binding of the cations to the DNA chain will give insight to the origins of complex life. This paper is intended as a first step in determining the effects geometry has on the entropy, occupation, and overall charge of the lattice. As such, the DNA strand, which is very much longer than it is wide, is considered to be an infinite one-dimensional lattice. It is to be filled by the simplest of all polymers: the dimer.

Bishop and McMullen used a similar approach in applying the lattice-gas model to the questions of DNA condensation and demonstrated that charge inversion can take place[6]. They are currently working on formulating an entropy from their findings.

However the lattice-gas model lacks any considerations of the geometry inherent in attaching a dimer flat against edge of a lattice because it considers each particle to be occupying a single lattice site (See Fig. 2). In an attempt to determine the

Fig. 1.: Lattice-gas representation of "dimers". The lattice-gas model considers all particles to occupy a single lattice site. Dimers which lie parallel to the surface should occupy two lattice sites.

Fig. 2.: Possible dimer occupation along the edge of a DNA helix. To correctly model DNA condensation the dimers which lie parallel to the surface must occupy two sites.

effects geometry has on the statistical mechanics this paper will draw upon the work done by Fisher on dimers [7]. The difference between the lattice-gas consideration of an occupied DNA strand and the more realistic representation of dimers is shown in Figs. 1 and 2.

Fisher correctly modeled how a rectangular lattice can be completely filled with dimers. His paper will be used to create a lattice of two rows which will then be mathematically cut in half between to create a model similar to Fig. 2.

This work seeks to show that charge inversion and DNA condensation will happen when the geometry of the lattice is taken into account. Additionally it is meant to serve as an introduction to applied statistical mechanics. As such much of Fisher's mathematics is carefully worked out herein. An understanding of these methods may be necessary to work on the binding of polymers which are larger than dimers to the DNA molecule.

CHAPTER 2

THE PARTITION FUNCTION AS A TRACE OF OPERATORS

Consider a plane rectangular lattice of m rows and n columns and $N = mn$ lattice sites. The lattice is populated entirely by N_x horizontal x-dimers, N_y vertical y-dimers, and N_z monomers so that every site is occupied and no sites are occupied more than once, as seen in Fig. 3. When the lattice is completely filled, called the close packed limit, it can be said that the number of particles on the lattice is

$$2N_x + 2N_y + N_z = mn \tag{2.1}$$

This is a particular treatment of the lattice gas model. All the thermodynamics of such a system can be derived from the configurational grand partition function,

$$Z_G = \mathrm{Tr}\left[e^{-\beta\left(H - \sum_\alpha \mu_\alpha N_\alpha\right)}\right], \tag{2.2}$$

where H is the Hamiltonian operator and N_α is the number operator for the species α. The μ_α parameter is the chemical potential. The species α in the system are the x-dimers, y-dimers, and z monomers.

2.1 Defining the Hamiltonian

The Hamiltonian of a system will usually have a kinetic $\frac{p^2}{2m}$ term followed by a couple of different potential energy terms. In this situation there would be a potential term for single-particle interactions and one for many-particle interactions. The work following this section will be hard enough without a convoluted Hamiltonian so only single-particle interactions are used. Coulomb repulsive and attractive forces are also
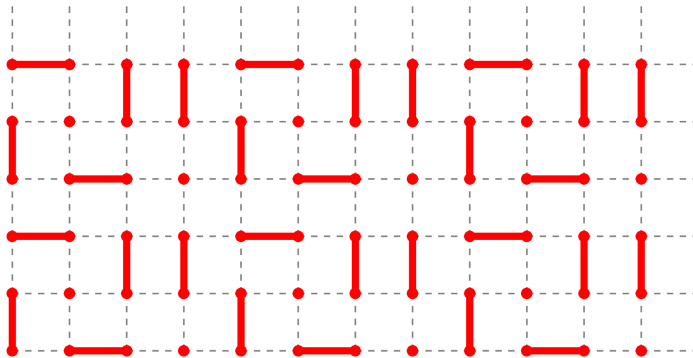
Fig. 3.: This figure shows a possible lattice configuration for x-dimers, y-dimers, and z monomers.

excluded since they are many particle interactions. It might be possible to treat them as a single particle mean field interaction, but that is outside the purview of this work.

This leaves a Hamiltonian that, when acting on a state, gives the potential energy of binding associated with that state. This potential energy should only depend on the number of particles of each species that are occupying the lattice. So for any given state Ψ

$$H \left| \Psi \right\rangle = E \left(N_x, N_y, N_z \right) \left| \Psi \right\rangle, \tag{2.3}$$

The state Ψ represents a possible arrangement of dimers and monomers on a lattice (Fig. 3). Because interaction terms are ignored, this energy should be simply

$$E \left(N_x, N_y, N_z \right) = E_x N_x + E_y N_y + E_z N_z. \tag{2.4}$$

Here E_x and E_y is the energy per dimer lying on the lattice in the x and y directions, respectively, and E_z is the energy per monomer. Based on the above equation any two states with the same number of dimers and monomers N_x, N_y, N_z, will necessarily have the same energy. This leads to many degenerate states that have different lattice configurations with the same number of respective particles.

From statistical mechanics the partition function of such a system can be written

as

$$Z_G = \text{Tr} \left\{ \exp \left[-\beta \left(E_x N_x - \mu_x N_x + E_y N_y - \mu_y N_y + E_z N_z - \mu_z N_z \right) \right] \right\}, \tag{2.5}$$

which can be made to be more descriptive by splitting the exponentials up according to the species of particles. Thus

$$Z_G = \text{Tr} \left[e^{-\beta N_x (E_x - \mu_x)} e^{-\beta N_y (E_y - \mu_y)} e^{-\beta N_z (E_z - \mu_z)} \right]. \tag{2.6}$$

2.2 Simplifying the Hamiltonion

From quantum mechanics Eq. (2.6) can be written as

$$Z_G = \sum_i \langle \Psi_i | \, e^{-N_x \beta (E_x - \mu_x)} e^{-N_y \beta (E_y - \mu_y)} e^{-N_z \beta (E_z - \mu_z)}{}_i \, | \Psi_i \rangle. \tag{2.7}$$

Here Ψ_i represents a possible pattern of placing the dimers and monomers on the lattice. Since there is no way to differentiate between configurations that have the same number of N_x, N_y, and N_z there will necessarily be some degeneracies. The degeneracy factor for a lattice with $mn = N$ lattice sites is represented by the term $g_{mn} \left(N_x, N_y, N_z \right)$.

The grand partition function Z_G can now be written as a sum over the number of particles instead of over the number of states,

$$Z_{mn} = \sum_{N_x, N_y, N_z} g_{mn} \left(N_x, N_y, N_z \right) e^{-N_x \beta (E_x - \mu_x)} e^{-N_y \beta (E_y - \mu_y)} e^{-N_z \beta (E_z - \mu_z)}, \tag{2.8}$$

where $Z_{mn} = Z_G$ is the grand partition function associated with an $m \times n$ lattice. This form of the partition function is mathematically and physically the same as Eq. (2.7).

The factor $e^{-N_x \beta (E_x - \mu_x)}$ is, thermodynamically, the activity of the x-dimers, and likewise for the y-dimers and z-monomers. For simplicity these activities will be

denoted as x, y, and z to arrive at

$$Z_{mn} = \sum_{N_x, N_y, N_z} g_{mn}\left(N_x, N_y, N_z\right) x^{N_x} y^{N_y} z^{N_z}. \tag{2.9}$$

This shows that a generating function for Z_{mn} must be found that depends only activities of the different species.

2.3 Finding the Correct Generating Function

To determine the correct generating function for monomers and dimers on a lattice it will be useful to first look at systems comprised solely of identical monomers without interactions. The case where there is one species of monomer is quite trivial. Imagine filling N lattice sites with N red balls. There are $N!$ ways of placing the balls on the lattice and since the balls are all alike, $N!$ ways of rearranging the balls that will result in the same configuration. Of course

$$\frac{N!}{N!} = 1, \tag{2.10}$$

and so there is only one way to completely fill a lattice with indistinguishable red balls.

When there are two species of monomer, say red balls and blue balls, the same starting point is used. There are $(N_b + N_r)! = N!$ ways to fill the N sites of the lattice where N_r and N_b are the numbers of red and blue balls respectively. This $N!$ is divided by $N_r!$, the number of ways to swap red balls and still have the same configuration, then by $N_b!$, the number of ways to swap blue balls without changing the configuration. Doing so gives

$$\frac{N!}{N_b!\, N_r!}, \tag{2.11}$$

6

which, since the lattice is completely filled, $N = N_r + N_b$, can be written as

$$\frac{N!}{(N - N_r)!\, N_r!}. \tag{2.12}$$

This counting expression would be the degeneracy factor $g_N(N_r, N_b)$ that would appear in the partition function that describes a lattice filled with red and blue balls. The state function for this situation is

$$Z_N = \sum_{N_r, N_b} g_N(N_r, N_b)\, r^{N_r} b^{N_b} = \sum_{N_r, N_b} \frac{N!}{N_r!\, (N - N_r)!} r^{N_r} b^{N_b}. \tag{2.13}$$

By the binomial theorem, this can be written as

$$Z_N = (b + r)^N, \tag{2.14}$$

which can be considered a generating function for red and blue balls on a lattice.

With a similar argument for three species of re, blue, and green balls, the degeneracy factor is

$$g_N(N_r, N_b, N_g) = \frac{N!}{N_r!\, N_b!\, N_g!}. \tag{2.15}$$

Then the state function is

$$Z_N = \sum_{N_r, N_b, N_g} \frac{N!}{N_r!\, N_b!\, N_g!} r^{N_r} b^{N_b} g^{N_g}, \tag{2.16}$$

where the right hand side is the definition of the multinomial theorem. This can then be expressed as

$$Z_N = (r + b + g)^N, \tag{2.17}$$

which is a generating function for three species of balls. Note that equivalently the green balls could be replaced with holes in the lattice. Additionally the power of N could be expressed as a product of N factors.

From these results it can be surmised that the proper form of Z_{mn} for a lattice
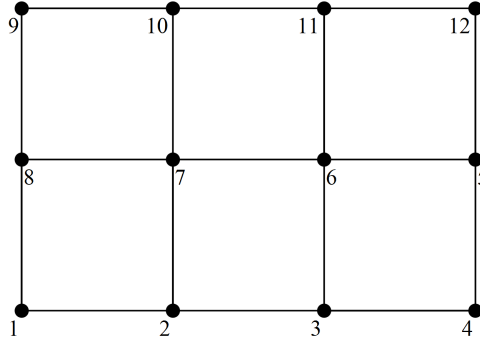
Fig. 4.: A 3×4 lattice.

filled with monomers and dimers will be a product over $N = mn$ sets of sums, where each sum is over the of species occupying the lattice points. From this point, the method for determining Z_{mn} is essentially guess and check until the function that gives the correct physical relationships is found. In the case of a lattice with mn points, each lattice site, k, can be considered to be occupied by a monomer or half of a dimer, which will be called a half-dimer. A half-dimer can be directed out from point k in any of four directions: right, left, up, or down. Later it will be required that the other half be attached to a nearest neighbor site.

An example configuration of 12 lattice sites arranged in a 3×4 lattice is shown in Fig. 4. In this figure, it would make no sense to consider connections, for example, from point 1 to points other than 2 or 8 because the length of a dimer is approximately the distance between nearest neighbors. The non-Cartesian numbering scheme of the lattice points in Fig. 4 is explained later. Thus the sum of species shall be limited to the nearest neighbors of the point k.

For any point k, the diagram in Fig. 5 will be used determine the nearest neighbors. Superimposing this diagram on the lattice in Fig. 4 shows that point 7 has the nearest neighbors $\ell_1 = 6$, $\ell_2 = 8$, $\ell_3 = 10$, and $\ell_4 = 2$.

Based on the intuition gained from this section and the previous section the best

8
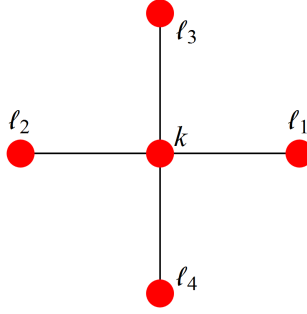
Fig. 5.: A lattice point k and the bonds to its four nearest neighbor sites.

guess for the form of a generating function for monomers and dimers on a lattice is

$$Z_{mn}^{\text{guess}} = \prod_{k=1}^{mn} \left[z_k I + x_{k,\ell_1}{}^{1/2} + x_{k,\ell_2}{}^{1/2} + y_{k,\ell_3}{}^{1/2} + y_{k,\ell_4}{}^{1/2} \right], \qquad (2.18)$$

where $x_{k\ell_1}{}^{1/2}$ indicates the half-dimer from k in the direction of the neighbor ℓ_1. As an example, the factor for $k = 2$ would be $\left[z_2 I + x_{2,3}{}^{1/2} + x_{2,1}{}^{1/2} + y_{2,7}{}^{1/2} + 0 \right]$. Note that at a lattice edge, where no bond can exist, a zero is used in place of the appropriate half-dimer in the sum.

However, upon taking the product in Eq. (2.18) there will be cross-terms that appear that make no physical sense (e.g $x_{1,2}{}^{1/2} x_{5,6}{}^{1/2}$) because dimers could not possibly connect them. These unfortunate cross terms can be eliminated by introducing a set of anticommuting operators A_i. Then any extraneous A_i operators can be eliminated by a trace operation. The correct generating function is

$$Z_{mn}(x,y,z) = t^{-1} \text{Tr} \left\{ \prod_{k=1}^{mn} \left[z_k I + x_{k,\ell_1}{}^{1/2} A_{k,\ell_1} + x_{k,\ell_2}{}^{1/2} A_{k,\ell_2} \right. \right.$$
$$\left. \left. + y_{k,\ell_3}{}^{1/2} A_{k,\ell_3} + y_{k,\ell_4}{}^{1/2} A_{k,\ell_4} \right] \right\}, \qquad (2.19)$$

as can be seen in Fisher's paper in his equation 6 [7]. Here t is the dimension of the operators A_i and of the identity matrix. The factor of t^{-1} is introduced to cancel out the t that appears when the trace is taken. The sum inside the square brackets will

9

sometimes be written as

$$Z_{mn}(x, y, z) = t^{-1}\text{Tr}\left\{ \prod_{k=1}^{mn} V_k\left(x, y, z\right) \right\}, \qquad (2.20)$$

where

$$V_k(x, y, z) = z_k I + x_{k,\ell_1}{}^{1/2} A_{k,\ell_1} + x_{k,\ell_2}{}^{1/2} A_{k,\ell_2} + y_{k,\ell_3}{}^{1/2} A_{k,\ell_3} + y_{k,\ell_4}{}^{1/2} A_{k,\ell_4}. \qquad (2.21)$$

The notation A_i is used to represent a general operator associated with bond i. Alternatively the operators can be written as $A_{k,\ell}$ where k labels one of the sites which the bond is touching and ℓ labels a nearest neighbor where ℓ can be the nearest neighbors ℓ_1 through ℓ_4 (See. Fig. 5). The $A_{k,\ell}$ notation requires that a bond, for example from point 1 to point 2, is represented by both $A_{1,2}$ and $A_{2,1}$ (Fig. 6). Since a distinct operator is needed for each bond i, a number ν is chosen so that 4ν operators or created. Here 4ν is equal to or greater than the number of bonds. Having extra operators that are not necessarily associated with bonds is not a problem.

The properties of the anticommuting A_i operators are given by

$$A_i A_j + A_j A_i = 2I\delta_{ij}. \qquad (2.22)$$

and

$$A_i^2 = I, \qquad (2.23)$$

which are the same relationships that one has for the Pauli matrices.
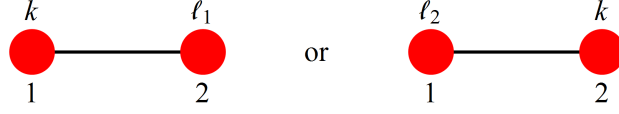
Fig. 6.: The superimposition of Fig. 5 onto a 1×2 lattice giving all of the k and their nearest neighbors.

The properties of the trace lead tto for the following relationships:

$$\text{Tr}\{A_i\} = 0, \qquad i \neq j$$

$$\text{Tr}\{A_i A_j\} = 0, \qquad i \neq j$$

$$\text{Tr}\{A_i A_j A_k\} = 0, \qquad i \neq j \neq k$$

$$\text{Tr}\{A_i^2\} = Tr\{I\} = t \quad , \tag{2.24}$$

where t is the dimension of the matrix I (see Appendix A). The trace operation is used to eliminate any term in the product that has an $A_{k,\ell}$ operator that is not paired with another $A_{k,\ell}$ which has the same subscripts. Note that $A_{1,2} = A_{1,2}$ since these are associated with the same bond. All paired operators will reduce to I and then, by the trace, to t.

This makes more sense with an example so consider a simple 1×2 lattice. As the diagram to determine nearest neighbors is moved through the lattice (Fig. 6) while taking the product the following relationships are developed.

The partition function for this case is, from Eq. (2.20),

$$Z_{mn} = t^{-1}\text{Tr}\left\{ \prod_{k=1}^{2} V_k\left(x, z\right) \right\} = t^{-1}\text{Tr}\left\{ \prod_{k=1}^{2} \left[z_k I + x_{k,\ell_1}{}^{1/2} A_{k,\ell_1} + x_{k,\ell_2}{}^{1/2} A_{k,\ell_2} \right] \right\}, \tag{2.25}$$

where $mn = 2$ and the y half-dimers have been excluded as the lattice has no vertical

11

component. Taking the product gives

$$
\begin{aligned}
Z_{mn} &= t^{-1} \mathrm{Tr} \left\{ \left(z_1 I + x_{1,2}{}^{1/2} A_{1,2} \right) \left(z_2 I + x_{2,1}{}^{1/2} A_{2,1} \right) \right\} \\
&= t^{-1} \mathrm{Tr} \left\{ z_1 z_2 I^2 + z_1 I x_{2,1}{}^{1/2} A_{2,1} + z_2 I x_{1,2}{}^{1/2} A_{1,2} + x_{1,2}{}^{1/2} A_{1,2} x_{2,1}{}^{1/2} A_{2,1} \right\},
\end{aligned}
\tag{2.26}
$$

where $x_{1,2} = x_{2,1}$ and $A_{1,2} = A_{2,1}$, since these correspond to the same bond. Then by taking the trace the equation becomes

$$
Z_{mn} = t^{-1} \left[z_1 z_2 \left(t \right) + x_{1,2} \left(t \right) \right] = z_1 z_2 + x_{1,2}.
\tag{2.27}
$$

This reduced partition function correctly shows that the only ways to completely fill this simple lattice are with either two monomers or with a single x-dimer. This procedure works for larger lattices as well. Each non-vanishing term in the partition function gives a possible lattice configuration. In general a term in the partition function is of the form

$$
(-1)^p z_{k_1} z_{k_2} ... z_{k_q} x_{k_r,\ell_r} x_{k_s,\ell_s} ... y_{k_u,\ell_u} y_{k_v,\ell_v} ... y_{k_w,\ell_w} t,
\tag{2.28}
$$

where p determines the parity. It is defined by the number of interchanges needed to move all the like operators next each other so that they may reduce to I. Each interchange creates a single negative sign.

Introducing the anticommuting operators works to eliminate unwanted cross terms but has the unfortunate side effect that they might produce negative signs in the terms in the partition function. To use this form of the partition function, even parity is required in each term in the trace since none of the terms can be negative. If the parity can be proved to be always even then it can be said that Eq. (2.20) is the correct partition function and will enumerate every possible combination of

monomers and dimers on the lattice.

CHAPTER 3

ENSURING EVEN PARITY

3.1 The Parity Problem

The parity of the expansion is determined by the order the factors appear in the product since it is based on the interchange of $A_{k,\ell}$ operators. This in turn depends on the order that the lattice points are numbered. Under normal circumstances one would usually number the lattice sites according to their Cartesian coordinates (r, s). In Cartesian the first point would be at $(0,0)$ and r and s would increase as $r = 0, 1, 2, ..., n - 1$ and $s = 0, 1, 2, ...m - 1$. The formula for lattice point k would then be

$$k = sn + r + 1, \tag{3.1}$$

which results in the common numbering scheme where each row is labeled in increasing order from left to right and bottom to top.

Some busy work with small lattices will show that this more obvious scheme of numbering lattice points will not resolve the parity problem. Some signs remain negative. Consider the 4×4 matrix in Fig. 7. From Eq. (2.20) the partition function for this case is
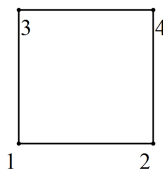
Fig. 7.: A simple square lattice with four points. Each row is numbered in the Cartesian style, left to right.

$$Z_{2,2}(x, y, z) =$$

$$t^{-1}\text{Tr}\left\{ \prod_{k=1}^{4} \left[z_k I + x_{k,\ell_1}{}^{1/2} A_{k,\ell_1} + x_{k,\ell_2}{}^{1/2} A_{k,\ell_2} + y_{k,\ell_3}{}^{1/2} A_{k,\ell_3} + y_{k,\ell_4}{}^{1/2} A_{k,\ell_4} \right] \right\},$$

$$(3.2)$$

which expands as

$$Z_{2,2}(x, y, z) =$$

$$t^{-1}\text{Tr}\left\{ \left(z_1 I + x_{1,\ell_1}{}^{1/2} A_{1,\ell_1} + x_{1,\ell_2}{}^{1/2} A_{1,\ell_2} + y_{1,\ell_3}{}^{1/2} A_{1,\ell_3} + y_{1,\ell_4}{}^{1/2} A_{1,\ell_1} \right) \right.$$

$$\left(z_2 I + x_{2,\ell_1}{}^{1/2} A_{2,\ell_1} + x_{2,\ell_2}{}^{1/2} A_{2,\ell_2} + y_{2,\ell_3}{}^{1/2} A_{2,\ell_3} + y_{2,\ell_4}{}^{1/2} A_{2,\ell_1} \right)$$

$$\left(z_3 I + x_{3,\ell_1}{}^{1/2} A_{3,\ell_1} + x_{3,\ell_2}{}^{1/2} A_{3,\ell_2} + y_{3,\ell_3}{}^{1/2} A_{3,\ell_3} + y_{3,\ell_4}{}^{1/2} A_{3,\ell_1} \right)$$

$$\left. \left(z_4 I + x_{4,\ell_1}{}^{1/2} A_{4,\ell_1} + x_{4,\ell_2}{}^{1/2} A_{4,\ell_2} + y_{4,\ell_3}{}^{1/2} A_{4,\ell_3} + y_{4,\ell_4}{}^{1/2} A_{4,\ell_1} \right) \right\}.$$

$$(3.3)$$

Then the diagram in Fig. 5 is superimposed over each point in the lattice to determine the values of ℓ_1, ℓ_2, ℓ_3, and ℓ_4. It is important to note that if the points outside of the 2×2 lattice had been labelled they would be naturally eliminated by not having paired operators in the trace. For simplicity, however, terms which have subscripts that would be referencing points outside the lattice are simply set to zero.

The result of applying Fig. 5 to the 2×2 matrix is

$$Z_{2,2}(x, y, z) = t^{-1}\text{Tr}\Bigg\{ \left(z_1 I + x_{1,2}{}^{1/2}A_{1,2} + y_{1,3}{}^{1/2}A_{1,3}\right)$$

$$\left(z_2 I + x_{2,1}{}^{1/2}A_{2,1} + y_{2,4}{}^{1/2}A_{2,4}\right)$$

$$\left(z_3 I + x_{3,4}{}^{1/2}A_{3,4} + y_{3,1}{}^{1/2}A_{3,1}\right)$$

$$\left(z_4 I + x_{4,3}{}^{1/2}A_{4,3} + y_{4,2}{}^{1/2}A_{4,2}\right) \Bigg\}. \tag{3.4}$$

Through the anticommutation laws described in Eq. (2.22) the $A_{k\ell}$ operators will eliminate all terms in which the operators are unpaired once the trace is taken. In the next equation they have been canceled somewhat prematurely to make this product small enough for demonstration purposes.

$$Z_{2,2}(x, y, z) = t^{-1}\text{Tr}\Bigg\{ \left(z_1 z_2 z_3 z_4 I^4\right)$$

$$+ \left(z_1 z_2 I^2 x_{3,4}{}^{1/2}A_{3,4}x_{4,3}{}^{1/2}A_{4,3}\right)$$

$$+ \left(z_3 z_4 I^2 x_{1,2}{}^{1/2}A_{1,2}x_{2,1}{}^{1/2}A_{2,1}\right)$$

$$+ \left(z_1 z_3 I^2 y_{2,4}{}^{1/2}A_{2,4}y_{4,2}{}^{1/2}A_{4,2}\right) \tag{3.5}$$

$$+ \left(z_2 z_4 I^2 y_{1,3}{}^{1/2}A_{1,3}y_{3,1}{}^{1/2}A_{3,1}\right)$$

$$+ \left(x_{1,2}{}^{1/2}A_{1,2}x_{2,1}{}^{1/2}A_{2,1}x_{3,4}{}^{1/2}A_{3,4}x_{4,3}{}^{1/2}A_{4,3}\right)$$

$$+ \left(y_{1,3}{}^{1/2}A_{1,3}y_{2,4}{}^{1/2}A_{2,4}y_{3,1}{}^{1/2}A_{3,1}y_{4,2}{}^{1/2}A_{4,2}\right) \Bigg\}.$$

When doing these expansions, care must be taken to keep all the operators in order. The operators can move through the x, y, and z factors with no problems, and of course the identity operator can move anywhere, but moving an anticommuting operator past an unlike anticommuting operator will create a negative sign. With everything rearranged except the anticommuting operators, and with the powers of

16

1/2 combined this is

$$Z_{2,2}(x, y, z) = t^{-1}\text{Tr}\Bigg\{ \left(z_1 z_2 z_3 z_4 I^4\right)$$

$$+ \left(z_1 z_2 I^2 x_{3,4} A_{3,4} A_{4,3}\right)$$

$$+ \left(z_3 z_4 I^2 x_{1,2} A_{1,2} A_{2,1}\right)$$

$$+ \left(z_1 z_3 I^2 y_{2,4} A_{2,4} A_{4,2}\right) \qquad (3.6)$$

$$+ \left(z_2 z_4 I^2 y_{1,3} A_{1,3} A_{3,1}\right)$$

$$+ \left(x_{1,2} A_{1,2} A_{2,1} x_{3,4} A_{3,4} A_{4,3}\right)$$

$$+ \left(y_{1,3} y_{2,4} A_{1,3} A_{2,4} A_{3,1} A_{4,2}\right) \Bigg\}.$$

The only term in which the operators need to be rearranged is the last term. All the other terms have the operators paired. Substituting $A_{2,4} A_{3,1} = -A_{3,1} A_{2,4}$ in the last term gives

$$Z_{2,2}(x, y, z) = t^{-1}\text{Tr}\Bigg\{ \left(z_1 z_2 z_3 z_4 I^4\right)$$

$$+ \left(z_1 z_2 I^2 x_{3,4} A_{3,4} A_{4,3}\right)$$

$$+ \left(z_3 z_4 I^2 x_{1,2} A_{1,2} A_{2,1}\right)$$

$$+ \left(z_1 z_3 I^2 y_{2,4} A_{2,4} A_{4,2}\right) \qquad (3.7)$$

$$+ \left(z_2 z_4 I^2 y_{1,3} A_{1,3} A_{3,1}\right)$$

$$+ \left(x_{1,2} A_{1,2} A_{2,1} x_{3,4} A_{3,4} A_{4,3}\right)$$

$$- \left(y_{1,3} y_{2,4} A_{1,3} A_{3,1} A_{2,4} A_{4,2}\right) \Bigg\},$$

which creates a negative sign. The partition function can be reduced through the

properties of the anticommuting operators to produce identity operators.

$$Z_{2,2}(x,y,z) = t^{-1}\text{Tr}\Big\{ (z_1 z_2 z_3 z_4 I) + (z_1 z_2 x_{3,4} I)$$
$$+ (z_3 z_4 x_{1,2} I) + (z_1 z_3 y_{2,4} I)$$
$$+ (z_2 z_4 y_{1,3} I) + (x_{1,2} A_{1,2} I)$$
$$- (y_{1,3} y_{2,4} I) \Big\}, \tag{3.8}$$

and because the trace of I is t, the order of the matrix, and the trace is distributive

$$Z_{2,2}(x,y,z) = (z_1 z_2 z_3 z_4) + (z_1 z_2 x_{3,4})$$
$$+ (z_3 z_4 x_{1,2}) + (z_1 z_3 y_{2,4})$$
$$+ (z_2 z_4 y_{1,3}) + (x_{1,2} x_{3,4})$$
$$- (y_{1,3} y_{2,4}) . \tag{3.9}$$

Each term here represents a possible configuration of monomers and dimers. For example the fourth time says that the lattice can be filled by a monomers at points 1 and 3 and a y-dimer connecting points 2 and 4.

Unfornately, the final term has a negative sign in front of it, making this trace unusable as a partition function. Kesteleyn solved this problem by affixing all y-dimers lying on odd columns of the lattice with a negative sign [8]. Temperley noted, in addition to Kasteleyn's method, that one could take every other y-dimer to be of opposite sign [9]. Fisher chose the more rigorous approach of proving that for a lattice occupied only by dimers the parity must be even, and every term will be positive [7]. The method he used may be useful in expanding from dimers to longer polymers, and so it is included here.

To move forward, an analogy is made to the one dimensional problem. A one dimensional lattice can be occupied only by x-dimers and z monomers. If it were
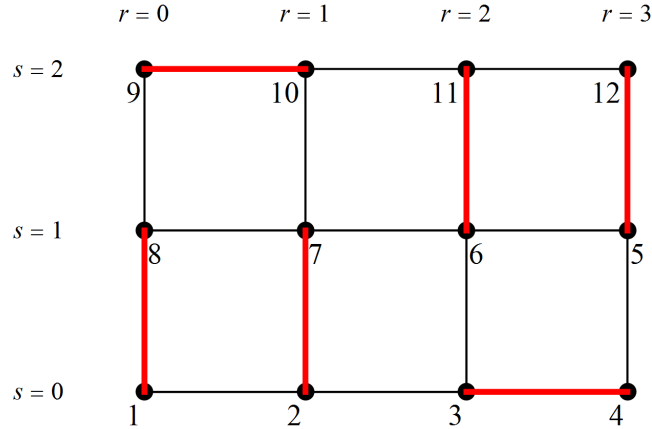
Fig. 8.: A possible lattice configuration for a 3 × 4 lattice numbered in a zig-zag pattern.

occupied only by x-dimers, then it is trivial to prove that the parity is even since every dimer occupies an even number of sites. Since y-dimers also occupy an even number of sites, it seems a promising avenue to number the lattice points so that the lattice can be "unfolded" into a one dimensional chain of lattice points.

What follows in this chapter is a proof that with a new numbering system even parity can be assured for every lattice that is completely occupied only by dimers, both in the x and y directions. Any monomers or holes may ruin the sign by the fact that they are not enumerated with anticommuting operators. Unfortunately, for the sake of parity, monomers must be excluded as a lattice element for now. This solution is currently valid only when the lattice is completely filled with dimers.

3.2   A New Numbering System

To make a lattice that can be unfolded into a chain, every other row needs to be numbered in the opposite direction as the previous row. Consider one configuration of an example lattice that has every site occupied by a dimer, shown in Fig. 8. Here the red lines are dimers. Note that every site is occupied by a dimer and no site can have two dimers attached to it. This is called the close-packed limit $z = 0$.
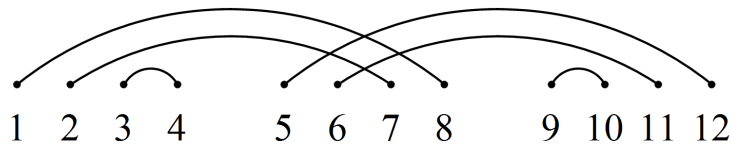
Fig. 9.: The one-dimensional array corresponding to unfolding the 3×4 lattice in Fig. 8

into a one-dimensional chain.

The zig-zag numbering of points in Fig. 8 ensures that each lattice point has one nearest neighbor that is one integer higher than it is. The lattice itself can now be viewed as a chain which is bent to zig-zag its way upwards. In terms of the Cartesian coordinates (r, s), shown in Fig. 8, if the lattice is placed so that the bottom left corner is on $(0, 0)$, then new definition of the $k(r, s)$ for an $m \times n$ lattice is

$$k(r, s) = \begin{cases} sn + r + 1, & s \text{ even} \\ \\ sn + n - r, & s \text{ odd} \end{cases}, \tag{3.10}$$

where $r = 0, 1, 2, ..., n - 2, n - 1$ and $s = 0, 1, 2, ..., m - 2, m - 1$ are the Cartesian coordinates and n is the number of columns and m is the number of rows. This numbering procedure causes all nonzero terms in the expansion of Eq. (2.20) to be positive when like operators are paired. It is easy to check this for small lattices, but a more detailed proof is required to continue.

For each term corresponding to given lattice configuration there will be N operators in that term. This N must be equal to the number of lattice sites because every bond that occupies two sites is indicated by two like operators. (e.g $A_{1,8} A_{8,1}$). This can be seen in the terms of Eq. (3.7) that involve only dimers. These operators and the bonds between them can be represented with a linear array of points, as seen in Fig. 9 which is a representation of the lattice configuration of Fig. 8. This linear array is analogous to "unfolding" the lattice into a linear chain.

20

In Fig. 9, the arcs above the array connect points in the lattice in Fig. 8 that are connected by dimers. These arcs are used to represent operators in the product which share the same subscripts. Each of the points can be imagined as representing the $A_{k,\ell}$ operators. In this notation, k is the number of the current point and ℓ is the number of the point to which it is connected. For example, the operator for the first point in the array is the operator $A_{1,8}$ since it is point 1 and the arc connects it to point 8. Naturally the operator at point 8 is $A_{8,1}$, which represents the other half of the dimer that connects the points in Fig. 8. So long as the product in Eq. (2.20) is taken carefully and in ascending order from 1 to mn, the order that the operators appear in the terms of the trace will correspond to the order they appear in the array in Fig. 9.

The gaps between sets of numbers are placeholders to indicate where, on the original lattice, one row gives way to the next. The x-dimers from the original lattice in Fig. 8 create arcs within rows. They will be called x-arcs. Similarly, y-dimers create y-arcs, and they traverse the gaps between the sets of rows. In the limit where z=0, every point on the lattice must be the end point of an arc.

It is important to note that some x-arcs or y-arcs that may seem possible, such as a y-arc from point 1 to point 5 (called $y(1,5)$) are precluded by the fact that the lattice in Fig. 8 used to create the array representation can have no dimer connecting point 1 to point 5. One must make sure that any array that is drawn represents a real lattice configuration.

From Fig. 9, it can be also seen that the arcs may intersect each other. The total number of intersections will be defined as f. It might be possible to stretch arcs around to change f. For example the arc from point 3 to 4 ($x(3,4)$) could be stretched way up above $y(2,7)$ and $y(1,8)$. This stretching would necessarily always create or eliminate an even number of intersections and will not change the parity of
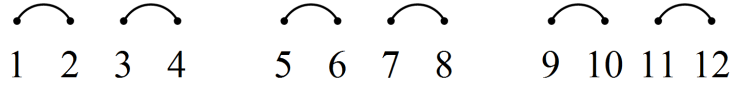
Fig. 10.: The trivial completely paired array where the number of intersections, f, is zero.

f. With this in mind it is best to always draw the arcs to minimize the number of intersections.

Taking the trace requires that all of the operators with the same subscripts are next to each other in the terms in the product of Eq. (3.7). Once they are paired, they can be reduce to the identity operator and survive the trace operation. Interchanging two unlike operators in the product corresponds to a similar interchange of points in our linear array. A change in points involving unlike operators will change the parity of f by creating or destroying a single intersection. This can be seen in the interchange of points 6 and 7. Swapping like operators that are already paired does nothing in the product and nothing in the array. An exchange of points 3 and 4 does not change the number of intersections in Fig. 9. Thus f, the array parity, and p, the parity of a term in the product, are related to each other. If one changes so does the other.

It is simple to determine that, in addition to being linked, the two parities must also have the same same sign. If all the operators in the array are paired to begin with then f is zero and the parity is even. Since this requires no interchange of points in the array to reorder the like operators we can conclude that all the similar operators are already paired with their neighbors. Therefore p is even in the product term corresponding to that array. So if f is even so too will p be even.

When talking about interchanging points in the array things can become rather confusing. It is imperative to make sure that the initial array is set up correctly when compared to the lattice or product term it corresponds to. After that consider the

22
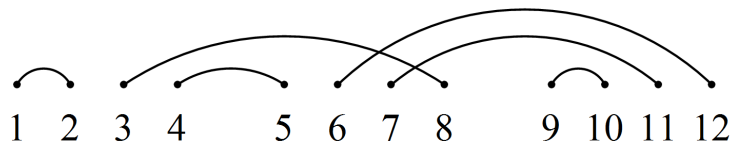
Fig. 11.: The first intermediate step of pairing off the lattice in Fig. 9 interchanging points until the arc representing the operator pair $A_{1,8}A_{8,1}$ spans array points 1 and 2.

array to be its own entity. When an interchange of points is made it is literally only an interchange of the points on the array, not the numbers labelling the array. With this in mind the arc going from label 1 to label 2 in the array might not represent $A_{1,2}A_{2,1}$ after interchanges are made. As an example the first step in pairing all the operators in Fig. 9 might be to move the point at the array label 8 all the way to the left next to array label 1 since the arc shows that those two points are like operators and should be paired off. This takes six interchanges of points to achieve and the result is show in Fig. 11. Note that this intermediate step to pairing off all of the operators is an array that does not represent a valid lattice dimer configuration because a dimer connection from 3 to 8 in Fig. 8 is not possible. This is common and not a problem so long as the array is eventually put back into a valid configuration.

Since an interchange of points will either create or destroy a single intersection (unless the points are already paired off by being next to each other) the parity of f will change once for each interchange. Therefore any array that can be created from the completely paired array via an even number of interchanges must have an even parity and vice versa. Therefore to prove that p is always even for the close packed limit where $z = 0$ we must only prove that f is always even for every array configuration created by a valid lattice or product term in the trace.
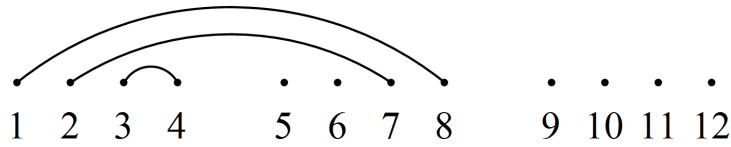
Fig. 12.: An array with arcs drawn from the $s = 0$ row to the $s = 1$ row of Fig. 9.
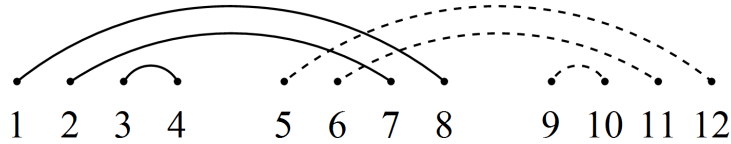
Fig. 13.: The addition of arcs from the $s = 1$ row to the $s = 2$ row causes intersections

3.3   Proving f and p are Even

To prove that f is always even the number of intersections of intersections that the arcs make above the array must be considered. The array in Fig. 9 will be used to determine the behavior of the arcs before generalizing to an array of arbitrary size. It should be noted that x-arcs can create no intersections as can be seen in Fig. 9.

In Fig. 12 the arcs from the $s = 0$ to $s = 1$ rows are shown. As is quickly evident, these arcs can create no intersections on their own. When arcs are added from the $s = 1$ row to the $s = 2$ row it can be seen in Fig 13 that new y-arcs will intersect each of the previously drawn arcs at most once. As before, no intersections are created by x-arcs.

From Fig. 13 it can be said that intersections are formed when arcs going from row $s = 0$ to row $s = 1$ cross over those y-arcs from $s = 1$ to $s = 2$ and that intersections are only formed when one of the arcs has an endpoint which lies between the ends of a different arc. For example the arc from 5 to 6 has an endpoint which is between the endpoints of the arc from 1 to 8. This is the condition under which intersections are created.

In general intersections are created when arcs going from row $s - 1$ to row s cross

24

Fig. 14.: The arbitrary row $2u$ and its neighboring rows.
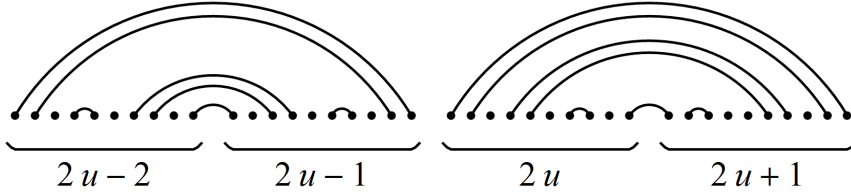
Fig. 15.: Arcs drawn for pairs of rows where the even row is the lower of the pair.

over those y-arcs from s to $s+1$. To generalize this to a lattice and array of arbitrary size consider the row $s = 2u$ in the arbitrarily large array of points show in Fig. 14.

The row $2u$ is chosen because the following explanation of how to generate this kind of array relies on pairing adjacent rows. It will be useful to know which of the rows in the pair is even and which is odd. To continue the rows in the array are paired off so that every even row is the lowest row in the pairs of adjacent rows. Then the appropriate arcs are added to connect the like operators in the pairs of rows as seen in Fig. 15. Adding these arcs to the array will create no intersections at all in the array just as none were created in Fig. 12.

Next the rows are re-paired so that the odd rows are the lowest in the pair. New dashed arcs are then added. These will be the arcs that create intersections. The intersections are highlighted in red in Fig. 16.

It was noted about Fig. 12 that when the array is drawn to minimize the number of intersections that the most a single y-arc can intersect another single y-arc is once. Assume all of the solid arcs are already drawn on the array. If it can be proven that no dashed arcs, when added to the array, are intersected by an odd number of other arcs, then f must be even. This will first be demonstrated using a single dimer on a

25
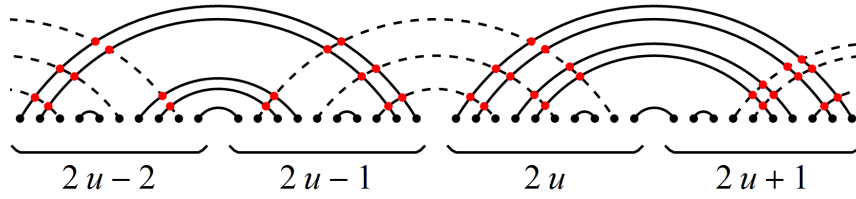
Fig. 16.: The addition of arcs for pairs of rows where the lower row is odd. The intersection points are colored red.
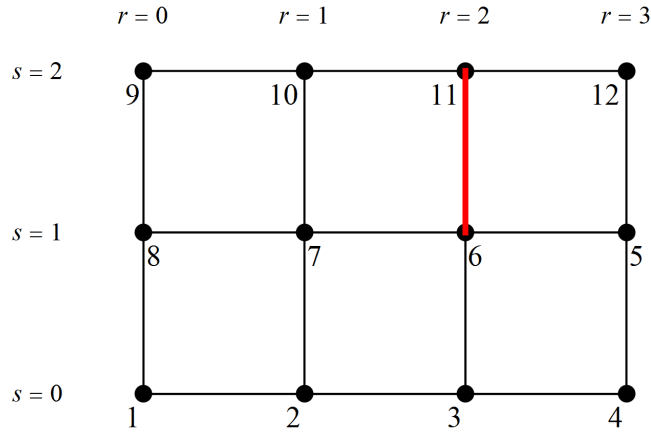
Fig. 17.: A lattice with a single dimer for demonstration purposes.

simple lattice (Fig. 17).

The dimer on the demonstration lattice goes from an s=odd row to an s=even row just as the dashed arcs did when added to the arbitrary array near the point $2u$. When this lattice configuration is unfolded into an array it has the form seen in Fig. 18.

Because intersections could only be formed in this example by arcs that have a single endpoint between the points 6 and 11 on the array, the key to ensuring even parity rests on showing that the number of these y-arcs is even.

Fig. 18.: An array with only one dimer.

When considering the lattice configuration note that a vertical dimer always has an even number of points between its start point and end point when taken in counting order. This is because the two points are both the same distance from the left hand edge of the lattice. The array, then, will have an even number of points between one end of a y-arc and the other end. This idea can be generalize to the array of arbitrary size. Recall the numbering scheme for lattice points from the previous section.

$$
k(r,s) = \begin{cases} sn + r + 1, & \text{s even} \\ sn + n - r, & \text{s odd} \end{cases}, \tag{3.11}
$$

With this it can be seen that points in the row $2u$ will be

$$
k(r, 2u) = 2un + r + 1. \tag{3.12}
$$

Those in row $2u - 1$ are defined as

$$
k(r, 2u) = 2un + r + 1. \tag{3.13}
$$

The points in $2u + 1$ are

$$
k(r, 2u) = 2un + r + 1. \tag{3.14}
$$

Consider an arbitrary $y - arc$ from row $2u - 1$ to row $2u$. This arc is necessarily going from the point $2un - r$ to the point $2un + r + 1$. This can be seen in lattice notation in Fig. 19.

The number of intermediate points spanned by the arc in the array is equal to twice the distance the dimer in Fig. 19 is from the edge. Thus there are $2r$ intermediate points, an even number, between the beginning and the end of any arbitrary y-arc from row $2u - 1$ to row $2u$. Each of these $2r$ points must be an endpoint of an arc because the array is completely filled with dimers.
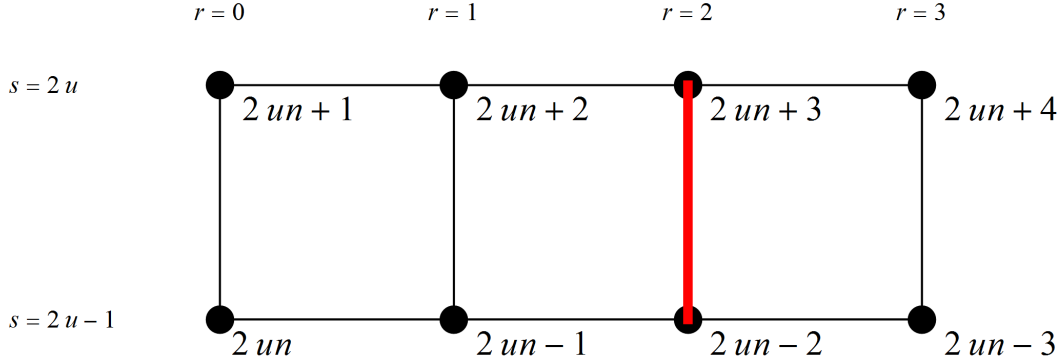
Fig. 19.: A single dimer connecting two points in an arbitrarily sized array.

There are three different species of arcs. There are x-arcs which occupy two points in a single row and create no intersections.

The second species are "internal" y-arcs. These are arcs that span the same pair of rows as the arbitrary y-arc. These arcs will arc either over or under the arc which is being considered. If they arc under they take up two points between the endpoints of the arbitrary arc. If they arc over then they take up no intermediate points points. They will never intersect the arbitrary y-arc which is under consideration. These are the other dotted lines which span two rows in Fig. 16.

The last species are "external" y-arcs. These arcs have one endpoint which lies on the pair of rows and one end point which is in an adjacent row, either $2u - 2$ or $2u + 1$. These are the only species that will create intersections. They will create an intersection when their sole endpoint in rows $2u - 1$ or $2u$ lies in the $2r$ intermediate points bounded by the endpoints of our arbitrary arc. These are solid lines in Fig. 16.

Since the arbitrary y-arc can intersect each of these external arcs only once the proof relies on there being an even number of external arcs that have an endpoint on the $2r$ intermediate points. Because the lattice must be completely filled this can be shown by subtracting all the other possible arcs that could occupy these $2r$ points.

The other two species, x-arcs and internal y-arcs, both occupy even numbers of

28

points. Only those which appear under the arbitrary y-arc need to be considered. Each of these w arcs made up of the combined number of x-arcs and internal y-arcs will occupy $2w$ points. Since $2r - 2w$ is necessarily an even number the remaining number of points is even. Each of these remaining points must be occupied by an external y-arc which passes through an arbitrary y-arc only once.

An even number of y-arcs each creating single intersections with other y-arcs creates an even number of intersections. Therefore f, and by extension, p, are even and Eq. (2.20) is an appropriate state function for the system of dimers on a plane lattice.

Note again that this solution is only valid for the limit in which $z = 0$ and the lattice is completely filled.

CHAPTER 4

THE PFAFFIAN

The partition function for this system can be expressed as the trace of a product of operators. Hurst and Green pointed out that any such trace of homogeneous operators of this form can be reduced to a Pfaffian [10]. A Pfaffian is a type of determinant that is taken on the upper right hand triangle of an antisymmetric matrix[11]. One of the most important properties of a Pfaffian is that its square is equal to the determinant of the corresponding antisymmetric matrix.

$$D = P^2 = \begin{vmatrix} 0 & a_{1,2} & a_{1,3} & a_{1,4} & \ldots & a_{1,2h} \\ -a_{1,2} & 0 & a_{2,3} & a_{2,4} & \ldots & a_{2,2h} \\ -a_{1,3} & -a_{2,3} & 0 & a_{3,4} & \ldots & a_{3,2h} \\ -a_{1,4} & -a_{2,4} & -a_{3,4} & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & a_{2h-1,2h} \\ -a_{1,2h} & -a_{2,2h} & -a_{3,2h} & \ldots & a_{2h-1,2h} & 0 \end{vmatrix} \tag{4.1}$$

The Pfaffian itself is written as $P = \backslash a_{r,s}|$. A Pfaffian of order 2h is a triangular array of h(2h-1) elements.

$$P = \begin{vmatrix} a_{1,2} & a_{1,3} & a_{1,4} & \ldots & a_{1,2h} \\ & a_{2,3} & a_{2,4} & \ldots & a_{2,2h} \\ & & a_{3,4} & \ldots & a_{3,2h} \\ & & & \ddots & \vdots \\ & & & & a_{2h-1,2h} \end{vmatrix} . \tag{4.2}$$

This is the upper right hand triangle of the determinant in Eq. (4.1). The Pfaffian expansion by the first row is just like a regular determinant with the notable caveat that both the rth row and column and the sth row and column are eliminated. Some basic Pfaffian manipulation follows as an introduction.

Consider a Pfaffian of order $2h = 6$,

$$P = \begin{vmatrix} a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} & a_{1,6} \\ & a_{2,3} & a_{2,4} & a_{2,5} & a_{2,6} \\ & & a_{3,4} & a_{3,5} & a_{3,6} \\ & & & a_{4,5} & a_{4,6} \\ & & & & a_{5,6} \end{vmatrix}. \tag{4.3}$$

To expand by this by the first row, the element $a_{r,s} = a_{1,2}$ is chosen and the rth row and column and the sth row and column are eliminated. Note the the columns and rows in the Pfaffian are labelled according to the numbering of the columns and rows of the corresponding antisymmetrix matrix so that the element $a_{1,2}$ is actually in the second column, $s = 2$. Elimination in this manner creates a minor which is then multipled by $a_{1,2}$ to give the term

$$a_{1,2} \times \begin{vmatrix} a_{3,4} & a_{3,5} & a_{3,6} \\ & a_{4,5} & a_{4,6} \\ & & a_{5,6} \end{vmatrix}. \tag{4.4}$$

Every element with $r, s = 1$ and $r, s = 2$ in the original Pfaffian has been eliminated to create the resulting minor.

In continuing the expansion recall that in a regular matrix determinant every other term gets a negative sign. The Pfaffian has the same alternating sign in its

31

expansion. The full expansion of first row terms with their minors is

$$P = a_{1,2} \times \begin{vmatrix} a_{3,4} & a_{3,5} & a_{3,6} \\ & a_{4,5} & a_{4,6} \\ & & a_{5,6} \end{vmatrix} - a_{1,3} \times \begin{vmatrix} a_{2,4} & a_{2,5} & a_{2,6} \\ & a_{4,5} & a_{4,6} \\ & & a_{5,6} \end{vmatrix} + a_{1,4} \times \begin{vmatrix} a_{2,3} & a_{2,5} & a_{2,6} \\ & a_{3,5} & a_{3,6} \\ & & a_{5,6} \end{vmatrix}$$

$$- a_{1,5} \times \begin{vmatrix} a_{2,3} & a_{2,4} & a_{2,6} \\ & a_{3,4} & a_{3,6} \\ & & a_{4,6} \end{vmatrix} + a_{1,6} \times \begin{vmatrix} a_{2,3} & a_{2,4} & a_{2,5} \\ & a_{3,4} & a_{3,5} \\ & & a_{4,5} \end{vmatrix} \tag{4.5}$$

It can be more easily seen in the full expansion that both the rth row and column and the sth row and column are eliminated to create each minor.

To complete the Pfaffian each of these minors is expanded by their first rows yielding

$$P = a_{1,2}[a_{3,4}a_{5,6} - a_{3,5}a_{4,6} + a_{3,6}a_{4,5}] - a_{1,3}[a_{2,4}a_{5,6} - a_{2,5}a_{4,6} + a_{2,6}a_{4,5}]$$

$$+ a_{1,4}[a_{2,3}a_{5,6} - a_{2,5}a_{3,6} + a_{2,6}a_{3,5}] - a_{1,5}[a_{2,3}a_{4,6} - a_{2,4}a_{3,6} + a_{2,6}a_{3,4}] \tag{4.6}$$

$$+ a_{1,6}[a_{2,3}a_{4,5} - a_{2,4}a_{3,5} + a_{2,5}a_{3,4}]$$

Simply distributing gives

$$P = a_{1,2}a_{3,4}a_{5,6} - a_{1,2}a_{3,5}a_{4,6} + a_{1,2}a_{3,6}a_{4,5} - a_{1,3}a_{2,4}a_{5,6} + a_{1,3}a_{2,5}a_{4,6} - a_{1,3}a_{2,6}a_{4,5}$$

$$+ a_{1,4}a_{2,3}a_{5,6} - a_{1,4}a_{2,5}a_{3,6} + a_{1,4}a_{2,6}a_{3,5} - a_{1,5}a_{2,3}a_{4,6} + a_{1,5}a_{2,4}a_{3,6} - a_{1,5}a_{2,6}a_{3,4}$$

$$+ a_{1,6}a_{2,3}a_{4,5} - a_{1,6}a_{2,4}a_{3,5} + a_{1,6}a_{2,5}a_{3,4}$$

$$\tag{4.7}$$

This is the full expansion of the Pfaffian. Squaring it will give the determinant of the original matrix.

The Pfaffian has been historically useful because it acts on only half of the matrix. This coupled with the double elimination of rows and columns in the expansion

ensures that each matrix element appears at most once in each term of the full expansion. It is, then, an excellent tool for counting and will save on some of the more complicated manipulations that were required by Kac and Ward [12].

4.1   The Pfaffian as a Sum

There is a way to determine the Pfaffian without having to write down matrices and determing their minors. For a Pfaffian of order 2h, the formula is

$$P = \sum_{permutations} (-1)^p (a_{\ell_1,\ell_2})(a_{\ell_3,\ell_4})...(a_{\ell_{2h-1},\ell_{2h}}) \tag{4.8}$$

The sum is over the $(2h-1)!! = (2h-1)(2h-3) \times ... \times 5 \times 3 \times 1$ permutations that satisfy the following two relationships:

$$\ell_1 < \ell_2, \quad \ell_3 < \ell_4, \quad ..., \quad \ell_2h-1 < \ell_2h, \tag{4.9}$$

and

$$\ell_1 < \ell_3 < \ell_5 < ... < \ell_2h - 1, \tag{4.10}$$

where $\ell_1 \neq \ell_2 \neq \ell_3... \neq \ell_2h$.

The parity, $(-1)^p$, is found by examining the order of the subscripts. If they are imagined to be in list form then the number of interchanges to get them into increasing numerical order directly determines whether p is even or odd. As a proof of concept the Pfaffian of order $2h = 6$ from Eq. (4.3) is used.

To do the sum in Eq. (4.8) ℓ_1 is set to equal 1. Note that because of the inequalities in Eqs. (4.9) and (4.10) that ℓ_1 can never be greater than 1. The simplest place to start working out the Pfaffian expansion is with the term in which all the subscripts are in counting order. The term is $a_{1,2}a_{3,4}a_{5,6}$. To find all of the terms that begin with $a_{1,2}$ the other subscripts need to be swapped around without violating the

inequalites in Eqs. (4.9) and (4.10). The second term is $-a_{1,2}a_{3,5}a_{4,6}$. The negative sign is inserted because it takes and odd number of exchanges in the order of the subscripts to produce this term. The last possible term starting with $a_{1,2}$ is $a_{1,2}a_{3,6}a_{4,5}$ which is positive because it takes an even number of exchanges to produce.

The three terms produced above are the first three terms in the Pfaffian expansion in Eq. (4.7). This process is repeated with $a_{1,3}$ and $a_{1,4}$ as the first elements. There should be $(2h-1)!! = 5!!$, or 15, terms in the expansion.

This has been a brief introduction to the Pfaffian. For the purposes of this paper the most important things about a Pfaffian are that each element appears at most once in a term and that the square of the Pfaffian is equal to the determinant of the corresponding antisymmetric matrix. This is commonly written as

$$P^2 = D = |d_{r,s}|, \tag{4.11}$$

where $d_{r,s}$ is an element of the matrix D and

$$d_{r,r} = 0, \qquad d_{r,s} = -d_{s,r}. \tag{4.12}$$

Fisher used the notation $a_{r,s} = (r, s)$ which might be less cumbersome. However since (r, s) is later used as the definition of an anticommutator ambiguity is avoided here by maintaining the notation $a_{r,s}$.

CHAPTER 5

A TRACE OF OPERATORS EXPRESSED AS A PFAFFIAN

From Eq. (2.20) the partition function is expressed as the product of a number of factors, each of which is an inhomogeneous combination of anticommuting operators. In Chapter 4, it was said that Hurst and Green claimed a product of homogeneous operators of this form can be reduced to a Pfaffian [10]. In this section it will be proved that the inhomogeneous set of operators in Eq. (2.20) has the same property. The Pfaffian is more convenient because it has fewer terms in its expansion than are obtained by doing the product in Eq. (2.20) directly. In fact the number of terms in Eq. (2.20) grows as 5^{mn} because each factor in the product expansion is a sum that has a monomer term and four half dimer terms.

The basic formula relating a Pfaffian to the trace of a product of operators U_r is

$$t^{-1}\mathrm{Tr}\left\{\prod_{r=1}^{2h} U_r\right\} = \left\backslash v_{r,0}v_{s,0} - \sum_{i=1}^{\mathcal{N}}(-1)^{r+s}v_{r,i}v_{s,i}\right|, \tag{5.1}$$

where

$$U_r = v_{r,0}I + \sum_{i=1}^{\mathcal{N}} v_{r,i}B_i, \tag{5.2}$$

and where the B_i follow the anticommuting laws in Eq. (2.22). Note the similiarity between U_r and $V_k = z_k I + x_{k,\ell_1}{}^{1/2}A_{k,\ell_1} + x_{k,\ell_2}{}^{1/2}A_{k,\ell_2} + y_{k,\ell_3}{}^{1/2}A_{k,\ell_3} + y_{k,\ell_4}{}^{1/2}A_{k,\ell_4}$.

Hurst and Green used a similar equation but assumed that $v_{r,0} = 0$ for all r. Eventually the same will be done in this paper since the $v_{r,0}$ is related to the monomer term in Eq. (5.1), which is currently required to be zero for the sake of parity. The Hurst and Green paper offered no formal proof of this relationship and instead stated that it could be shown through demonstration that it gave the correct terms for their

35

Pfaffian[10].

Fisher proved the relationship in Eq. (5.1) and, for the sake of completeness, that proof is shown in this chapter. It is important to note that Fisher's sum ranged from $i = 1$ to $i = 4\nu$ in his expression that is equivalent to Eq. (5.1). If the same 4ν used in Sec. 2.3 to label the A_i bonds then Eq. (5.1) might lead to some nonsensical relationships. It is possible that through some manner of cancellation this is the correct index but here the index $\mathcal{N}$ is used until the appropriate index can be determined.

To prove that Eq. (5.1) is true, all of the operators in the trace must be turned into homogeneous operators. Homogeneous in this case is the requirement that each term in U_r be of the same order. Or, in other words, each term in U_r must be multiplied by the same kind of operator. This will allow operators to be moved past each other in the product more easily.

5.1 Complementary Operators and Their Properties

To make U_r homogeneous a set of operators that are complementary to the B_i are introdcued. They are defined by

$$C_0 = \prod_{j=1}^{\mathcal{N}} B_j, \tag{5.3}$$

and by

$$C_j = iB_jC_0 = -iC_0B_j, \qquad (i^2 = 1). \tag{5.4}$$

These C_j will follow similar laws of anticommutation as Eq. (2.22) and will be in the form

$$C_iC_j + C_jC_i = 2\delta_{i,j}, \tag{5.5}$$

and

$$C_0{}^2 = I \tag{5.6}$$

These relationships must be proved true, however first the correct index on the sum is found. The full expansion of Eq. (5.6) is

$$C_0{}^2 = B_1 B_2 B_3 B_4 B_5 B_6 B_7 B_8 ... B_\mathcal{N} B_1 B_2 B_3 B_4 B_5 B_6 B_7 B_8 ... B_\mathcal{N}. \qquad (5.7)$$

To find the correct index, $\mathcal{N}$, the operators must be moved past each other in the product until all like operators are next to each other. The second B_1 operator must move $\mathcal{N} - 1$ times to the left to be next to the first B_1. B_2 must then move $\mathcal{N} - 2$ times, and so on for all the operators until the last two move left one time and then zero times.

If $\mathcal{N}$ is odd, then the terms with odd subscripts must always move an even number of times to get to the appropriate place (assuming the second set of operators are moved left in ascending order of their subscripts every time). However those with an even subscript must move an odd number of times. Each even subscript, therefore, creates one negative sign. To ensure that $C_0{}^2$ is positive, the number of even subscripts in C_0 must also be even, thus creating an even number of negative signs which subsequently cancel. To ensure that the number of even subscripts is even, $\mathcal{N}$ must be even.

If $\mathcal{N}$ is even, when moving the operators to their correct places in ascending numerical order the odd numbered subscripts take an odd number of interchanges. The even subscripts take an even number. Since $C_0{}^2$ must be positive, there must be an even number of odd subscripts in the expansion. Therefore $\mathcal{N}$ must be double even, $\mathcal{N} = 4\mathtt{n}$, where $\mathtt{n}$ is an integer and is different from the number of columns in the lattice, n. It appears this is where Fisher's index of 4ν was derived. His ν appears to be unrelated to the number of A_i operators that were used earlier to label every

bond site. Eq. (5.1) can now be written

$$t^{-1}\mathrm{Tr}\left\{\prod_{r=1}^{2h} U_r\right\} = \left\backslash v_{r,0}v_{s,0} - \sum_{i=1}^{4n}(-1)^{r+s}v_{r,i}v_{s,i}\right|,$$ (5.8)

where

$$U_r = v_{r,0}I + \sum_{i=1}^{4n} v_{r,i}B_i.$$ (5.9)

Limiting $\mathcal{N}$ in this manner is not a drawback in this situation because V_k will have at most four nearest neighbor sites under the sum.

Now that the correct index is known it would be useful to make sure the anti-commuting relationships are consistent. To prove Eq. (5.5) is true the definition of C_j is plugged into it. Eq. (5.5) now becomes

$$(iB_iC_0)(iB_jC_0) + (iB_jC_0)(iB_iC_0) = 2\delta_{i,j},$$ (5.10)

where $i^2 = -1$. Multiplying together the factors of i gives

$$-(B_iC_0)(B_jC_0) - (B_jC_0)(B_iC_0) = 2\delta_{i,j},$$ (5.11)

To prove that the relationship is true, the C_0 should be moved to be next to each other in order to create an identity operator. Then the relationship will be the same as Eq. (2.22) since the B_i were defined as having the same properties as the A_i operators. To do this, one must consider how the C_0 and B_i operators move past each other. For instance, the multiplication of a single B_j operator with a C_0 is given by

$$B_jC_0 = B_j\prod_{i=1}^{4n} B_i = B_jB_1B_2B_3B_4B_5B_6B_7B_8...B_j...B_{4n-1}B_{4n}$$ (5.12)

For an even number of terms in C_0, it takes any single B_j an even number of interchanges to pass all the way through C_0. Each of those interchanges will create a negative sign except the interchange where B_j must pass through itself. This results

38

in an odd number, $4\mathbf{n} - 1$, of negative signs, and so interchanging a B_j with a C_0 must give a sign change. Incidentally, this proves Eq. (5.4) is correct. Returning to Eq. (5.11), $B_jC_0 = -C_0B_j$ is substituted into the first term and $B_iC_0 = -C_0B_i$ is substituted into the secong giving,

$$(B_iC_0)(C_0B_j) + (B_jC_0)(C_0B_i) = 2\delta_{i,j}. \qquad (5.13)$$

After using the identity property, $C_0^2 = I$, this turns into

$$(B_iB_j) + (B_jB_i) = 2\delta_{i,j}, \qquad (5.14)$$

which is exactly the communation that there B_i were defined as following when they were introduced in Eq. 5.2.

5.2   Operators with Homogeneity

With these complementary C_i operators, the U_i operators can be put into a homogeneous form. This again follows from Fisher's work but he notes that his proof is an extension of the work done by Caianiell and Fubini on Dirac spurs[13]. The product of operators is written as

$$\left\{\prod_{r=1}^{2h} U_r\right\} = (U_1C_0)(C_0U_2)(U_3C_0)(C_0U_4)...(U_{2h-1}C_0)(C_0U_{2h}) = \left\{\prod_{r=1}^{2h} Q_r\right\}, \quad (5.15)$$

where $I = C_0$ has been inserted between each pair of operators. Each of the U_rC_0 or C_0U_r will become a single homogeneous operator named Q_r. All the terms with odd r have C_0 to the right of U_r and terms with even r have C_0 to the left.

Each term with r odd looks like

$$U_{r=odd} = (v_{r=odd,0}I + \sum_{i=1}^{4\mathbf{n}} v_{r=odd,i}B_i)C_0. \qquad (5.16)$$

39

Distributing C_0 gives

$$U_{r=odd} = v_{r=odd,0}C_0 + \sum_{i=1}^{4n} v_{r=odd,i}B_iC_0, \tag{5.17}$$

and using the relationship in Eq. (5.4) yields

$$U_{r=odd} = v_{r=odd,0}C_0 + \sum_{i=1}^{4n} v_{r=odd,i}(-1)iC_i. \tag{5.18}$$

This form is homogeneous as defined above as each team is multiplied by an operator of the same order. By the same logic the terms with r even are

$$U_{r=even} = C_0 v_{r=even,0} + \sum_{i=1}^{4n} iC_i v_{r=even,i}. \tag{5.19}$$

The even and odd versions can be rearranged and combined so long as no operators pass through each other. This gives for r even or odd,

$$U_r = v_{r,0}C_0 + \sum_{i=1}^{4n} (-1)^r i v_{r,i} C_i. \tag{5.20}$$

This is the homogeneous linear combination of operators that will be defined as Q_r. With a change of indices in the sum, Q_r is

$$Q_r = \sum_{i=0}^{4n} q_{r,i} C_i, \tag{5.21}$$

where the coefficients are

$$q_{r,0} = v_{r,0}, \qquad q_{r,j} = i(-1)^r v_{r,j}. \tag{5.22}$$

The Q_r operators will follow a different anticommutation paradigm. It can be found by setting up the anticommutation relation

$$Q_r Q_s + Q_s Q_r = \sum_{i=0}^{4n} q_{r,i} C_i \sum_{j=0}^{4n} q_{s,j} C_j + \sum_{j=0}^{4n} q_{s,j} C_j \sum_{i=0}^{4n} q_{r,i} C_i. \tag{5.23}$$

40

After some rearranging and factoring the sums this becomes

$$Q_r Q_s + Q_s Q_r = \sum_{i=0}^{4n} \sum_{j=0}^{4n} q_{r,i} q_{s,j} (C_i C_j + C_j C_i). \tag{5.24}$$

By Eq. (5.5), the operator term on the right can be written in terms of a Kronecker delta,

$$Q_r Q_s + Q_s Q_r = \sum_{i=0}^{4n} \sum_{j=0}^{4n} q_{r,i} q_{s,j} 2\delta_{ij}. \tag{5.25}$$

Evaluating the sum over j with the Kronecker delta gives

$$Q_r Q_s + Q_s Q_r = 2 \sum_{i=0}^{4n} q_{r,i} q_{s,i} \tag{5.26}$$

The notation can be made a little more compact by defining the anticommutator (r, s):

$$(r, s) = \sum_{i=0}^{4n} q_{r,i} q_{s,i}, \tag{5.27}$$

and finally the anticommutation relationship for the Q_r operators is

$$Q_r Q_s + Q_s Q_r = 2(r, s). \tag{5.28}$$

5.3 The Trace of the Product Gives the Same Terms as the Pfaffian

Now that each operator U_r has been manipulated into the homogeneous form Q_r, the product of operators can be shown to expand with the same rules as the Pfaffian expansion in Sec. 4.1. The product of operators is

$$\prod_{r=1}^{2h} Q_r = Q_1 Q_2 Q_3 Q_4 Q_5 ... Q_{2h}. \tag{5.29}$$

This product is under the trace in Eq. (5.1) and the trace is invariant under cyclic permutations of products of matrices. With this in mind the next step will be to cycle Q_1 from the beginning of this product to the end of the product to determine how

41

this new product acts under the trace operation. Each interchange of operators will be accompanied by the anticommutation relationship givien in Eq. (5.28). Changing the order of Q_1 and Q_2 in Eq. (5.29) gives

$$Q_1Q_2Q_3Q_4Q_5...Q_2h = (-1)Q_2Q_1Q_3Q_4Q_5...Q_{2h} + 2(1,2)Q_3Q_4Q_5...Q_{2h}, \qquad (5.30)$$

which was found by manipulating Eq. (5.28) into the form $Q_1Q_2 = -Q_2Q_1 + 2(1,2)$ and then appending the rest of the product over Q_r to the right hand side of the anticommutation relation.

The term on the left hand side is simply the product of all the Q_r. The first term on the right has Q_1 displaced one place and this term can be written as

$$(-1)Q_2Q_1 \prod_{r=3}^{2h} Q_r = (-1)Q_2Q_1Q_3Q_4Q_5...Q_2h. \qquad (5.31)$$

Using the same anticommutation rule to exchange the positions of Q_1 and Q_3 gives

$$(-1)Q_2Q_1 \prod_{r=3}^{2h} Q_r = (-1)[(-1)Q_2Q_3Q_1Q_4Q_5...Q_{2h} + 2(1,3)Q_2Q_4Q_5...Q_{2h}]. \qquad (5.32)$$

Plugging this back into Eq. (5.30) yields

$$\prod_{r=1}^{2h} Q_r = (-1)^2 Q_2Q_3Q_1Q_4Q_5...Q_{2h} + 2(1,2)Q_3Q_4Q_5...Q_{2h}$$
$$- 2(1,3)Q_2Q_4Q_5...Q_{2h}. \qquad (5.33)$$

A pattern begins to appear here. The first term on the right shows the correct position of Q_1 as it moves through the product and it is multiplied by -1 for each interchange. Each new interchange creates an anticommutator, as seen in the other terms on the right hand side. The commutators that have been already created are multiplied by the products of all the Q_r and Q_s that are not included in that (r,s) anticommutator. The signs of the anticommutator terms are determined by whether

s is odd or even. Therefore, after the $2h - 1$ interchanges needed to move Q_1 from one end of the product to the other, the equation can be expressed in terms of the sum over all the commutators $(1, s)$ as,

$$\prod_{r=1}^{2h} Q_r = (-1)^{2h-1} Q_2 Q_3 Q_4 Q_5 ... Q_{2h} Q_1$$
$$+ 2 \sum_{s=2}^{2h} (-1)^s (1, s) Q_2 Q_3 ... Q_{s-1} Q_{s+1} ... Q_{2h}. \qquad (5.34)$$

The trace can now be taken on both sides of this equation. Since the trace is invariant under cyclic interchange, Q_1 can be moved back to the front of the product. Additionally $2h - 1$ is always odd. Therefore the trace of the product of Q_r is

$$\text{Tr} \left\{ \prod_{r=1}^{2h} Q_r \right\} = \text{Tr} \left\{ -Q_1 Q_2 Q_3 Q_4 Q_5 ... Q_{2h} \right\}$$
$$+ \text{Tr} \left\{ 2 \sum_{s=2}^{2h} (-1)^s (1, s) Q_2 Q_3 ... Q_{s-1} Q_{s+1} ... Q_{2h} \right\}. \qquad (5.35)$$

Moving the first term on the right hand side to the left hand side will cancel out the factor of 2. The sum over s and the sign factor can be pulled out of the trace on the right hand side to give

$$\text{Tr} \left\{ \prod_{r=1}^{2h} Q_r \right\} = \sum_{s=2}^{2h} (-1)^s (1, s) \text{Tr} \left\{ Q_2 Q_3 ... Q_{s-1} Q_{s+1} ... Q_2 h \right\}. \qquad (5.36)$$

At this point the process of carrying Q_1 all the way through the product has pulled out the anticommutator $(1, s)$ on the right hand side, but the left hand side has remained the same product. The sum on the right will contain every possible anticommutator $(1, s)$ as s goes from 2 to $2h$. If the elements of the Pfaffian are of the form $a_{r,s} = (r, s)$ then the anticommutator $(1, s)$, where $2 \leq s \leq 2h$, looks like the first row of the Pfaffian. Even the negative sign appears in the correct place with every other term being negative. However this is not a sufficient proof yet. The

process needs to be iterated on the remaining product on the right hand side. There are two cases. Either $s = 2$, which is exactly the same process, or $s > 2$ in which case there is a gap in the product and the sign of the result needs to be investigated. If the part of the right hand side of Eq. (5.36) that is in the sum is defined as

$$O(s) = (-1)^s (1, s) \text{Tr} \left\{ Q_2 Q_3 ... Q_{s-1} Q_{s+1} ... Q_2 h \right\}, \tag{5.37}$$

then for $s = 2$

$$O(2) = (1, 2)(-1)^2 \text{Tr} \left\{ Q_3 Q_4 ... Q_2 h \right\}. \tag{5.38}$$

Moving Q_3 through the entire product on the right hand side will give a similar result the movement of Q_1 through the entire product. After this operation the product on the right hand side produces the anticommutator (t, u). Then $O(2)$ becomes

$$O(2) = (1, 2) \sum_{u=4}^{2h} (-1)^u (3, u) \text{Tr} \left\{ Q_4 Q_5 ... Q_{u-1} Q_{u+1} ... Q_2 h \right\}. \tag{5.39}$$

This can be plugged back into the full product to give

$$\text{Tr} \left\{ \prod_{r=1}^{2h} Q_r \right\} = (1, 2) \sum_{u=4}^{2h} (-1)^u (3, u) \text{Tr} \left\{ Q_4 Q_5 ... Q_{u-1} Q_{u+1} ... Q_2 h \right\}. \tag{5.40}$$

It is obvious here that if the lowest possible value is always chosen to fill in the anticommutator then this product will produce the term $(1, 2)(3, 4)...(2h - 1, 2h)$ and that the term will be positive because the second argument of each anticommutator is even.

The expansion of the product is more interesting when $s > 2$ because of the gap this produces in the middle of the product on the right hand side which produces a change in the sign. When $u > s$ the sign will be as above where the anticommutator earned a $(-1)^u$ in front of it. However when $u < s$ the sign will be $(-1)^{u+1}$ to account

for the missing operator in the Q_s place that was removed by the first anticommutator. The product expansion of the two cases combined can be written as

$$\text{Tr}\left\{\prod_{r=1}^{2h} Q_r\right\} = \sum_{s=2}^{2h}(-1)^s(1,s)$$
$$\times \begin{cases} \sum_{u=3}^{2h}(-1)^u(t,u) & (u>s) \\ \sum_{u=3}^{2h}(-1)^{u+1}(t,u) & (u<s) \end{cases} \times \text{Tr}\left\{Q_\tau...Q_{u-1}Q_{u+1}...Q_{s-1}Q_{s+1}...Q_2h\right\}, \quad (5.41)$$

where Q_τ is used as a placeholder representing the lowest value yet to be taken by an anticommutator and where the positions of the missing operators in the product, Q_u and Q_s might not appear in the order shown here. Though it is not shown here, when $s > 2$ it must be true that $t = 2$ because the Q_τ is necessarily Q_2 in that case. Additionally when $s = 2$ then $t = 3$. Note this this t is not the same t that was used to define the size of A_i anticommuting operators. The first argument of each anticommutator is determined by the first operator in the product used to produce that anticommutator based on how the process of pulling commutators out is defined.

The next anticommutator which can be pulled out of the product on the right hand side is called (v, w). It adds further complications to the signs because, while v takes the value of τ, w can greater than both s and u, or between them, or less than both. Each of these will have a different parity term given respectively by $(-1)^w$, $(-1)^{w+1}$, $(-1)^{w+2}$. Because of the complications caused by tracking all these signs, the problem of the sign convention is left to the next section. For the rest of this section the Sign (Sgn) function will be used to prevent the formulae from becoming

overloaded with multiple cases for each possible sign. Thus

$$\text{Tr}\left\{\prod_{r=1}^{2h} Q_r\right\} = \sum_{s=2}^{2h} \text{Sgn}(s)(1,s) \sum_{\substack{u=3 \\ u \neq s,t}}^{2h} \text{Sgn}(u)(t,u) \sum_{\substack{w=4 \\ w \neq s,t,u,v}}^{2h} \text{Sgn}(w)(v,w)$$
$$\times \text{Tr}\left\{Q_\tau...Q_{u-1}Q_{u+1}...Q_{w-1}Q_{w+1}...Q_{s-1}Q_{s+1}...Q_2h\right\}, \tag{5.42}$$

where, again, the missing operators in the product might not occur in the order above. This is the same pattern used to define the Pfaffian expansion in Sec. 4.1, which can be seen from the expansion of this product. In each case the first argument in each anticommutator must be larger than the first term in the previous anticommutator, $r < t < v...$, since it is determined by the lowest remaining value in the product, Q_τ. The second argument in each anticommutator must be larger than the first argument in that same anticommutator, $r < s, \quad t < u, \quad v < w$. These are the same rules used in the Pfaffian expansion. Commutators can continue to be pulled out of the product in this manner until only two Q_r remain in the product, which will be named Q_A and Q_Ω. The remaining trace on the right hand side is

$$\text{Tr}\left\{Q_A Q_\Omega\right\}. \tag{5.43}$$

These final two operators are handled by breaking them into their coefficients to give

$$\text{Tr}\left\{Q_A Q_\Omega\right\} = \text{Tr}\left\{\sum_{i=0}^{4n} q_{A,i} C_i \sum_{j=0}^{4n} q_{\Omega,j} C_j\right\}, \tag{5.44}$$

which can be expanded even further using the definition of C_j in Eq. (5.4) yielding

$$\text{Tr}\left\{Q_A Q_\Omega\right\} = \text{Tr}\left\{\sum_{i=0}^{4n} q_{A,i} i B_i C_0 \sum_{j=0}^{4n} q_{\Omega,j} i B_j C_0\right\}. \tag{5.45}$$

Again, using the properties of Eq. (5.4), this can be rearranged in the form

$$\text{Tr}\left\{Q_A Q_\Omega\right\} = \text{Tr}\left\{-(i^2)\sum_{i=0}^{4n}\sum_{j=0}^{4n} q_{A,i} q_{\Omega,j} B_i C_0 C_0 B_j\right\}, \tag{5.46}$$

which, since $C_0^2 = I$, reduces to

$$\text{Tr}\left\{Q_A Q_\Omega\right\} = \sum_{i=0}^{4n}\sum_{j=0}^{4n} q_{A,i} q_{\Omega,j} \text{Tr}\left\{B_i B_j\right\}, \tag{5.47}$$

The anticommutation rules in Eq. (2.24) state that $\text{Tr}\left\{B_i B_j\right\} = 0$ when $i \neq j$. Otherwise it is equal to t, the dimension of the identity matrix $I = B_j^2$. The trace on the right hand side can be replaced with a Kronecker delta.

$$\text{Tr}\left\{Q_A Q_\Omega\right\} = \sum_{i=0}^{4n}\sum_{j=0}^{4n} q_{A,i} q_{\Omega,j} \delta_{ij} t. \tag{5.48}$$

Taking the sum over j and moving the t to the other side gives

$$t^{-1}\text{Tr}\left\{Q_A Q_\Omega\right\} = \sum_{i=0}^{4n} q_{A,i} q_{\Omega,i}. \tag{5.49}$$

The right hand side is exactly the definition given for the anticommutator in Eq. (5.27). Therefore

$$t^{-1}\text{Tr}\left\{Q_A Q_\Omega\right\} = (q_{A,i}, q_{\Omega,i}), \tag{5.50}$$

where A is greater than the first argument of every anticommutator that was already removed from the sum, and where Ω is greater that A. It is important for the next section that these last two operators contribute no sign to the term that includes them since they are removed without swapping. This completes the proof that the expansion of the product of operators gives the same terms as the expansion of the Pfaffian. The next section will deal with making sure both expansions give the same sign.

5.4 The Sign in the Product

If, in the Pfaffian, the elements $a_{r,s}$ are defined as

$$a_{r,s} = (r, s), \tag{5.51}$$

then, as was shown in Sec. 5.3 the trace in Eq. (5.1) has the same terms as the Pfaffian expansion of (r, s). If the signs of the terms in the trace can be determined in the same manner as the signs in the Pfaffian expansion, then that is sufficient proof of the relationship given in Eq. (5.1). In the Pfaffian, the sign of each term was determined by determining the number of interchanges it would take to put all of the subscripts in counting order. The subscripts of the elements match the arguments of the commutators (r, s), and so it must be determined whether the sign of the terms can be found by interchanging the arguments of the commutators until they are in counting order. This section will show that this is the case. Though the exponents do not directly count the number of interchanges required to achieve counting order, they do completely account for the parity that would be required for each argument, in a term of this expansion consisting of a product of commutators, to be put into counting order.

As can be seen in Eq. (5.41), the sign of a particular term is determined by

$$(-1)^s \times \begin{cases} (-1)^u, & (u > s) \\ (-1)^{u+1}, & (u < s) \end{cases}. \tag{5.52}$$

48

This relationship continues into the (v, w) anticommutator with the sign given by

$$(-1)^s \times \begin{cases} (-1)^u, & (u > s) \\ (-1)^{u+1}, & (u < s) \end{cases} \times \begin{cases} (-1)^w, & (w > s, u) \\ (-1)^{w+1}, & (s < w < u) \quad \text{or} \quad (u < w < s) \\ (-1)^{w+2}, & (w < s, u) \end{cases} \cdot \quad (5.53)$$

The sign of each new anticommutator depends on how many of the previous commutators contain a number in the second argument which is higher than the second argument in the new anticommutator. This can be seen in the remaining trace in Eq. (5.41), which is shown below,

$$\text{Tr}\left\{ Q_\tau ... Q_{u-1} Q_{u+1} ... Q_{w-1} Q_{w+1} ... Q_{s-1} Q_{s+1} ... Q_2 h \right\}. \quad (5.54)$$

If a new anticommutator is taken out, the first argument will be Q_τ. The second argument can be any operator whose subscript follows the rules expanding the product, or the rules for expanding the Pfaffian. Note again that the order in which the gaps appear in the product above is arbitrary for the moment.

The first step is to show that the term with all of the arguments in counting order has even parity because this is true for the Pfaffian term with all the subscripts in counting order. The term is

$$(1, 2)(3, 4)(5, 6)...(2h - 1, 2h), \quad (5.55)$$

and the parity is definitely even. The sign contributed by each anticommutator comes from the s of the anticommutator and each s is an even number when the term is completely ordered. The positions of the numbers in the completely paired term will be called the "correct" positions for this section. It takes an even number of swaps to move a number from a position of one parity to a position of like parity. It takes

an odd number of interchanges to move to a position of opposite parity. This means that any term which can be reduced to, or created from, the "correct" term via an even number of interchanges must have even parity and be a positive. Of course if an odd number of interchanges is required, then the term is negative. This is the same rule used in the Pfaffian expansion.

Since the sign of each term is determined by the argument in the s place of each anticommutator, it must be shown that counting order can be achieved only through moving the numbers in the s positions to their correct positions. This is easy to realize since the products are created using the same inequality rules as the Pfaffian. Therefore all of the arguments in r positions will already be in ascending order. Consider

$$(1,5)(2,4)(3,6)(7,8), \tag{5.56}$$

which has r arguments 1, 2, 3, and 7. This restriction on r means the only arguments that could possibly not be in ascending order are those in s positions. Moving all of the numbers in the s positions to their correct positions will necessarily force all of the r to be correct because of the inequality rules.

Now consider the general term

$$(r,s)(t,u)(v,w)...(2h-1,2h). \tag{5.57}$$

The sign contributed by the first anticommutator is $(-1)^s$. This checks whether the number s is odd or even. If it is even, it will take an even number of interchanges to move it to its correct spot. If odd, then it will take an odd number of swaps. This is the correct relationship because in the fully correct term the numbers occupying the places $s, u, w, ..., 2h$ must all be even.

The second anticommutator has two possible signs. If $u > s$ then the sign is

$(-1)^u$. This is because u is not shifted any positions in the general term by moving s through it. This corresponds to a product of Q's that has no gap that Q_u must pass through. If $s > u$ then the sign is $(-1)^{u+1}$. The plus one comes from the fact to the move everything to the correct position s must pass through u. This will naturally shift u one place to the left. Therefore if u were even, it would take an odd number of interchanges is move it to the correct spot, and so the total sign contribution of the anticommutator would be odd, and vice versa.

The same logic can be applied to w. It will either be shifted zero, one, or two places in Eq. (5.57) by numbers that are greater than it being to the left of it in the general product of anticommutators. This is sufficient to show that even though the sign terms do not track how many total interchanges are required, they do correctly track the parity required by considering whether the number in question is even or odd and then determining how the other numbers shift it around as they move.

Two things should be noted about this section. The first is that while each of the arguments is being interchanged in order to arrange the arguments into counting order they might create a term that is not possible with the inequality rules. Just as with the shifting of the arcs in Sec. 3.2, this is permissible so long as the starting term was valid and the ending term is valid. The second is that the exchanges between arguments were done by starting at the left of each term and moving the numbers which were not in counting order to the right until they were in what would be considered their "correct" positions in counting order. However, this is not necessary since there are an even number of arguments in each term. The sign should end up the same either way. Starting at the left and moving arguments to the right was the simplest method, and the most appropriate, given how the commutators and their signs were produced.

Now that both the sign and the terms themselves are proven to the same as the

Pfaffian expansion, it can be said that

$$t^{-1}\text{Tr}\left\{\prod_{r=1}^{2h} Q_r\right\} = \backslash(r, s)|,$$ (5.58)

where (r, s) is the anticommutator in Eq. (5.27). If Eqs. (5.27) and (5.22) are substituted into Eq. (5.58) along the definitions of all of the commutators given Eqs. (5.21), (5.3), and (5.4) it can be seen that Eq. (5.1) is true.

CHAPTER 6

THE PARTITION FUNCTION AS A PFAFFIAN

In Chapter 5 it was seen that the partition function can be expressed as a Pfaffian. The partition function for a lattice completely covered only in dimers is

$$
\begin{aligned}
Z_{mn} &= t^{-1} \mathrm{Tr} \left\{ \prod_{k=1}^{mn} V_k \right\} \\
&= t^{-1} \mathrm{Tr} \left\{ \prod_{k=1}^{mn} \left[x_{k,\ell_1}{}^{1/2} A_{k,\ell_1} + x_{k,\ell_2}{}^{1/2} A_{k,\ell_2} + y_{k,\ell_3}{}^{1/2} A_{k,\ell_3} + y_{k,\ell_4}{}^{1/2} A_{k,\ell_4} \right] \right\},
\end{aligned}
\tag{6.1}
$$

where ℓ_1 to ℓ_4 represent the nearest neighbors of the point k, as in Fig 5. The $z_k I$ term has be dropped for the sake of parity, which means that there will be no monomers occupying the lattice. This form of the partition function is large and unwieldy. In a 3×4 lattice the product is taken from $k = 1$ to $k = 12$ and will give a total of $4^{12} = 16,777,216$ terms, which must then be paired and eliminated in the trace. By contrast the Pfaffian related to this term will have $(2h - 1)!! = (11)!! = 10395$ terms. Part of the purpose behind using the Pfaffian is to take advantage of this drastic reduction in order of magnitude when determining possible states of a lattice.

6.1 Equating the General Trace to the Trace in the State Function

With the general case for the relationship between the Pfaffian and the trace of a product of anticommuting operators proven, the next step is to see what the form the Pfaffian takes when applied to the specific set of operators used in this paper. In

general it can be stated that

$$Z_{mn} = t^{-1}\text{Tr}\left\{\prod_{k=1}^{mn}\left[x_{k,\ell_1}^{1/2}A_{k,\ell_1} + x_{k,\ell_2}^{1/2}A_{k,\ell_2} + y_{k,\ell_3}^{1/2}A_{k,\ell_3} + y_{k,\ell_4}^{1/2}A_{k,\ell_4}\right]\right\}$$

$$= t^{-1}\text{Tr}\left\{\prod_{r=1}^{2h}v_{r,0}I + \sum_{i=1}^{4n}v_{r,i}B_i\right\} = \left|\begin{array}{c} v_{r,0}v_{s,0} - \sum_{i=1}^{4n}(-1)^{r+s}v_{r,i}v_{s,i}\end{array}\right|, \qquad (6.2)$$

where $v_{r,0} = 0$ for all r by the fact that the state function has no monomers. To make these have more similar forms, a generalized symbol for the half-bonds in the state functions is introduced. χ_{k,ℓ_j} represents the bonds $x_{k,\ell_1}^{1/2}$, $x_{k,\ell_2}^{1/2}$, $y_{k,\ell_3}^{1/2}$, and $y_{k,\ell_{41}}^{1/2}$. Then the state function becomes

$$Z_{mn} = t^{-1}\text{Tr}\left\{\prod_{k=1}^{mn}\sum_{j=1}^{4}\chi_{k,\ell_j}A_{k,\ell_i}\right\}$$

$$= t^{-1}\text{Tr}\left\{\prod_{r=1}^{2h}\sum_{i=1}^{4n}v_{r,i}B_i\right\} = \left|\begin{array}{c} -\sum_{i=1}^{4n}(-1)^{r+s}v_{r,i}v_{s,i}\end{array}\right|. \qquad (6.3)$$

The two traces are starting to look similar so a discussion of the Pfaffian is postponed. It is immediately obvious that r must equal k and that $2h$ must equal mn. Though it was not stated explicitly before, mn must be even because the lattice is to be complately filled with dimers. The idea that mn is even is reinforced by the properties of the Pfaffian. The Pfaffian must have an even order, $2h$, because it is defined as the upper right triangle of an antisymmetric matrix. The determinant of an antisymmetrix matrix vanishes if the number of rows and columns is odd. Since the Pfaffian has one fewer rows than the matrix, the Pfaffian is only defined if there are an odd number of columns in it $(2h - 1)$. The order of the Pfaffian is determined by the related antisymmetrix matrix. The formula for the partition function can now be written

$$Z_{mn} = t^{-1}\text{Tr}\left\{\prod_{k=1}^{mn}\sum_{j=1}^{4}\chi_{k,\ell_j}A_{k,\ell_i}\right\} = t^{-1}\text{Tr}\left\{\prod_{k=1}^{mn}\sum_{i=1}^{4n}v_{r,i}B_i\right\}, \qquad (6.4)$$

where the B_i are defined as following the same anticommutation rules as A_i. To make the general form on the right equal the partition function on the left, the sum over all bonds, $\sum_{i=1}^{4n} v_{r,i} B_i$, must be converted to a sum only over nearest neighbors. This is done be defining $i = k\ell_i$ and setting $\mathbf{n} = 1$. This would cause the v to take the subscripts v_{k,k,ℓ_i}. The double labeling in k is unnecessary and therefore dropped. The formula can now be written

$$Z_{mn} = t^{-1}\text{Tr}\left\{\prod_{k=1}^{mn}\sum_{j=1}^{4}\chi_{k,\ell_j}A_{k,\ell_i}\right\} = t^{-1}\text{Tr}\left\{\prod_{k=1}^{mn}\sum_{i=1}^{4}v_{k,\ell_i}A_{k,\ell_i}\right\}. \qquad (6.5)$$

It can clearly be seen that v_{k,ℓ_i} must be equal to χ_{k,ℓ_j}. The two traces are now equal. It must now be determined how the changes affect the form in the Pfaffian.

6.2 Equating the State Function to the Pfaffian

The transformations applied to the general trace in the previous section can be be applied directly to the Pfaffian

$$\left\backslash -\sum_{i=1}^{4n}(-1)^{r+s}v_{r,i}v_{s,i}\right| . \qquad (6.6)$$

The main difficulty lies in converting the sum over all bonds, which would include bonds that dimers can not possibly occupy, to a sum over nearest neighbors. The Pfaffian lacks the anticommuting operators which paired to give 0 or 1 in the trace, depending on whether the bond is even possible.

The problem can be demonstrated quickly. If the changes were applied directly, then the Pfaffian would have the form

$$\left\backslash -\sum_{i=1}^{4}(-1)^{k+\ell}v_{k,\ell_i}v_{\ell,\ell_i}\right| , \qquad (6.7)$$

where $s = \ell$, and ℓ is any point on the lattice greater than k because $s > r$ was defined in the Pfaffian expansion. Immediately there is a problem. The ℓ_i are the

nearest neighbors to k, and so the second v is indicating a bond that goes from any point ℓ to the nearest neighbors of k.

Even if the ℓ_i are considered to be the nearest neighbors to the subscript with which they are paired, this form fails. Imagine a simple lattice with two points. The Pfaffian would be

$$\left. \diagdown \; - \sum_{i=1}^{2} (-1)^{1+2l} v_{1,\ell_i} v_{2,\ell_i} \right|. \tag{6.8}$$

To be correct, the Pfaffian must return the term $x_{1,2}$ just as the trace would. However there is no way for v_{1,ℓ_i} to equal $v_{1,2}$ while v_{2,ℓ_i} equals $v_{2,1}$ (recall that $x_{1,2}{}^{1/2} = x_{2,1}{}^{1/2}$), because Figs. 5 and 6 shows that the v_{2,ℓ_i} should pick the half bond to its own right. It will not be pointing from 2 to 1.

To fix this the definition of the anticommutator (r,s) must be slightly modified. Setting up the same anticommuting rules from before gives

$$Q_r Q_s + Q_s Q_r = \sum_{i=0}^{4\mathbf{n}} q_{r,i} C_i \sum_{j=0}^{4\mathbf{n}} q_{s,j} C_j + \sum_{j=0}^{4\mathbf{n}} q_{s,j} C_j \sum_{i=0}^{4\mathbf{n}} q_{r,i} C_i. \tag{6.9}$$

Now the changes that were made earlier in the traces must also be made to this anticommuting rule. Specifically $r = k$, $s = \ell$ are substituted in and the sums should be taken over i, while the arguments of the sums undergo the change $i = k\ell_i$ to make the terms correspond to nearest neighbors. Substituting gives

$$Q_k Q_\ell + Q_\ell Q_k = \sum_{i=0}^{4} q_{k,\ell_i} C_{k\ell_i} \sum_{j=0}^{4} q_{\ell,\ell_j} C_{\ell,\ell_j} + \sum_{j=0}^{4} q_{\ell,\ell_j} C_{\ell,\ell_j} \sum_{i=0}^{4} q_{k,\ell_i} C_{k,\ell_i}, \tag{6.10}$$

where, as before, anything that was labeled with the same subscript twice has had the extra subscript eliminated. Like terms can be factored giving

$$Q_k Q_\ell + Q_\ell Q_k = \sum_{i=0}^{4} \sum_{j=0}^{4} q_{k,\ell_i} q_{\ell,\ell_j} (C_{k\ell_i} C_{\ell,\ell_j} + C_{\ell,\ell_j} C_{k,\ell_i}). \tag{6.11}$$

To get to Eq. (5.26), a Kronecker delta $\delta_{i,j}$ was used to eliminate one of the sums.

This is not the proper way to proceed when the sum must be over nearest neighbors. Instead a term is required that will give 1 when $k\ell_i = k\ell_j$ and zero otherwise (the order of the subscripts doesn't matter). Instead of the Kronecker delta, a new symbol is defined.

$$(C_{k\ell_i}C_{\ell,\ell_j} + C_{\ell,\ell_j}C_{k,\ell_i}) = 2\Gamma_{k,\ell}, \tag{6.12}$$

where $\Gamma_{k,\ell}$ is equal to 1 if a dimer bond could exist between the points k and ℓ and is 0 otherwise. Then

$$Q_k Q_\ell + Q_\ell Q_k = 2(k,\ell), \tag{6.13}$$

where the anticommutator (k,ℓ) is

$$(k,\ell) = \sum_{i=1}^{4}\sum_{j=0}^{4} q_{k,\ell_i}q_{\ell,\ell_j}\Gamma_{k,\ell}. \tag{6.14}$$

The anticommutator can then be returned to the form

$$(k,\ell) = -\sum_{i=1}^{4}\sum_{j=0}^{4}(-1)^{k+\ell}v_{k,\ell_i}v_{\ell,\ell_i}\Gamma_{k,\ell}. \tag{6.15}$$

It can be seen by investigating small lattices that the factor $(-1)^{k+\ell}$ must be -1. Dimer bonds can only exist from an odd point to an even point on the lattice or vice versa. Therefore the anticommutator is

$$(k,\ell) = \Gamma_{k,\ell}\sum_{i=1}^{4}\sum_{j=0}^{4}v_{k,\ell_i}v_{\ell,\ell_i}. \tag{6.16}$$

Because of the $\Gamma_{k,\ell}$ factor, no matter what order the two sums are evaluated, the only nonzero will be those for which $v_{k,\ell_i} = v_{\ell,\ell_i}$. The anticommutator must therefore be equivalent to

$$(k,\ell) = \Gamma_{k,\ell}v_{k,\ell}{}^2. \tag{6.17}$$

The partition function can now be correctly written as a Pfaffian with the form

$$Z_{mn} = t^{-1}\text{Tr}\left\{\prod_{k=1}^{mn}\sum_{j=1}^{4}\chi_{k,\ell_j}A_{k,\ell_i}\right\} = \left\backslash\overline{\Gamma_{k,\ell}\upsilon_{k,\ell}}^2\right|, \tag{6.18}$$

where $\upsilon_{k,\ell} = \chi_{k,\ell_j}$ can be substituted to give

$$Z_{mn} = t^{-1}\text{Tr}\left\{\prod_{k=1}^{mn}\sum_{j=1}^{4}\chi_{k,\ell_j}A_{k,\ell_i}\right\} = \left\backslash\overline{\Gamma_{k,\ell}\chi_{k,\ell_j}}^2\right|. \tag{6.19}$$

It can be said that $\chi_{k,\ell_j} = x_{k,\ell}\xi_{l,\ell} + y_{k,\ell}\eta_{k,\ell}$, where $\xi_{k,\ell}$ and $\eta_{k,\ell}$ are 1 or 0 depending on whether an x or y bond exists. This is possible because one part of the sum must always be zero. Therefore

$$Z_{mn} = t^{-1}\text{Tr}\left\{\prod_{k=1}^{mn}\sum_{j=1}^{4}\chi_{k,\ell_j}A_{k,\ell_i}\right\} = \left\backslash\overline{(x_{k,\ell}\xi_{l,\ell} + y_{k,\ell}\eta_{k,\ell})}\right|. \tag{6.20}$$

To appreciate that is is the correct form consider a simple one dimensional lattice with four points, all of which are oriented in the x direction. The Pfaffian expansion of (k,ℓ) is

$$\backslash(k,\ell)|= (1,2)(3,4) - (1,3)(2,4) + (1,4)(2,3), \tag{6.21}$$

where it can be quickly seen that the last two terms will not survive the $\Gamma_{k,\ell}$ operation. The first term gives $x_{1,2}x_{3,4}$ which is the only possible filled lattice for this particular one dimensional case.

One of the interesting properties of the combination of the Pfaffian and the zig-zag pattern for numbering lattice rows is that every term in the Pfaffian expansion that is negative will be eliminated by $\Gamma_{k,\ell}$. This is necessary for the Pfaffian to be a state function. This property further reduces the number of terms that need to be investigated by about half. If the same 3×4 lattice example at the beginning of this chapter is used, the terms are further reduced from 10395 terms to $\frac{(2h-1)!!-1}{2} = 5198$.

CHAPTER 7

EVALUATION OF THE DETERMINANT

In Ch. 4 it was stated that one of the most important properties of the Pfaffian is that its square is equal to the determinant of the corresponding antisymmetric matrix, specifically

$$P^2 = |D| = \backslash a_{r,s}|. \tag{7.1}$$

Matrix manipulation is harder to do in the Pfaffian form, and so the Pfaffian will be used to build an antisymmetric matrix. The new matrix, D, will then be manipulated to find a general form for the state function via

$$Z_{mn}{}^2 = P^2 = |D|. \tag{7.2}$$

The state function is given as

$$Z_{mn} = \backslash (k, \ell)|, \tag{7.3}$$

whose form will vary slightly. The Pfaffian will generally take a form similar to

$$\begin{vmatrix}
x & & & & & & & & & y & & & & & & \\
& x & & & & & & & y & & & & & & \\
& & \ddots & & & \cdot^{\cdot^{\cdot}} & & & & & & & & & \\
& & & x & & y & & & & & & & & & \\
& & & & y & & & & & & & & & & \\
& & & & & x & & & & & & & y & & \\
& & & & & & x & & & & & y & & & \\
& & & & & & & \ddots & & & \cdot^{\cdot^{\cdot}} & & & & \\
& & & & & & & & x & y & & & & & \\
& & & & & & & & y & & & & & & \\
& & & & & & & x & & & & & & & y \\
& & & & & & x & & & & & & & y & \\
& & & & & & & & \ddots & & & \cdot^{\cdot^{\cdot}} & & & \\
& & & & & & & & & x & y & & & & \\
& & & & & & & & & y & & & & & \\
& & & & & & & & & x & & & & & \\
& & & & & & & & x & & & & & & \\
& & & & & & & & & \ddots & & & & \\
& & & & & & & & x & & & & & &
\end{vmatrix}, \qquad (7.4)$$

where the blank spaces in the matrix are filled with zeroes. This is the Pfaffian that would be created by making each term the anticommutator (k, ℓ) and then doing the $\Gamma_{k,\ell}$ operation. The subscripts have been removed from the x and y elements since the partition function only cares about the numbers of x and y and not their position on the lattice. If they need to be recovered they are simply the element discriptors in the matrix corresponding to this Pfaffian.

Fig. 20.: A 4 × 3 lattice

7.1 The Formulation of the Block Matrix

To better understand the placement of the x and y elements in the Pfaffian and corresponding matrix consider the lattice in Fig. 20. The state function in Pfaffian form for this lattice is

$$
\begin{vmatrix}
x & & & & y & & & & & & & \\
 & x & & y & & & & & & & & \\
 & & y & & & & & & & & & \\
 & & & x & & & & y & & & & \\
 & & & & x & & y & & & & & \\
 & & & & & y & & & & & & \\
 & & & & & & x & & & & y & \\
 & & & & & & & x & & y & & \\
 & & & & & & & & y & & & \\
 & & & & & & & & & x & & \\
 & & & & & & & & & & x & \\
 & & & & & & & & & & & x
\end{vmatrix} ,
\tag{7.5}
$$

where all the blanks are filled with zeroes which have been left out to ease visibility. The x's appear only down the diagonal edge of the matrix. There will be $n-1$ x's between each y because the n horizontal points in the lattice can be connected by $n-1$ bonds. A y appears on the diagonal when the numbers in the counting scheme jump up to the next row of the lattice. This then repeats down the diagonal. Each y on the diagonal is part of a set of y's extending counter to the diagonal. The number of y's in each set is equal to $m-1$, which is the total number of possible vertical bonds in a lattice column.

A Pfaffian can be quickly constructed for any lattice by filling in only those places in the Pfaffian where a bond could exist with the appropriate x or y element. Completing the Pfaffian determinant will give the state function of that lattice. This is fine to do by hand for smaller lattices but for very large lattices a more general form is required. To find this requires finding the general form of the determinant of the antisymmetrix matrix corresponding with the Pfaffian. As an example the matrix

associated with the above Pfaffian is

$$
Z^2_{mn} =
\begin{bmatrix}
0 & x & & & & y & & & & & & \\
-x & 0 & x & & y & & & & & & & \\
& -x & 0 & y & & & & & & & & \\
& & -y & 0 & x & & & & y & & & \\
& -y & & -x & 0 & x & & y & & & & \\
-y & & & & -x & 0 & y & & & & & \\
& & & & & -y & 0 & x & & & y & \\
& & & & -y & & -x & 0 & x & & y & \\
& & & -y & & & & -x & 0 & y & & \\
& & & & & & & & -y & 0 & x & \\
& & & & & & & -y & & -x & 0 & x \\
& & & & & & -y & & & & -x & 0
\end{bmatrix} .
\tag{7.6}
$$

This matix can be written as a tridiagonal block matrix of the form

$$
Z^2_{mn} =
\begin{bmatrix}
X & Y & & \\
-Y & X & Y & \\
& -Y & X & Y \\
& & -Y & X
\end{bmatrix} ,
\tag{7.7}
$$

where

$$
X =
\begin{pmatrix}
0 & x & 0 \\
-x & 0 & x \\
0 & -x & 0
\end{pmatrix} ,
\tag{7.8}
$$

and where

$$
Y =
\begin{pmatrix}
0 & 0 & y \\
0 & y & 0 \\
y & 0 & 0
\end{pmatrix} .
\tag{7.9}
$$

63

The general case for the block matrix, called $\mathbf{D}$, is

$$\mathbf{D} = \begin{pmatrix} X & Y & & \\ -Y & X & \ddots & \\ & \ddots & \ddots & Y \\ & & -Y & X \end{pmatrix},$$

(7.10)

where

$$X = \begin{pmatrix} 0 & x & & \\ -x & 0 & \ddots & \\ & \ddots & \ddots & x \\ & & -x & 0 \end{pmatrix},$$

(7.11)

and where

$$Y = \begin{pmatrix} & & y \\ & \iddots & \\ y & & \end{pmatrix}.$$

(7.12)

The block matrix $\mathbf{D}$ is an $m \times m$ block matrix composed of blocks which are $n \times n$ in size.

7.2   Finding the Eigenvectors and Eigenvalues

The matrix $\mathbf{D}$ is a tridiagonal antisymmetric matrix. These kind of matrices generally all have some eigenvectors and eigenvalues, and much of the following could be avoided with some clever shortcuts. However for the sake of completeness each of the eigenvalues and eigenvectors will be worked out. For a tridiagonal matrix it is usually easier to find the eigenvectors before the eigenvalues because the eigenvalue equations will appear in the form of a difference equation. The eigenvectors are those vectors $\mathbf{U}$ that satisfy

$$\mathbf{D}\mathbf{U} = \lambda\mathbf{U},$$

(7.13)

64

where $\mathbf{U}$ is a vector with elements u_ℓ with ℓ going from 1 to m. So

$$\mathbf{DU} = \begin{pmatrix} X & Y & & & \\ -Y & X & \ddots & & \\ & \ddots & \ddots & Y & \\ & & -Y & X & Y \\ & & & -Y & X \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_\ell \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} Xu_1 + Yu_2 \\ \vdots \\ -Yu_{\ell-1} + Xu_\ell + Yu_{\ell+1} \\ \vdots - Yu_{m-1} + Xu_m \end{pmatrix}. \quad (7.14)$$

The eigenvectors can be found by examining the rows of Eq. (7.14), which are given by

$$Yu_{\ell+1} + Xu_\ell - Yu_{\ell-1} = \lambda u_\ell, \quad (7.15)$$

with the boundary conditions

$$u_0 = u_{m+1} = 0. \quad (7.16)$$

A difference equation of this sort has solutions in the form $e^{i\theta\ell}$ and $e^{-i\theta\ell}$. An adjustment should be made to the form $e^{-i\theta\ell}$ to ease the math later. It is better to specify that, since ℓ is always positive, that the negative must come from the angular argument of the exponent. The eigenvectors can thus be written in the form

$$u_\ell = Ae^{i\theta\ell} + Be^{i(\pi-\theta)\ell}. \quad (7.17)$$

By using the boundary conditions, B can be found in terms of A, where

$$u_0 = 0 = A + B, \quad (7.18)$$

and so,

$$B = -A. \quad (7.19)$$

The eigenvector can now be written as

$$u_\ell = A[e^{i\theta\ell} - e^{i(\pi-\theta)\ell}]. \quad (7.20)$$

At this point either of the two terms in u_ℓ could be used to find the eigenvalues. Naturally it is easiest to use the first term as it is less complicated. It is plugged into the difference equation above to give

$$Ye^{i\theta(\ell+1)} + Xe^{i\theta\ell} - Ye^{i\theta(\ell-1)} = \lambda e^{i\theta\ell}, \tag{7.21}$$

where the A has already been cancelled from both sides. To solve for λ both sides are divided by $e^{i\theta\ell}$ to give

$$Ye^{i\theta} + X - Ye^{-i\theta} = \lambda. \tag{7.22}$$

This is easily reduced to

$$\lambda = X + 2iY\sin\theta, \tag{7.23}$$

using the relationship between exponential and trig functions.

The second boundary condition places limits on θ. It is $u_{m+1} = 0$, which gives

$$u_{m+1} = A[e^{i\theta(m+1)l} - e^{i(\pi-\theta)(m+1)}] = 0. \tag{7.24}$$

After dividing by A and moving the second term across the equal sign this becomes

$$e^{i\theta(m+1)} = e^{i(\pi-\theta)(m+1)}. \tag{7.25}$$

Then natural log can be taken on both sides to giving

$$i\theta(m+1) = i(\pi-\theta)(m+1) + 2\pi i q \tag{7.26}$$

where $q = 1, 2, 3, ..., m$. The new term $2\pi i q$ is added because the natural log is many-valued in the complex plane. This new term counts the number of full rotations about the origin in the complex plane. In general $e^{w+2i\pi} = e^w$. Making sure this term is included guarantees that every possible eigenvector is represented. From this point

it is simple to solve for θ. It is renamed θ_q as a reminder that it is many-valued.

$$\theta_q = \frac{\pi}{2} + \frac{\pi q}{m+1}. \tag{7.27}$$

These are the values of θ that are solutions to the difference equation according to the boundary conditions. The eigenvectors can now be written as

$$u_{\ell q} = A[e^{i\theta_q \ell} - e^{i(\pi - \theta_q)\ell}], \tag{7.28}$$

with eigenvalues given by

$$\lambda_q = X + 2iY \sin \theta_q. \tag{7.29}$$

To determine the value of A the eigenvectors $u_{q,\ell}$ must be normalized. This is most easily done after a change in form. First the expanded form of θ_q is plugged into the eigenvectors giving

$$u_{\ell q} = A[e^{i(\frac{\pi}{2} + \frac{\pi q}{m+1})\ell} - e^{i(\pi - (\frac{\pi}{2} + \frac{\pi q}{m+1}))\ell}], \tag{7.30}$$

which has exponents that can be split up into

$$u_{\ell q} = A[e^{\frac{i\pi\ell}{2}} e^{\frac{i\pi q\ell}{m+1}\ell} - e^{\frac{i\pi\ell}{2}} e^{-\frac{i\pi q\ell}{m+1}}]. \tag{7.31}$$

The term $e^{\frac{i\pi\ell}{2}}$ can be factored out. The remaining difference in exponentials is related to the sine function. The form of the eigenvectors is now

$$u_{\ell q} = A e^{\frac{i\pi\ell}{2}} [2i \sin \frac{\pi q \ell}{m+1}]. \tag{7.32}$$

Euler's formula reveals that $e^{\frac{i\pi\ell}{2}}$ is equal to i^{ℓ} which gives

$$u_{\ell q} = 2Ai^{\ell+1} \sin \frac{\pi q \ell}{m+1}. \tag{7.33}$$

This form is simple to normalize. It is done through the usual method of normalizing

via complex conjugates:

$$\sum_{\ell=1}^{m} u^{*}_{\ell q} u_{\ell q} = 1. \tag{7.34}$$

The values of $u_{\ell q}*$ and $u_{\ell q}$ can be plugged in producing

$$\sum_{\ell=1}^{m} \left(2A(-i^{\ell+1}) \sin \frac{\pi q \ell}{m+1} \right) \left(2A(i)^{\ell+1} \sin \frac{\pi q \ell}{m+1} \right) = 1, \tag{7.35}$$

which reduces to

$$\sum_{\ell=1}^{m} 4|A|^2 (-i)^{\ell+1} (i)^{\ell+1} \sin^2 \frac{\pi q \ell}{m+1} = 1. \tag{7.36}$$

It can be easily proved that $(-i)^{\ell+1}(i)^{\ell+1}$ always equals 1. Therefore

$$4|A|^2 \sum_{\ell=1}^{m} \sin^2 \frac{\pi q \ell}{m+1} = 1. \tag{7.37}$$

The half-angle identity for $\sin^2 x$ can be used to manipulate this further. It gives

$$2|A|^2 \sum_{\ell=1}^{m} \left(1 - \cos \frac{2\pi q \ell}{m+1} \right) = 1. \tag{7.38}$$

In general if a sum is taken over 1 it will give

$$\sum_{n=1}^{N} 1 = N. \tag{7.39}$$

Using this in the normalization yields

$$2|A|^2 \left[m - \sum_{\ell=1}^{m} \cos \left(\frac{2\pi q \ell}{m+1} \right) \right] = 1. \tag{7.40}$$

Mathematica was used to evaluate the remaining sum over cosines; it is equal to -1.

Thus

$$2|A|^2 (m+1) = 1, \tag{7.41}$$

and solving for A easily gives

$$|A| = \frac{1}{2} \sqrt{\frac{2}{m+1}}, \tag{7.42}$$

68

after rationalizing $\frac{1}{\sqrt{2}}$. Assuming A to be real, the normalized eigenvectors are

$$u_{\ell q} = \sqrt{\frac{2}{m+1}} i^{\ell+1} \sin \frac{\pi q \ell}{m+1}, \tag{7.43}$$

and the eigenvalues associated with these eigenvectors are given by

$$\lambda_q = X + 2iY \sin \theta_q, \tag{7.44}$$

where $\theta_q = \frac{\pi}{2} + \frac{\pi q}{m+1}$ and q goes from 1 to m.

7.3 Block Diagonalizing the Matrix D

To reduce the difficulty in taking the determinant of $\mathbf{D}$ it is necessary to use the eigenvectors to complete the diagonalization. This can be done through the unitary transformation

$$\mathbf{U}^\dagger \mathbf{D} \mathbf{U} = \bar{D}, \tag{7.45}$$

where $\mathbf{U}^\dagger$ is the conjugate transpose of $\mathbf{U}$ and $\bar{D}$ is the diagonalized matrix. The matrix $\mathbf{U}$ is composed of the elements $u_{\ell q}$ with ℓ increasing downwards and q increasing to the right so that

$$\mathbf{U} = \begin{pmatrix} u_{11} & \cdots & u_{1q} & \cdots & u_{1m} \\ \vdots & \ddots & & & \\ u_{\ell 1} & & u_{\ell=q,\ell=q} & & \\ \vdots & & & \ddots & \\ u_{m1} & & & & u_{mm} \end{pmatrix}. \tag{7.46}$$

The conjugate transponse is found by reflecting $\mathbf{U}$ across its main diagonal and then taking the complex conjugate of every element. The math can be check but here a shortcut is taken. It is known that when diagonalizing a matrix with its eigenvectors

69

the diagonal will have the eigenvalues corresponding to those eigenvectors.

$$
\bar{D} = \begin{pmatrix}
X + 2iY\sin(\theta_1) & & & & \\
& X + 2iY\sin(\theta_2) & & & \\
& & X + 2iY\sin(\theta_3) & & \\
& & & \ddots & \\
& & & & X + 2iY\sin(\theta_m)
\end{pmatrix}. \tag{7.47}
$$

The state function now takes the form

$$
Z_{mn}{}^2 = |\mathbf{D}| = |\bar{D}| = \prod_{q=1}^{m} \overline{D_q} \tag{7.48}
$$

where

$$
\overline{D_q} = X + 2iY\sin(\theta_q), \tag{7.49}
$$

is an $n \times n$ matrix because both X and Y are $n \times n$.

7.4 Changing the Form of $\overline{D_q}$

The block diagonal matrix $\bar{D}$ has elements $\overline{D_q}$ where each $\overline{D_q}$ has the form

$$
\overline{D_q} = \begin{pmatrix}
0 & x & & & & & 2i\sin(\theta_q)y \\
-x & 0 & x & 2i\sin(\theta_q)y & & & \\
& -x & \ddots & \ddots & & & \\
& & \ddots & \ddots & x & & \\
& 2i\sin(\theta_q)y & & -x & 0 & x & \\
2i\sin(\theta_q)y & & & & -x & 0
\end{pmatrix}. \tag{7.50}
$$

This form is still hard to work with. It would be better to diagonalize X. The process of finding the eigenvectors and eigenvalues is exactly the same as the process of diagonlizing $\mathbf{D}$ with the exception that the diagonal is filled with zeros. However

70

applying the unitary transformation is tricky because whatever actions are performed on X must also be performed on $2iY\sin(\theta_q)$ because they are both part of $\overline{D_q}$.

The normalized eigenvectors of X are

$$u_{kr} = \sqrt{\frac{2}{n+1}} i^{k+1} \sin\frac{\pi r k}{n+1}, \tag{7.51}$$

where n is the number of rows or columns in X and where k and r are analogous to the ℓ and q, respectively, from the eigenvector $u_{q\ell}$. These eigenvectors were found in the same way the eigenvectors for D were found. Similar boundary conidtions apply: $u_{0r} = u_{n+1,r} = 0$. The eigenvalues of X are

$$\lambda_r = 2ix\sin(\phi_r), \tag{7.52}$$

where

$$\phi_r = \frac{\pi}{2} + \frac{\pi r}{n+1}. \tag{7.53}$$

Now X is diagonlized by the same type of unitary transform

$$V^\dagger X V = \bar{X}, \tag{7.54}$$

where $\bar{X}$ is a diagonal matrix with its eigenvectors down the diagonal. the matrix V is made up of the elements u_{kr} and looks like

$$V = \begin{pmatrix} u_{11} & \cdots & u_{1r} & \cdots & u_{1n} \\ \vdots & \ddots & & & \\ u_{k1} & & u_{k=r,k=r} & & \\ \vdots & & & \ddots & \\ u_{n1} & & & & u_{nn} \end{pmatrix}. \tag{7.55}$$

The same unitary transformation that diagonalizes X must be applied to the matrix Y. The transform is $V^\dagger 2i\sin(\theta_q)YV = \bar{Y}$, where $\bar{Y}$ is the transformed matrix.

Written out in matrix form, this transform is

$$
\begin{pmatrix} u_{11}^* & \cdots & u_{n1}^* \\ \vdots & \ddots & \\ u_{1n}^* & & u_{nn}^* \end{pmatrix} 2i\sin(\theta_q) \begin{pmatrix} & & y \\ & \cdot^{\cdot^{\cdot}} & \\ y & & \end{pmatrix} \begin{pmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & \ddots & \\ u_{n1} & & u_{nn} \end{pmatrix} = \bar{Y} \tag{7.56}
$$

The first thing to do is to factor out $2iy\sin(\theta_q)$ and divide it to the other side giving

$$
\frac{\bar{Y}}{2iy\sin\theta_q} = \begin{pmatrix} u_{11}^* & \cdots & u_{n1}^* \\ \vdots & \ddots & \\ u_{1n}^* & & u_{nn}^* \end{pmatrix} \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix} \begin{pmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & \ddots & \\ u_{n1} & & u_{nn} \end{pmatrix}. \tag{7.57}
$$

The antidiagonal matrix of all ones on the right hand side will act to create a mirror image of any matrix it is multiplied by. If multiplied by a matrix that precedes it then it will reflect that matrix horizontially across. When operating on a matrix to the right of it that matrix is flipped vertically. The equation can now be written

$$
\frac{\bar{Y}}{2iy\sin\theta_q} = \begin{pmatrix} u_{11}^* & \cdots & u_{n1}^* \\ \vdots & \ddots & \\ u_{1n}^* & & u_{nn}^* \end{pmatrix} \begin{pmatrix} u_{n1} & & u_{nn} \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ u_{11} & \cdots & u_{1n} \end{pmatrix}, \tag{7.58}
$$

where

$$
u_{kr} = \sqrt{\frac{2}{n+1}} i^{k+1} \sin\frac{\pi r k}{n+1}. \tag{7.59}
$$

The expansion of Eq. (7.58) is large, and so only the first term in the resulting matrix is considered, which will be called t_{11}, where $t_{\alpha\beta}$ is an element in the matrix $\frac{\bar{Y}}{2iy\sin\theta_q}$. This first element is

$$
t_{11} = u_{11}^* u_{n1} + u_{21}^* u_{n-1,1} + u_{31}^* u_{n-2,1} + \ldots + u_{n1}^* u_{11}. \tag{7.60}
$$

72

The second element in the first row will be the term

$$t_{12} = u_{11}^* u_{n2} + u_{21}^* u_{n-1,2} + u_{31}^* u_{n-2,2} + \ldots + u_{n1}^* u_{12}. \tag{7.61}$$

From these the pattern for a general element can be seen. It can be written as

$$t_{r,r'} = \sum_{k=1}^{n} u_{kr}^* u_{k'r'} = \frac{2}{n+1} \sum_{k=1}^{n} (-i)^{k+1} \sin\left(\frac{\pi rk}{n+1}\right) (i)^{k'+1} \sin\left(\frac{\pi r'k'}{n+1}\right), \tag{7.62}$$

where the unconjugated matrix has been given prime indices to distinguish between the two sets of variables. The terms k and k' are related through

$$k' = n - k + 1, \tag{7.63}$$

which can be determined by examining the terms in some resulting $t_{r,r'}$ matrices in *Mathematica*. Thus $t_{r,r'}$ can be written

$$t_{r,r'} = \frac{2}{n+1} \sum_{k=1}^{n} (-i)^{k+1} \sin\left(\frac{\pi rk}{n+1}\right) (i)^{n-k+1+1} \sin\left(\frac{\pi r'(n-k+1)}{n+1}\right). \tag{7.64}$$

Further investigation of some small matrices will reveal that r and r' are not only variables but also indicate the position of the element in the resulting $t_{r,r'}$ matrix. They will be renamed as $r = \alpha$ and $r' = \beta$, since they can now be said to label elements in the resulting matrix. From this a general element is

$$t_{\alpha\beta} = \frac{2}{n+1} \sum_{k=1}^{n} (-i)^{k+1} \sin\left(\frac{\pi \alpha k}{n+1}\right) (i)^{n-k+1'+1} \sin\left(\frac{\pi \beta(n-k+1)}{n+1}\right). \tag{7.65}$$

The constant can be factored out of the sum to give the form

$$t_{\alpha\beta} = \frac{2i^{n+3}}{n+1} \sum_{k=1}^{n} (-1)^{k+1} \sin\left(\frac{\pi \alpha k}{n+1}\right) \sin\left(\frac{\pi \beta(n+1-k)}{n+1}\right). \tag{7.66}$$

7.5   The General Form of $t_{\alpha\beta}$

When investigating small lattices with terms in this form using *Mathematica*, it can be seen that all the terms are zero except those that lie on the antidiagonal from the bottom left of the matrix to the upper right. To show that this is true requires further manipulation from the form above. First the product of sines should be turned into a sum of cosines via the identity $\sin u \sin v = \frac{1}{2}[\cos(u-v) - \cos(u+v)]$. This gives

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1} \sum_{k=1}^{n} \left\{ (-1)^{k+1} \left[\cos\left(\frac{\pi\alpha k}{n+1} - \frac{\pi\beta(n+1-k)}{n+1} \right) - \cos\left(\frac{\pi\alpha k}{n+1} + \frac{\pi\beta(n+1-k)}{n+1} \right) \right] \right\},$$

$$(7.67)$$

which can be put into exponential form as

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1} \text{Re} \left[\sum_{k=1}^{n} \left\{ (-1)^{k+1} \left(e^{\frac{i(\pi\alpha k - \pi\beta(n+1-k))}{n+1}} - e^{\frac{i(\pi\alpha k + \pi\beta(n+1-k))}{n+1}} \right) \right\} \right], \qquad (7.68)$$

where Re indicates that that the real part of the sum is required so that $\text{Re}\,[e^{ix}] = \text{Re}\,[\cos(x) + i\sin(x)] = \cos(x)$. Because β has to be a whole number $e^{i\pi\beta} = e^{-i\pi\beta} = \cos(\pi\beta) = (-1)^{\beta}$ and a factor of $e^{\pm\frac{\pi\beta(n+1)}{n+1}}$ can be pulled out of each exponential in the form of $(-1)^{\beta}$.

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1} \text{Re} \left[\sum_{k=1}^{n} \left\{ (-1)^{k+1}(-1)^{\beta} \left(e^{\frac{i\pi\alpha k + i\pi\beta k}{n+1}} - e^{\frac{i\pi\alpha k - i\pi\beta k}{n+1}} \right) \right\} \right], \qquad (7.69)$$

A factor of $(-1)^{k}$ is now moved into the exponentials. The sign terms are then factored of the sum and a matrix element is now

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1}(-1)(-1)^{\beta} \text{Re} \left[\sum_{k=1}^{n} \left\{ \left(-e^{\frac{i\pi\alpha + i\pi\beta}{n+1}} \right)^{k} - \left(-e^{\frac{i\pi\alpha - i\pi\beta}{n+1}} \right)^{k} \right\} \right], \qquad (7.70)$$

The form of the exponentials are very nearly geometric progressions, which can be solved using Eq. (26) in Dwight which is given below [14].

$$\sum_{k=1}^{n} r^{k-1} = \frac{r^n - 1}{r - 1}. \tag{7.71}$$

To manipulate the exponentials in the elements into the correct form a single exponential from the product of k exponentials is factored out giving

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1}(-1)(-1)^{\beta}\text{Re}\left[\left(-e^{\frac{i\pi\alpha+i\pi\beta}{n+1}}\right)\sum_{k=1}^{n}\left(-e^{\frac{i\pi\alpha+i\pi\beta}{n+1}}\right)^{k-1} - \left(-e^{\frac{i\pi\alpha-i\pi\beta}{n+1}}\right)\sum_{k=1}^{n}\left(-e^{\frac{i\pi\alpha-i\pi\beta}{n+1}}\right)^{k-1}\right]. \tag{7.72}$$

The terms in the sum are transformed using the Dwight relationship and this becomes

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1}(-1)^{\beta+1}\text{Re}\left[\left(-e^{\frac{i\pi\alpha+i\pi\beta}{n+1}}\right)\frac{\left(-e^{\frac{i\pi\alpha+i\pi\beta}{n+1}}\right)^n - 1}{\left(-e^{\frac{i\pi\alpha+i\pi\beta}{n+1}}\right) - 1} - \left(-e^{\frac{i\pi\alpha-i\pi\beta}{n+1}}\right)\frac{\left(-e^{\frac{i\pi\alpha-i\pi\beta}{n+1}}\right)^n - 1}{\left(-e^{\frac{i\pi\alpha-i\pi\beta}{n+1}}\right) - 1}\right]. \tag{7.73}$$

Distributing the exponentials in each term will cause the first exponential in each numerator contain $\frac{n+1}{n+1}$ and reduce to siimpler exponentials as

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1}(-1)^{\beta+1}\text{Re}\left[\frac{(-1)^{n+1}\left(e^{i\pi\alpha+i\pi\beta}\right) + e^{\frac{i\pi\alpha+i\pi\beta}{n+1}}}{\left(-e^{\frac{i\pi\alpha+i\pi\beta}{n+1}}\right) - 1} - \frac{(-1)^{n+1}\left(e^{i\pi\alpha-i\pi\beta}\right) + e^{\frac{i\pi\alpha-i\pi\beta}{n+1}}}{\left(-e^{\frac{i\pi\alpha-i\pi\beta}{n+1}}\right) - 1}\right]. \tag{7.74}$$

By Euler's Formula the first term in each numerator reduces to $(-1)^{\alpha+\beta+n+1}$ and $(-1)^{\alpha-\beta+n+1}$ respectively, and pulling a minus sign out of the denominator of each term gives

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1}(-1)^{\beta+1}\text{Re}\left[-\frac{(-1)^{\alpha+\beta+n+1} + e^{\frac{i\pi(\alpha+\beta)}{n+1}}}{1 + e^{\frac{i\pi\alpha+i\pi\beta}{n+1}}} + \frac{(-1)^{\alpha-\beta+n+1} + e^{\frac{i\pi(\alpha-\beta)}{n+1}}}{1 + e^{\frac{i\pi\alpha-i\pi\beta}{n+1}}}\right]. \tag{7.75}$$

The denominator of the first term in the brackets is discontinuous at $\alpha + \beta = n + 1$, and so this geometric expansion can not be solved for instances when $\alpha + \beta = n + 1$.

Those instances will be in the next section.

Moving forward from here depends on whether the exponents of the sign terms inside the Re operation are odd or even. If $\alpha + \beta$ is even then $\alpha - \beta$ must also be even so both terms inside the square brackets have the same parity in the first term in their numerators. When $\alpha + \beta + n + 1$ and $\alpha - \beta + n + 1$ are both even, it is easy to see that $t_{\alpha\beta}$ reduces to

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1}(-1)^{\beta+1}\text{Re}\left[(-1)-(-1)\right] = 0. \tag{7.76}$$

When $\alpha + \beta + n + 1$ and $\alpha - \beta + n + 1$ are both odd the situation is more complicated. The first terms in each of the numerators equal -1, and the expression simplifies to the form

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1}(-1)^{\beta+1}\text{Re}\left[\frac{-e^{\frac{i\pi(\alpha+\beta)}{n+1}}-1}{e^{\frac{i\pi(\alpha+\beta)}{n+1}}+1} + \frac{e^{\frac{i\pi(\alpha-\beta)}{n+1}}-1}{e^{\frac{i\pi(\alpha-\beta)}{n+1}}+1}\right]. \tag{7.77}$$

From here the numerators and denominators are close to being sines and cosines. The factor $e^{-\frac{i\pi(\alpha+\beta)}{2(n+1)}}$ is multiplied both the numerator and denominator of the first exponential term, and $e^{-\frac{i\pi(\alpha-\beta)}{2(n+1)}}$ is multiplied on the top and bottom of the second term. This is not an uncommon trick and multiplying by these factors gives

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1}(-1)^{\beta+1}\text{Re}\left[-\frac{e^{\frac{i\pi(\alpha+\beta)}{2(n+1)}}-e^{-\frac{i\pi(\alpha+\beta)}{2(n+1)}}}{e^{\frac{i\pi(\alpha+\beta)}{2(n+1)}}+e^{-\frac{i\pi(\alpha+\beta)}{2(n+1)}}} + \frac{e^{\frac{i\pi(\alpha-\beta)}{2(n+1)}}-e^{-\frac{i\pi(\alpha-\beta)}{2(n+1)}}}{e^{\frac{i\pi(\alpha-\beta)}{2(n+1)}}+e^{-\frac{i\pi(\alpha-\beta)}{2(n+1)}}}\right]. \tag{7.78}$$

Now both terms in brackets have a sine in the numerator and a cosine in the denominator. Since $e^{ix} - e^{-ix} = 2i\sin x$ and $e^{ix} + e^{-ix} = 2\cos x$, $t_{\alpha\beta}$ is

$$t_{\alpha\beta} = \frac{i^{n+3}}{n+1}(-1)^{\beta+1}\text{Re}\left[-\frac{2i\sin\left(\frac{(\alpha+\beta)\pi}{2(n+1)}\right)}{2\cos\left(\frac{(\alpha+\beta)\pi}{2(n+1)}\right)} + \frac{2i\sin\left(\frac{(\alpha-\beta)\pi}{2(n+1)}\right)}{2\cos\left(\frac{(\alpha-\beta)\pi}{2(n+1)}\right)}\right]. \tag{7.79}$$

Because each term inside the Re operation is purely imaginary this is identically 0. Therefore the matrix $\frac{\bar{Y}}{2iy\sin\theta_q}$ will be zero everywhere except those elements where

$\alpha + \beta = n + 1$. These elements in the matrix could not be determined through this method because they cause a singularity in a denominator. They will be determined in the next secion

7.6 Finding $t_{\alpha\beta}$ When $\alpha = \beta$

To show the behavior of the elements of $t_{\alpha\beta}$ when $\alpha = \beta$, one must backtrack to the general form

$$(t_{\alpha\beta})_{\alpha+\beta=n+1} = \frac{2i^{n+3}}{n+1} \sum_{k=1}^{n} (-1)^{k+1} \sin\left(\frac{\pi\alpha k}{n+1}\right) \sin\left(\frac{\pi\beta(n+1-k)}{n+1}\right), \qquad (7.80)$$

and plug in $\alpha = n + 1 - \beta$ to get

$$(t_{\alpha\beta})_{\alpha+\beta=n+1} = \frac{2i^{n+3}}{n+1} \sum_{k=1}^{n} (-1)^{k+1} \sin\left(\frac{\pi(n+1-\beta)k}{n+1}\right) \sin\left(\frac{\pi\beta(n+1-k)}{n+1}\right). \tag{7.81}$$

It is useful to break the arguments of the sines into pieces to look for identities. This gives

$$(t_{\alpha\beta})_{\alpha+\beta=n+1} = \frac{2i^{n+3}}{n+1} \sum_{k=1}^{n} (-1)^{k+1} \sin\left(\pi k - \frac{\pi\beta k}{n+1}\right) \sin\left(\pi\beta - \frac{\pi\beta k}{n+1}\right). \tag{7.82}$$

The trigonometric identity $\sin(u-v) = \sin u \cos u - \cos u \sin v$ allows this to be written as

$$(t_{\alpha\beta})_{\alpha+\beta=n+1} = \frac{2i^{n+3}}{n+1} \sum_{k=1}^{n} (-1)^{k+1} \left[\sin(\pi k) \cos\left(\frac{\pi\beta k}{n+1}\right) - \cos(\pi k) \sin\left(\frac{\pi\beta k}{n+1}\right)\right]$$
$$\times \left[\sin(\pi\beta) \cos\left(\frac{\pi\beta k}{n+1}\right) - \cos(\pi\beta) \sin\left(\frac{\pi\beta k}{n+1}\right)\right]. \tag{7.83}$$

Because k and β are both integers this reduces to

$$(t_{\alpha\beta})_{\alpha+\beta=n+1} = \frac{2i^{n+3}}{n+1} \sum_{k=1}^{n} (-1)^{k+1} \left[(-1)^{k+1} \sin\left(\frac{\pi\beta k}{n+1}\right) \right] \left[(-1)^{\beta+1} \sin\left(\frac{\pi\beta k}{n+1}\right) \right]$$

$$= \frac{2i^{n+3}}{n+1} (-1)^{\beta-1} \sum_{k=1}^{n} \sin^2\left(\frac{\pi\beta k}{n+1}\right),$$

$$(7.84)$$

where the change in signs occurs via $(-1)^{k+1}(-1)^{k+1}(-1)^{\beta+1} = (-1)^{2k}(-1)^3(-1)^{\beta} = (-1)^{\beta+1} = (-1)^{\beta-1}$. Now the trig identity $2\sin^2(u) = 1 - \cos(2u)$ is used to give

$$(t_{\alpha\beta})_{\alpha+\beta=n+1} = \frac{i^{n+3}}{n+1} (-1)^{\beta-1} \sum_{k=1}^{n} \left[1 - \cos\left(\frac{2\pi\beta k}{n+1}\right) \right], \qquad (7.85)$$

which has been seen before in solving the normalization of $u_{q\ell}$, and *Mathematica* was used to determine that this sum is $n+1$, therefore

$$(t_{\alpha\beta})_{\alpha+\beta=n+1} = i^{n+3}(-1)^{\beta-1}. \qquad (7.86)$$

Recalling that $i^n = i^{n \mod 4}$ allows this to be written

$$(t_{\alpha\beta})_{\alpha+\beta=n+1} = i^{n-1}(-1)^{\beta-1}. \qquad (7.87)$$

These are the terms that appear on the antidiagonal in $\bar{Y}$. The terms anywhere other than the antidiagonal are zero. Recall that β indicates the column of the element in the $n \times n$ matrix $\bar{Y}$. The matrix can finally be written as

$$\bar{Y} = 2iy\sin(\theta_q) \begin{pmatrix} & & & i^{n-1}(-1)^{n-1} \\ & & \iddots & \\ & i^{n-1}(-1)^{2-1} & & \\ i^{n-1}(-1)^{1-1} & & & \end{pmatrix}. \qquad (7.88)$$

The matrix $\overline{D_q} = X + 2i\sin(\theta_q)Y$ can now be expressed as $\overline{D_q} = \bar{X} + 2i\sin(\theta_q)\bar{Y}$,

78

where $\bar{X}$ is the diagonalized matrix X with its eigenvalues on the diagonal and where $\bar{Y}$ is the antidiagonal matrix in Eq. (7.88). Then, written explicitly, $\overline{D_q}$ is

$$\overline{D_q} = \begin{pmatrix} 2ix\sin(\phi_1) & & & & & (-1)^{n-1}2i^n y\sin(\theta_q) \\ & 2ix\sin(\phi_2) & & & (-1)^{n-2}2i^n y\sin(\theta_q) & \\ & & \ddots & \reflectbox{$\ddots$} & & \\ & & \reflectbox{$\ddots$} & \ddots & & \\ & (-1)^1 2i^n y\sin(\theta_q) & & & 2ix\sin(\phi_{n-1}) & \\ (-1)^0 2i^n y\sin(\theta_q) & & & & & 2ix\sin(\phi_n) \end{pmatrix} .$$

$$(7.89)$$

Fisher called this matrix a cruciform matrix, with nonzer elements only on the diagonal and antidiagonal.

7.7 Defining the Partition Function

As stated before, the state function is found by

$$Z_{mn}{}^2 = |\mathbf{D}| = |\bar{D}| = \prod_{q=1}^{m} |\overline{D_q}|, \tag{7.90}$$

where $\mathbf{D}$ is the tridiagonal block matrix, $\bar{D}$ is the diagonalized block matrix, and the $\overline{D_q}$ are the cruciform matrices above which occupy the diagonal of $\bar{D}$. The determinant of $\overline{D_q}$ can be determined by expanding by the first row of the matrix and then in each minor expanding by the last row. Then each new minor is expanded by the first row and each of the resulting minors by their bottom rows. This repetition creates $\lfloor \frac{1}{2}(n+1)\rfloor$ factors of the type $(d_{11}d_{nn} - d_{1n}d_{n1})$. The brackets $\lfloor x \rfloor$ indicate the floor function, which gives the highest interger lower than x.

It is not immediately obvious that this is the correct expansion. To get a better

idea consider the general cruciform matrix $\bar{C}$

$$|\bar{C}| = \begin{pmatrix} d_{11} & & & & & d_{1,n} \\ & d_{22} & & & d_{2,n-1} & \\ & & \ddots & \reflectbox{$\ddots$} & & \\ & & \reflectbox{$\ddots$} & \ddots & & \\ & d_{n-1,2} & & & d_{n-1,n-1} & \\ d_{n1} & & & & & d_{nn} \end{pmatrix}. \tag{7.91}$$

The expansion begins with the first row as

$$|\bar{C}| = d_{11} \begin{vmatrix} d_{22} & & d_{2,n-1} & 0 \\ & \ddots & \reflectbox{$\ddots$} & \\ & \reflectbox{$\ddots$} & \ddots & \\ d_{n-1,2} & & d_{n-1,n-1} & \\ 0 & & & d_{nn} \end{vmatrix}$$

$$+ (-1)^{n-1} d_{1,n} \begin{vmatrix} 0 & d_{22} & & d_{2,n-1} \\ & & \ddots & \reflectbox{$\ddots$} \\ & & \reflectbox{$\ddots$} & \ddots \\ & d_{n-1,2} & & d_{n-1,n-1} \\ d_{n1} & & & 0 \end{vmatrix}, \tag{7.92}$$

where the only two elements in row one are multiplied by their minors. A factor of $(-1)^{n-1}$ is necessary to get the correct sign since it depends on the number of rows in the matrix. Then the remaining minors are expanded by their bottom rows. Note

that this minor has $n - 1$ rows, since its first row was eliminated,

$$|\bar{C}| = d_{11}d_{nn} \begin{vmatrix} d_{22} & & d_{2,n-1} \\ & \ddots & \iddots \\ & \iddots & \ddots \\ d_{n-1,2} & & d_{n-1,n-1} \end{vmatrix}$$

$$+ (-1)^{n-2}(-1)^{n-1}d_{1n}d_{n1} \begin{vmatrix} d_{22} & & d_{2,n-1} \\ & \ddots & \iddots \\ & \iddots & \ddots \\ d_{n-1,2} & & d_{n-1,n-1} \end{vmatrix}. \qquad (7.93)$$

Notice that the second term in this expansion d_{n1} was in the $n - 1$ row and first column and so there is an additional term of $(-1)^{(}n - 2)$. In general, when taking a determinant, a negative sign is included with any element whose row and column positions sum to an odd number. The first term in the expansion of $|\bar{C}|$ will always be positive because the subscripts of the elements on the main diagonal always sum to an even number. The sums of the row and column numbers of the elements on the antidiagonal will depend on whether n is odd or even. However if the matrix $\bar{C}$ has an odd number of rows and columns, then the minor obtained by expanding by a row will have an even number of rows and columns and vice versa. Therefore either the element in the position d_{1n} will have a negative sign or the one in position d_{n1} will have a negative sign but they will not both have one. The sign of the second term is then always negative, since $(-1)^{n-1}(-1)^{n-2}$=-1. After factoring out the remaining

determinant, the expansion by the first row can be written as

$$|\bar{C}| = (d_{11}d_{nn} - d_{1n}d_{n1}) \begin{vmatrix} d_{22} & & & d_{2,n-1} \\ & \ddots & \cdot^{\cdot^{\cdot}} & \\ & \cdot^{\cdot^{\cdot}} & \ddots & \\ d_{n-1,2} & & & d_{n-1,n-1} \end{vmatrix}. \tag{7.94}$$

The remaining determinant is expanded in the same way. This process is repeated until the matrix is fully expanded. The only thing to be aware of at this point is that if the matrix has odd dimension, the final remaining determinant will be of a single term.

If the cruciform matrix related to the state function is expanded in this manner it is found that first iteration is

$$|\overline{D_q}| = \{[2ix\sin(\phi_1)][2ix\sin(\phi_n)] - [(-1)^{n-1}2i^n y\sin(\theta_q)][(-1)^0 2i^n y\sin(\theta_q)]\} \times$$

$$\times \begin{vmatrix} 2ix\sin(\phi_2) & & & (-1)^{n-2}2i^n y\sin(\theta_q) \\ & \ddots & \cdot^{\cdot^{\cdot}} & \\ & \cdot^{\cdot^{\cdot}} & \ddots & \\ (-1)^1 2i^n y\sin(\theta_q) & & & 2ix\sin(\phi_{n-1}) \end{vmatrix}. \tag{7.95}$$

A second iteration gives

$$|\overline{D_q}| = \{[2ix\sin(\phi_1)][2ix\sin(\phi_n)] - [(-1)^{n-1}2i^n y\sin(\theta_q)][(-1)^0 2i^n y\sin(\theta_q)]\}$$

$$\times \{[2ix\sin(\phi_2)][2ix\sin(\phi_{n-1})] - [(-1)^{n-2}2i^n y\sin(\theta_q)][(-1)^1 2i^n y\sin(\theta_q)]\}$$

$$\times \begin{vmatrix} 2ix\sin(\phi_3) & & & (-1)^{n-3}2i^n y\sin(\theta_q) \\ & \ddots & \cdot^{\cdot^{\cdot}} & \\ & \cdot^{\cdot^{\cdot}} & \ddots & \\ (-1)^2 2i^n y\sin(\theta_q) & & & 2ix\sin(\phi_{n-2}) \end{vmatrix},$$

$$\tag{7.96}$$

82

which is enough to see the pattern in the terms in the resulting product. In general it can be said that

$$|\overline{D_q}| = \prod_{s=1}^{\lfloor \frac{n}{2} \rfloor} \left\{ [2ix \sin(\phi_s)][2ix \sin(\phi_{n+1-s})] - (-1)^{n-s}(-1)^{s-1} [2i^n y \sin(\theta_q)]^2 \right\}$$

$$\times \begin{cases} 1, & n \text{ even} \\ \left\{ 2ix \sin\left(\phi_{\frac{(n+1)}{2}}\right) + (-1)^{n-\frac{(n+1)}{2}} 2i^n y \sin(\theta_q) \right\}, & n \text{ odd} \end{cases} \quad , \quad (7.97)$$

where s is a placeholder positive integer for now. When n is odd, the center term of the matrix is multiplied after the product is taken. This is what the last term is in the odd expansion.

Ideally a product is needed that is a product of r over the entire matrix n, but the extra term in the middle for odd matrices disallows a single product expression over r. Recall that r was the original subscript to ϕ and that s is just a placeholder. To make the change to a product over r, the expression must be rewritten. The factor of $i^{(n-1)^2}$ pulled out of the square in the second term. Then a factor of i^2 is pulled out of all of the terms. This leaves

$$|\overline{D_q}| = \prod_{s=1}^{\lfloor \frac{n}{2} \rfloor} (i^2) \left\{ [2x \sin(\phi_s)][2x \sin(\phi_{n+1-s})] - (-1)^{n-1}(i)^{(n-2)^2} [2iy \sin(\theta_q)]^2 \right\}$$

$$\times \begin{cases} 1, & n \text{ even} \\ \left\{ 2ix sin\left(\phi_{\frac{(n+1)}{2}}\right) + (-1)^{\frac{(n-1)}{2}} 2i^n y \sin(\theta_q) \right\}, & n \text{ odd} \end{cases} \quad . \quad (7.98)$$

Because of the behavior of i when exponentiated the sign terms inside the bracket reduce via $(-1)^{n-1}(i)^{(n-2)^2} = (-1)^{n-1}(i)^{2n-4} = (-1)^{n-1}(i)^{2n} = (-1)^{n-1}(-1)^n = -1$.

The sign wil always be -1 since $i^2 = -1$. Therefore $\overline{D_q}$ can be written as

$$|\overline{D_q}| = (-1)^{\lfloor \frac{n}{2} \rfloor} \prod_{s=1}^{\lfloor \frac{n}{2} \rfloor} \left\{ [2x \sin (\phi_s)][2x \sin (\phi_{n+1-s})] + [2iy \sin (\theta_q)]^2 \right\}$$

$$\times \begin{cases} 1, & n \text{ even} \\ \left\{ 2ix\sin \left(\phi_{\frac{(n+1)}{2}} \right) + (-1)^{\frac{(n-1)}{2}} 2i^n y \sin (\theta_q) \right\}, & n \text{ odd} \end{cases} \tag{7.99}$$

where i^2 has been moved outside of the product.

The next step in working toward a sum over r is to factor the first bracketed term. This yields

$$|\overline{D_q}| = (-1)^{\lfloor \frac{n}{2} \rfloor} \prod_{s=1}^{\lfloor \frac{n}{2} \rfloor} \left\{ [2x \sin (\phi_s) + 2iy \sin (\theta_q)] [2x \sin (\phi_{n+1-s}) + 2iy \sin (\theta_q)] \right\}$$

$$\times \begin{cases} 1, & n \text{ even} \\ \left\{ 2ix\sin \left(\phi_{\frac{(n+1)}{2}} \right) + (-1)^{\frac{(n-1)}{2}} 2i^n y \sin (\theta_q) \right\}, & n \text{ odd} \end{cases}$$

$$\tag{7.100}$$

where s skips over $\frac{n+1}{2}$ in the odd case for the moment. To arrive at this factorization it is necessary to realize, from Eq. (7.53), that $\sin(\phi_s) = -\sin(\phi_{n+1-s})$ which makes the cross terms cancel. This is proved below. Since $\phi_s = \frac{\pi}{2} + \frac{\pi s}{n+1}$,

$$\sin(\phi_{n+1-s}) = \sin \left(\frac{\pi}{2} + \frac{\pi(n+1-s)}{n+1} \right). \tag{7.101}$$

After splitting the fraction up this is

$$\sin(\phi_{n+1-s}) = \sin \left(\frac{\pi}{2} + \frac{\pi(n+1)}{n+1} - \frac{\pi s}{n+1} \right), \tag{7.102}$$

which becomes

$$\sin(\phi_{n+1-s}) = \sin \left(\frac{\pi}{2} + \pi - \frac{\pi s}{n+1} \right). \tag{7.103}$$

84

Through the trig identity $\sin(u \pm v) = \sin u \cos v \pm \cos u \sin v$ it can be said that

$$\sin(\phi_{n+1-s}) = \sin\left(\frac{\pi}{2}\right)\cos\left(\pi - \frac{\pi s}{n+1}\right) + \cos\left(\frac{\pi}{2}\right)\sin\left(\pi - \frac{\pi s}{n+1}\right). \tag{7.104}$$

This is quickly reduced to

$$\sin(\phi_{n+1-s}) = \cos\left(\pi - \frac{\pi s}{n+1}\right). \tag{7.105}$$

Then by $\cos(u \pm v) = \cos u \cos v \mp \sin u \sin v$ the relationship is expressed as

$$\sin(\phi_{n+1-s}) = \cos(\pi)\cos\left(\frac{\pi s}{n+1}\right) + \sin(\pi)\sin\left(\frac{\pi s}{n+1}\right), \tag{7.106}$$

which reduces to

$$\sin(\phi_{n+1-s}) = -\cos\left(\frac{\pi s}{n+1}\right). \tag{7.107}$$

A factor of $\sin\left(\frac{\pi}{2}\right)$ is introduced and then the relationship $\sin(u \pm v) = \sin u \cos v \pm \cos u \sin v$ is used in reverse giving

$$\sin(\phi_{n+1-s}) = -\sin\left(\frac{\pi}{2} + \frac{\pi s}{n+1}\right) = \sin\left(\phi_s\right). \tag{7.108}$$

This relationship causes all of the cross terms to cancel out and allows the factoring of the first term in the product of $\overline{D_q}$.

This relationship is also useful for reducing the extra term in the odd expansion. In the case where $s = \frac{n+1}{2}$ it is shown here that $\sin(\phi_s) = 0$.

$$\sin \phi_s = \sin(\phi_{\frac{n+1}{2}}) = -\sin\left(\frac{\pi}{2} + \frac{\pi}{n+1}\frac{n+1}{2}\right) = \sin(\pi) = 0. \tag{7.109}$$

The term $\left\{2ix\sin\left(\phi_{\frac{(n+1)}{2}}\right) + (-1)^{\frac{(n-1)}{2}}2i^n y \sin\left(\theta_q\right)\right\}$ can now be easily manipulated. The first term is 0 by the relationship above. Now the factor $(-1)^{\frac{(n-1)}{2}} i^n$ in front of the $\sin(\theta_q)$ needs to be investigated. First say

$$(-1)^{\frac{(n-1)}{2}} i^n = (i^2)^{\frac{(n-1)}{2}} i^n = (i)^{2n-n-1} i^n = (i)^{2n}(i)^{-1}. \tag{7.110}$$

Then becuase $i^{-1} = -i$, and because this is the case where n is odd, $i^{2n} = -1$ the above product reduces to

$$(-1)^{\frac{(n11)}{2}} i^n = i. \qquad (7.111)$$

Now the leftover term in the odd expansion can be written as $2iy \sin(\theta_q)$. The product can be written as a single product from $r = 1$ to n instead of being split into cases and instead of going to $\lfloor \frac{n}{2} \rfloor$.

$$|\overline{D_q}| = (-1)^{\lfloor \frac{n}{2} \rfloor} 2^n \prod_{r=1}^{n} [x \sin(\phi_r) + iy \sin(\theta_q)], \qquad (7.112)$$

where a 2 has been factored out of each term in the product, giving a $2n$ out front.. The two different sines, $\sin(\phi_r)$ and $\sin(\phi_{n+1-r})$ will take opposite signs as r goes from 1 to $\lfloor \frac{n}{2} \rfloor$. In this form of $\overline{D_q}$ the term $\sin(\phi_r)$ will take all of the values that were formerly split between the combination of the two sines because the product now goes to n.

The arguments in $\overline{D_q}$ can be reduced slightly. By the trig identity $\sin(u \pm v) = \sin u \cos v \pm \cos u \sin v$, any angle in the form $\left(\frac{\pi}{2} + \frac{\pi s}{n+1}\right)$ can be written as $\cos\left(\frac{\pi r}{n+1}\right)$. Combining this with the definition of the state function in Eq. (7.90) gives

$$Z_{mn}{}^2 = \prod_{q=1}^{m} (-1)^{\lfloor \frac{n}{2} \rfloor} 2^n \prod_{r=1}^{n} \left[x \cos\left(\frac{\pi r}{n+1}\right) + iy \cos\left(\frac{\pi q}{m+1}\right) \right]. \qquad (7.113)$$

After pulling the terms between the products out, the state function is

$$Z_{mn}{}^2 = (-1)^{m \lfloor \frac{n}{2} \rfloor} 2^{mn} \prod_{q=1}^{m} \prod_{r=1}^{n} \left[x \cos\left(\frac{\pi r}{n+1}\right) + iy \cos\left(\frac{\pi q}{m+1}\right) \right]. \qquad (7.114)$$

CHAPTER 8

ANALYSIS OF THE PARTITION FUNCTION

Eq. (7.114) is the exact partition function of a finite $m \times n$ lattice as a product of mn terms. If both n and m are odd, then there is a term in the product for which $r = \frac{1}{2}(n+1)$ and $q = \frac{1}{2}(m+1)$, which means that both the x and y parts vanish identically. The whole product is then multiplied by zero correctly reflecting the fact that a lattice with an odd number of points can not be occupied solely by dimers. Therefore the lattice is required to have an even number of points and either m or n or both must be even.

Because eventually the entropy will be required, it would be best to have a parition function that is not squared. If it is assumed that n is even, the partition function can be manipulated into a more useable form. Holding n even instead of m is an arbitrary choice.

8.1 A Change in Form

Moving the 2 back inside the products makes the partition function take the form

$$Z_{mn}{}^2 = (-1)^{m\lfloor \frac{n}{2} \rfloor} \prod_{q=1}^{m} \prod_{r=1}^{n} \left[2x \cos\left(\frac{\pi r}{n+1} \right) + 2iy \cos\left(\frac{\pi q}{m+1} \right) \right]. \tag{8.1}$$

In the case where m is odd, and focusing only on the product over q, it is possible to write this as

$$Z_{\substack{n=\text{even} \\ m=\text{odd}}}{}^2 = (-1)^{m\lfloor \frac{n}{2} \rfloor} \prod_{q=1}^{m} \prod_{r=1}^{n} \left[a + 2iy \cos\left(\frac{\pi q}{m+1} \right) \right], \tag{8.2}$$

87

where $a = 2x\cos\left(\frac{\pi r}{n+1}\right)$. Comparing the factor with $q = 1$ to the factor with $q = m-1$ in the cosine shows that one of those terms is the negative of the other via

$$\cos\left(\frac{\pi m}{m+1}\right) = \cos\left(\pi - \frac{\pi}{m+1}\right) = -\cos\left(\frac{\pi}{m+1}\right) \tag{8.3}$$

This relationship holds for the pair $q = 2$ and $q = m - 1$ and so on. Thus for each term in the product there will appear a complex conjugate of that term. The only exception is the unpaired term where $\cos\left(\frac{\pi m}{m+1}\right) = \cos\left(\frac{\pi}{2}\right) = 0$, and so that term must be written separately. With the terms paired off, the product can be cut in half.

$$Z_{\substack{n=\text{even}\\m=\text{odd}}}^2 = (-1)^{m\lfloor\frac{n}{2}\rfloor}\left\{\prod_{r=1}^{n}\prod_{q=1}^{\lfloor\frac{m}{2}\rfloor}\left[a^2 - \left(2iy\cos\left(\frac{\pi q}{m+1}\right)\right)^2\right]\right\}\left\{\prod_{r=1}^{n}(a+0)\right\}. \tag{8.4}$$

Plugging for a gives

$$Z_{\substack{n=\text{even}\\m=\text{odd}}}^2 = (-1)^{m\lfloor\frac{n}{2}\rfloor}\left\{\prod_{r=1}^{n}\prod_{q=1}^{\lfloor\frac{m}{2}\rfloor}\left[4x^2\cos^2\left(\frac{\pi r}{n+1}\right) + 4y^2\cos^2\left(\frac{\pi q}{m+1}\right)\right]\right\}\left\{\prod_{r=1}^{n}2x\cos\left(\frac{\pi r}{n+1}\right)\right\}. \tag{8.5}$$

The last product reduces to $x^n(-1)^{\frac{n}{2}}$ via the identity $\prod_{k=1}^{n-1}\cos\frac{k\pi}{n} = \frac{\sin\frac{\pi n}{2}}{2^{n-1}}$, as can be seen here,

$$\prod_{r=1}^{n}2x\cos\left(\frac{\pi r}{n+1}\right) = 2^n x^n\prod_{r=1}^{n}\cos\left(\frac{\pi r}{n+1}\right) = 2^n x^n\frac{\sin\frac{\pi(n+1))}{2}}{2^n} = x^n(-1)^{\frac{n}{2}}. \tag{8.6}$$

The sign factor $(-1)^{\frac{n}{2}}$ may be combined with the already existing sign factor to cancel as $(-1)^{m\lfloor\frac{n}{2}\rfloor}(-1)^{\frac{n}{2}} = 1$ when n is even and m is odd. Then

$$Z_{\substack{n=\text{even}\\m=\text{odd}}}^2 = x^n\prod_{r=1}^{n}\prod_{q=1}^{\lfloor\frac{m}{2}\rfloor}\left[4x^2\cos^2\left(\frac{\pi r}{n+1}\right) + 4y^2\cos^2\left(\frac{\pi q}{m+1}\right)\right]. \tag{8.7}$$

Using the same sort of expansion over the n as was used for the product over m gives

$$Z_{\substack{n=\text{even}\\m=\text{odd}}}^2 = x^n\prod_{r=1}^{\frac{n}{2}}\prod_{q=1}^{\lfloor\frac{m}{2}\rfloor}\left[4x^2\cos^2\left(\frac{\pi r}{n+1}\right) + 4y^2\cos^2\left(\frac{\pi q}{m+1}\right)\right]^2. \tag{8.8}$$

As before, each cosine was paired with its conjugate. However, since the cosines in this case were already squared the negative signs vanished. Also, since n is taken to be even, there is no point at which the argument of the cosine is $\frac{\pi}{2}$, and so there is no extra unpaired term, as there was when the product over m was cut in half. Now the square root is taken on both sides and the square cosines are manipulated by the the half-angle identity $\cos^2 x = \frac{1}{2}(1 + \cos 2x)$ to yield

$$Z_{\substack{n=\text{even} \\ m=\text{odd}}} = 2^{\frac{n}{2}\lfloor\frac{m}{2}\rfloor} x^{\frac{n}{2}} \prod_{r=1}^{\frac{n}{2}} \prod_{q=1}^{\lfloor\frac{m}{2}\rfloor} \left[x^2 + y^2 + x^2 \cos\left(\frac{2\pi r}{n+1}\right) + y^2 \cos\left(\frac{2\pi q}{m+1}\right) \right]. \quad (8.9)$$

In the case for which m is even, the solution follows the same process. The key differences are that the product over m will not have an unpaired term when m is even and the sign term in Z_{mn} reduces identically to 1. Thus the partition function is

$$Z_{mn} = 2^{\frac{n}{2}\lfloor\frac{m}{2}\rfloor} \prod_{r=1}^{\frac{n}{2}} \prod_{q=1}^{\lfloor\frac{m}{2}\rfloor} \left[x^2 + y^2 + x^2 \cos\left(\frac{2\pi r}{n+1}\right) + y^2 \cos\left(\frac{2\pi q}{m+1}\right) \right] \begin{cases} \times 1 & m \text{ even} \\ \times x^{\frac{n}{2}} & m \text{ odd} \end{cases}. \quad (8.10)$$

8.2   The Free Energy

The entropy is found by $S_{site} = -\frac{\partial}{\partial T}\Omega_G$, where $\Omega_G = -k_B T \ln[Z_G]$. The natural log of Z_G is the grand canoncial potential of the system. In the case of the system used here, where the chemical potential is a constant, it will be a Gibbs free energy called G. The free energy is needed to find the entropy and so it is derived in this section. For a large lattice, the limiting behavior of the partition function is

$$Z_{mn}(x, y) \sim [Z(x, y)]^{mn}. \quad (8.11)$$

Upon taking the natrual log of both sides this becomes

$$\ln[Z(x,y)] = \lim_{m,n\to\infty} \frac{1}{mn} \ln[Z_{mn}(x,y)], \tag{8.12}$$

in the limit as m and n go to infinity. If a factor of $-k_B T$ were included then this would be the free energy per lattice site which is written as

$$G_{site} = -k_B T \lim_{m,n\to\infty} \frac{1}{mn} \ln[Z_{mn}(x,y)], \tag{8.13}$$

For this limit it is assumed that m is even. For a large lattice adding or subtracting a single row will have very little overall effect on the statistical mechanics. Now Eq. (8.10) is expressed in the limit of large lattices as

$$G_{site} = -k_B T \lim_{m,n\to\infty} \frac{1}{mn} \ln\left[2^{\frac{n}{2}\frac{m}{2}} \prod_{r=1}^{\frac{n}{2}} \prod_{q=1}^{\frac{m}{2}} x^2 + y^2 + x^2 \cos\left(\frac{2\pi r}{n+1}\right) + y^2 \cos\left(\frac{2\pi q}{m+1}\right) \right]. \tag{8.14}$$

The logarithm of a product can be written as the sum of logarithms so this is

$$G_{site} = -k_B T \lim_{m,n\to\infty} \frac{1}{mn} \left[\frac{mn}{4} \ln 2 + \ln \prod_{r=1}^{\frac{n}{2}} \prod_{q=1}^{\frac{m}{2}} x^2 + y^2 + x^2 \cos\left(\frac{2\pi r}{n+1}\right) + y^2 \cos\left(\frac{2\pi q}{m+1}\right) \right]. \tag{8.15}$$

In the limit as m goes to infinity, $m+1 \to m$. The arguments of the cosine terms can be rewritten. Also since the log of a product is the sum of logs, this can be written as

$$G_{site} = -k_B T \lim_{m,n\to\infty} \frac{1}{mn} \left[\frac{mn}{4} \ln 2 + \sum_{r=1}^{\frac{n}{2}} \sum_{q=1}^{\frac{m}{2}} \ln\left(x^2 + y^2 + x^2 \cos\left(\frac{2\pi r}{n}\right) + y^2 \cos\left(\frac{2\pi q}{m}\right) \right) \right]. \tag{8.16}$$

Infinite sums may be converted to integrals via the rule from basic calculus that

$$\int_a^b f(x)dx = \sum_{i=1}^k f(x_i^*)\Delta x, \tag{8.17}$$

90

where $x_i^* = a + i\Delta x$ and $\Delta x = \frac{b-a}{k}$, in the limit as $k \to \infty$. Working first on the sum over r as given by

$$\lim_{n\to\infty} \sum_{r=1}^{\frac{n}{2}} \ln \left(x^2 + y^2 + x^2 \cos \left(\frac{2\pi r}{n} \right) + y \cos \left(\frac{2\pi q}{m} \right) \right),$$ (8.18)

the limts of the resulting integral can be found by plugging in the limits of the sum. When $r = 1$, $\cos \left(\frac{2\pi r}{n} \right)$ is the same as $\cos(0)$, and so the lower limit of the integral is 0. By the same argument the upper limit is $b = \pi$. Combining Eqs. (8.17) and (8.18) with the information so far gives

$$\int_0^\pi f(\alpha)d\alpha = \lim_{n\to\infty} \sum_{r=1}^{\frac{n}{2}} f \left(0 + \frac{\pi - 0}{\frac{n}{2}} r \right) \frac{\pi - 0}{\frac{n}{2}},$$ (8.19)

where $\alpha = \frac{2\pi r}{n}$. This can be further reduced to

$$\int_0^\pi f(\alpha)d\alpha = \lim_{n\to\infty} \sum_{r=1}^{\frac{n}{2}} f \left(\frac{2\pi r}{n} \right) \frac{2\pi}{n},$$ (8.20)

Since the only argument inside the natural log of Eq. (8.18) that depends on r is the same argument $\frac{2\pi r}{n}$ the above equation allows a conversion of Eq. (8.18) from a sum to an integral at this point. The some is solved for to obtain

$$\lim_{m\to\infty} \sum_{r=1}^{\frac{n}{2}} f \left(\frac{2\pi r}{n} \right) = \frac{n}{2\pi} \int_0^\pi f(\alpha)d\alpha = \frac{n}{2\pi} \int_0^\pi \ln \left(x^2 + y^2 + x^2 \cos\alpha + y^2 \cos \left(\frac{2\pi q}{m} \right) \right)$$ (8.21)

This can now be plugged back into Eq. (8.16). The same process is applied to the sum over q and the natural log of the partition function can be written

$$G_{site} = -k_B T \frac{1}{mn} \left[\frac{mn}{4} \ln 2 + \frac{n}{2\pi} \frac{m}{2\pi} \int_0^\pi \int_0^\pi \ln \left(x^2 + y^2 + x^2 \cos\alpha + y^2 \cos\beta \right) d\alpha d\beta \right],$$ (8.22)
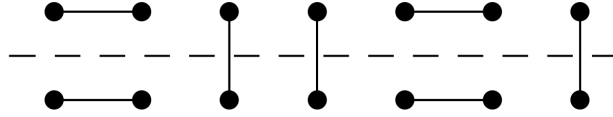
Fig. 21.: A two row lattice that will be cut in half along the dashed line to produce a one-dimensional lattice. Afterwards, the y-dimers can be considered to be monomers and pairs of x-dimers to be single x-dimers.

or, distributing $\frac{1}{mn}$,

$$G_{site} = -k_B T \frac{\ln 2}{4} + \frac{2}{(2\pi)^2} \int_0^\pi \int_0^\pi \ln \left(x^2 + y^2 + x^2 \cos \alpha + y^2 \cos \beta \right) d\alpha d\beta. \quad (8.23)$$

This is as far as this paper will treat the close packed limit of both x and y. Fisher has an excellent analysis of the results that will not be reiterated here. Instead the case of x-dimers, y-dimers and monomers on the edge of a strand of DNA will be considered.

8.3 One-Dimensional Partition Function

While there are more simple ways to to determine the statistical mechanics of one-dimensional lattices, the preceding work is more mathematically interesting and may be valuable in moving on to polymers longer than dimers. Recall that in this thesis the interest lies charge inversion. Specifically, inversion caused by dimers parallel to the edge of the DNA strand and perpendicular to the DNA strand. If a two-dimensional lattice were to be cut in half down the middle, as seen in Fig. 21, it would be a good representation of this kind of system.

If the lattice were cut horizontally across the middle, then each vertically paired x-dimer in the two row lattice will appear as a single x-dimer in the one row lattice. This can be represented mathematically by replacing each x Eq. (8.10) by $x^{(1/2)}$. For this section, each y-dimer would appear as a monomer, which is a change from y to

z in the equation. When $m = 2$

$$Z_n(x, z) = 2^{\frac{n}{2}} \prod_{r=1}^{\frac{n}{2}} \prod_{q=1}^{1} \left[x + z^2 + x \cos \left(\frac{2\pi r}{n+1} \right) + z^2 \cos \left(\frac{2\pi q}{3} \right) \right]. \tag{8.24}$$

Since the product over q has only one factor it can be easily evaluated, giving

$$Z_n(x, z) = 2^{\frac{n}{2}} \prod_{r=1}^{\frac{n}{2}} \left[x + z^2 + x \cos \left(\frac{2\pi r}{n+1} \right) - \frac{z^2}{2} \right]. \tag{8.25}$$

Combining terms in the $2^{\frac{n}{2}}$ into the product produces

$$Z_n(x, z) = \prod_{r=1}^{\frac{n}{2}} \left[2x + z^2 + 2x \cos \left(\frac{2\pi r}{n+1} \right) \right]. \tag{8.26}$$

Again the limiting behavior for large lattices is

$$Z_n(x, z) \sim [Z(x, z)]^n, \tag{8.27}$$

which means that

$$\ln[Z(x, z)] = \lim_{n \to \infty} \frac{1}{n} \ln[Z_n(x, z)]. \tag{8.28}$$

This is, again, related to the free energy per lattice site, which can be used to find the entropy per lattice site, $G_{site} = -k_B T ln[Z_G]$. Plugging in Z_n gives

$$G_{site} = -k_B T \lim_{n \to \infty} \frac{1}{n} \ln \left[\prod_{r=1}^{\frac{n}{2}} 2x + z^2 + 2x \cos \left(\frac{2\pi r}{n+1} \right) \right]. \tag{8.29}$$

As before, the limit allows the change $n + 1 \sim n$ and the natural log allows a change from a product to a sum. The formula is now

$$G_{site} = -k_B T \lim_{n \to \infty} \frac{1}{n} \sum_{r=1}^{\frac{n}{2}} \ln \left[2x + z^2 + 2x \cos \left(\frac{2\pi r}{n} \right) \right]. \tag{8.30}$$

Using the same conversion as in Eq. (8.17) to convert this sum to an integral with

$\omega = \frac{2\pi r}{n}$ yields the integral

$$G_{site} = -k_B T \frac{1}{2\pi} \int_0^\pi \ln\left[2x + z^2 + 2x\cos(\omega)\right] d\omega. \tag{8.31}$$

This integral is exactly solvable using G.R. (4.224.9)[15] which says

$$\int_0^\pi \ln[a + b\cos(x)]dx = \pi \ln\left[\frac{a + \sqrt{a^2 - b^2}}{2}\right], a \geq |b| > b. \tag{8.32}$$

Evaluating the integral gives

$$G_{site} = -k_B T \frac{1}{2\pi} \pi \ln\left[\frac{(z^2 + 2x) + \sqrt{(z^2 + 2x)^2 - (2x)^2}}{2}\right]. \tag{8.33}$$

Expanding inside the square root gives

$$G_{site} = -k_B T \frac{1}{2} \ln\left[\frac{z^2 + 2x + \sqrt{z^4 + 4xz^2 + 4x^2 - 4x^2}}{2}\right]. \tag{8.34}$$

Multiplying the argument of the natural log by $\frac{2}{2}$ and factoring a z^2 from the square root allows this to be written as

$$G_{site} = -k_B T \frac{1}{2} \ln\left[\frac{2z^2 + 2z\sqrt{z^2 + 4x} + 4x}{4}\right], \tag{8.35}$$

which completes the square in the numerator. This can be factored into

$$G_{site} = -k_B T \frac{1}{2} \ln\left[\left(\frac{z + \sqrt{z^2 + 4x}}{2}\right)^2\right]. \tag{8.36}$$

By the rules of the natural log the exponent is brought down in front to cancel with the $\frac{1}{2}$. The final form is

$$G_{site} = -k_B T \ln\left[\frac{1}{2}\left(z + \sqrt{z^2 + 4x}\right)\right]. \tag{8.37}$$

This form will be used as a representation of the edge of a strand of DNA in solution. A y-dimer is analytically the same as a monomer in the formulation of the problem

in this one-dimensional case, since it occupies a single site on the lattice.

CHAPTER 9

THE CALCULATION OF THE ENTROPIES

This chapter is concerned with the calculation of the per-site entropy of the system with only monomers and of the one-dimensional system with monomers and dimers. When doing calculations of real applications, one must always allow for unfilled lattice sites. Without vacancies, or holes, in the lattice, equations will not properly show how the combination of binding energy and chemical potential affects how the lattice is filled.

9.1 The Entropy of Monomers and Holes

From Sec. 2.3, the solution for the partition function of the monomer only lattice was shown to be

$$Z(r, b, g)_N = (r + b + g)^N. \tag{9.1}$$

In the limit as N becomes large the limiting behavior of that lattice is

$$Z(r, b, g)_N \sim Z(r, b, g)^N, \tag{9.2}$$

and as before in the limit as $N \to \infty$ this becomes

$$\ln\left[Z(r, b, g)_N\right] = \lim_{N \to \infty} \frac{1}{N} \ln\left[Z(r, b, g)^N\right] = \lim_{N \to \infty} \frac{1}{N} \ln\left[(r + b + g)^N\right], \tag{9.3}$$

which, through the properties of the natural log, can be written as

$$\ln\left[Z(r, b, g)_N\right] = \ln\left[r + b + g\right]. \tag{9.4}$$

This is related to the free energy per lattice site by the term $-k_B T$ and can be used to find the entropy per lattice site. The entropy is given by

$$S_{site} = -\frac{\partial}{\partial T}\Omega_G, \tag{9.5}$$

where

$$\Omega_G = -k_B T \ln[Z_G] = -\frac{1}{\beta}\ln[Z_G]. \tag{9.6}$$

The entropy per lattice site is

$$S_{site} = -\frac{\partial}{\partial T} - \left\{\frac{1}{\beta}\ln[r+b+g]\right\}, \tag{9.7}$$

where this can be changed to the partial derivate with respect to β, where $\beta = \frac{1}{k_B T}$ giving

$$S_{site} = \frac{\partial}{\partial \beta}\left\{\frac{1}{\beta}\ln[r+b+g]\right\}\frac{\partial \beta}{\partial T}. \tag{9.8}$$

Applying the same assumptions to r, b, and g that were applied to x, y, and z in the formulation of the Hamiltonian gives the activities of $r = e^{-\beta(E_r - \mu_r)}$, $b = e^{-\beta(E_b - \mu_b)}$, and $g = e^{-\beta(E_g - \mu_g)}$. The per site entropy is now

$$S_{site} = -k_B \beta^2 \frac{\partial}{\partial \beta}\left\{\frac{1}{\beta}\ln\left[e^{-\beta(E_r - \mu_r)} + e^{-\beta(E_b - \mu_b)} + e^{-\beta(E_g - \mu_g)}\right]\right\}. \tag{9.9}$$

This can be expanded by the product rule of differentiation as

$$
\begin{aligned}
S_{site} = -k_B \beta^2 \Big\{ &\ln\left[e^{-\beta(E_r - \mu_r)} + e^{-\beta(E_b - \mu_b)} + e^{-\beta(E_g - \mu_g)}\right]\frac{\partial}{\partial \beta}\frac{1}{\beta} \\
&+\frac{1}{\beta}\frac{\partial}{\partial \beta}\ln\left[e^{-\beta(E_r - \mu_r)} + e^{-\beta(E_b - \mu_b)} + e^{-\beta(E_g - \mu_g)}\right]\Big\}.
\end{aligned}
\tag{9.10}
$$

For convenience the exponents can be compressed by defining $\gamma_a = E_a - \mu_a$. The partials are taken while remembering that $Z(r,b,g) = (r+b+g)$, resulting in

$$S_{site} = -k_B \beta^2 \left\{-\frac{1}{\beta^2}\ln[Z(r,b,g)] + \frac{1}{\beta}\frac{1}{Z(r,b,g)}\left[-r\gamma_r - b\gamma_b - g\gamma_g\right]\right\}. \tag{9.11}$$

After distrubuting a $-\beta^2$ this is finally written as

$$S_{site} = k_B \left\{ \ln\left[Z(r,b,g)\right] + \frac{r\beta\gamma_r}{Z(r,b,g)} + \frac{b\beta\gamma_b}{Z(r,b,g)} + \frac{g\beta\gamma_g}{Z(r,b,g)} \right\}. \tag{9.12}$$

This is the entropy per lattice site for a system filled with three kinds of monomers, red, blue, and green. It is simple enough to see the pattern to create the entropy if more species of monomer are required. If vacancies are required then the species g can be considered to be. In this case $e^{-\beta(E_g - \mu_g)} = 1$ and $\gamma_g = 0$ as the binding energy and chemical potential of a vacancy, or hole, are both zero. The entropy of a system with two species of monomer and holes occupying the lattice can be expressed as

$$S_{site} = k_B \left\{ \ln\left[Z(r,b,1)\right] + \frac{r\beta\gamma_r}{Z(r,b,1)} + \frac{b\beta\gamma_b}{Z(r,b,1)} \right\}, \tag{9.13}$$

where the term $\frac{g\beta\gamma_g}{Z(r,b,1)}$ is dropped because γ_g is zero. This might be more appropriately written as

$$S_{site} = k_B \left\{ \ln\left[Z(r,b)\right] + \frac{r\beta\gamma_r}{Z(r,b)} + \frac{b\beta\gamma_b}{Z(r,b)} \right\}, \tag{9.14}$$

where

$$Z(r,b) = (r + b + 1), \tag{9.15}$$

and where it is assumed that there must be vacancies for this equation to represent a real world physical system. This form of the entropy is analogous to the lattice-gas model of dimers filling the edge of the DNA lattice. Each particle is considered as a monomer that can occupy a single site.

9.2   The Entropy of Monomers, Dimers, and Holes

The addition of vacancies to the monomer-dimer model in Eq. (8.37) is done by considering the monomer terms inside $G_{site} = -\frac{1}{\beta} \ln\left[\frac{1}{2}\left(z^2 + \sqrt{z^2 + 4x}\right)\right]$ to instead be a combination of species, each of which could occupy a single point. Specifically,

that there are z monomers and h holes. This is now written as

$$\ln[Z(x, z, h)] = \ln\left[\frac{1}{2}\left((z + h) + \sqrt{(z + h)^2 + 4x}\right)\right]. \tag{9.16}$$

As in the monomer-hole problem in the previous section, the chemical potential and the energy of a hole are both taken to be zero, and so $h = 1$. It can be said that

$$\ln[Z(x, z, 1)] = \ln\left[\frac{1}{2}\left(z + 1 + \sqrt{(z + 1)^2 + 4x}\right)\right]. \tag{9.17}$$

Just as before, the negative partial derivative with respect to β is taken with the conversion factor of $\frac{\partial \beta}{\partial T}$ to give the per site entropy as

$$S_{site} = \frac{\partial}{\partial \beta}\left\{\frac{1}{\beta}\ln\left[\frac{1}{2}\left(e^{-\beta\gamma_z} + 1 + \sqrt{(e^{-\beta\gamma_z} + 1)^2 + 4e^{-\beta\gamma_x}}\right)\right]\right\}\frac{\partial \beta}{\partial T}. \tag{9.18}$$

By the product rule this is

$$S_{site} = \left\{\frac{1}{\beta}\frac{\partial}{\partial \beta}\ln\left[\frac{1}{2}\left(e^{-\beta\gamma_z} + 1 + \sqrt{(e^{-\beta\gamma_z} + 1)^2 + 4e^{-\beta\gamma_x}}\right)\right]\right.$$
$$\left. + \ln\left[\frac{1}{2}\left(e^{-\beta\gamma_z} + 1 + \sqrt{(e^{-\beta\gamma_z} + 1)^2 + 4e^{-\beta\gamma_x}}\right)\right]\frac{\partial}{\partial \beta}\frac{1}{\beta}\right\}\frac{\partial \beta}{\partial T}. \tag{9.19}$$

After a few iterations of the chain rule this gives

$$S_{site} = -k_B\beta^2\left\{\frac{1}{\beta}\frac{1}{Z(x, z)}\left(\frac{1}{2}\right)\left(-z\gamma_z + \frac{2(z + 1)(-z\gamma_z) - 4x\gamma_x}{2\sqrt{(z + 1)^2 + 4x}}\right) + \ln[Z(x, z)]\left(\frac{1}{\beta}\right)\right\}. \tag{9.20}$$

The $-\beta^2$ can be distrbuted into the braces and the factors outside of the parentheses can be distributed inside to give the form

$$S_{site} = k_B\left\{\ln[Z(x, z)] + \frac{z\beta\gamma_z}{2Z(x, z)} + \frac{z^2\beta\gamma_z}{2Z(x, z)\sqrt{(z + 1)^2 + 4x}}\right.$$
$$\left. + \frac{z\beta\gamma_z}{2Z(x, z)\sqrt{(z + 1)^2 + 4x}} + \frac{x\beta\gamma_x}{Z(x, z)\sqrt{(z + 1)^2 + 4x}}\right\}, \tag{9.21}$$

where

$$Z(x, z, 1) = Z(x, z) = \frac{1}{2}\left(z + 1 + \sqrt{(z+1)^2 + 4x}\right). \tag{9.22}$$

CHAPTER 10

PLOTS OF THE ENTROPY, AVERAGE OCCUPATION, AND TOTAL CHARGE

At this point it is possible to draw comparisons between the lattice-gas model and the Fisher dimer model. There should be a great difference between the two entropies since the Fisher model takes into account which sites are excluded from further occuptaion by the geometry and therefore changes the number of ways in which the particles may be arranged on the lattice. Plots for the for the average occupation of species will also be given in this section and from these occupation plots the total charge of the lattice can be calculated.

10.1 Lattice Gas

As stated in the introduction to this paper the lattice-gas model considers all particles to be monomers occupying single sites. To make them appear as dimers the model assigns the monomers different energies. Fig. 1 shows an example of how this model considers sites to be occupied. The appropriate entropy to use is that with two types of monomers and holes which was given in Eq. (9.14) as

$$S_{site} = k_B \left\{ \ln \left[Z(r,b) \right] + \frac{r\beta\gamma_r}{Z(r,b)} + \frac{b\beta\gamma_b}{Z(r,b)} \right\}, \tag{10.1}$$

where

$$Z(r,b) = (r + b + 1). \tag{10.2}$$

The activities of the monomers are $r = e^{-\beta(E_r - \mu_r)}$, and $b = e^{-\beta(E_b - \mu_b)}$ and $\gamma_r = E_r - \mu_r$, and $\gamma_b = E_b - \mu_b$. If the lattice is considered to be filled by two of different

types they could be described as being attached perpendicular to the direction of the lattice or parallel to it. To attempt to model dimers with the lattice-gas model some of the monomers are considered to be parallel to the lattice and some perpendicular. The monomers which are considered to be perpendicular are given half the binding energy when compared to those lying parallel. This can be represented mathematically as $z_\perp = e^{-\beta(0.5\mathcal{E}-\mu)}$ and $z_\parallel = e^{-\beta(\mathcal{E}-\mu)}$ where $z_\perp$ and $z_\parallel$ are monomers in the lattice-gas model that are taken to be perpendicular or parallel to the edge of the lattice respectively. The entropy is written as

$$S_{site} = k_B \left\{ \ln \left[Z(z_\perp, z_\parallel) \right] + \frac{z_\perp(0.5\beta\mathcal{E} - \beta\mu)}{Z(z_\perp, z_\parallel)} + \frac{z_\parallel(\beta\mathcal{E} - \beta\mu)}{Z(z_\perp, z_\parallel)} \right\}, \qquad (10.3)$$

where

$$Z(z_\perp, z_\parallel) = z_\perp + z_\parallel + 1. \qquad (10.4)$$

A factor of $\beta\mu = 0.79$ for a dimer is used. This comes from a pending paper by Bishop and McMullen but is derived using the work of Nguyen and Shklovskii on the energy of polyelectrolytes in solution[16]. Nguyen and Shklovskii show that the chemical potential of a free polyelectrolyte in solution can be expressed as the self-energy of a rigid rod. Following through their calculations with dimers will result in a $\beta\mu$ of 0.79 when a lattice constant of ~ 0.68nm is used. The result of the entropy calculation as a function of $\beta\mathcal{E}$ is shown in Fig. 22.

Fig. 22 reveals that as $\beta\mathcal{E}$ becomes very positive or very negative there are fewer ways to occupy the lattice. This makes sense because if the binding energy is very positive the lattices resists binding and will be filled mostly by holes. At large negative values for $\beta\mathcal{E}$ the number of ways to effectively rearrange the particles on the lattice also decreases because the lattice will become occupied by only one species. Since this is a model with dimers represented by monomers and all other effects are ignored
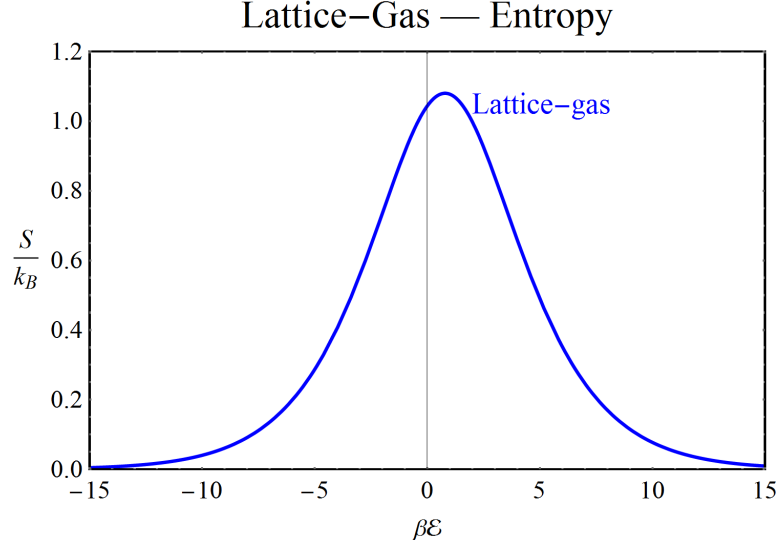
Lattice–Gas — Entropy

Fig. 22.: The entropy of the lattice-gas representation of an occupied lattice as a function of binding energy.

the lattice should tend to be occupied by the species with the highest binding energy as $\beta\mathcal{E}$ becomes more and more negative. This is of course the species $z_{parallel}$.

The average occupation of a species can be calculated. It is defined here as

$$\langle N_a \rangle = \frac{1}{Z_{mn}} \sum_N g_{mn} N_a e^{-\beta \sum_a (E_a - \mu_a) N_a}. \tag{10.5}$$

This is exactly the partial of the state function in this paper with respect to its chemical potential multiplied by a factor of $\frac{1}{\beta Z_{mn}}$. For example if the average number, $\langle N \rangle$, of the species x is required then

$$\frac{\partial}{\partial \mu_x} Z_{mn} = \frac{\partial}{\partial \mu_x} \sum_{N_x, N_y, N_z} g_{mn}(N_x, N_y, N_z) x^{N_x} y^{N_y} z^{N_z}. \tag{10.6}$$

Since everything other than the exponent of $x^{N_X} = e^{-\beta(E_x - \mu_x)N_x}$ is a constant the partial derivative gives

$$\frac{\partial}{\partial \mu_x} Z_{mn} = \sum_{N_x, N_y, N_z} g_{mn}(N_x, N_y, N_z) x^{N_x} y^{N_y} z^{N_z} (\beta N_x). \tag{10.7}$$

103

Thus it can be said that in general

$$\langle N_a \rangle = \frac{1}{\beta} \frac{1}{Z_{mn}} \frac{\partial}{\partial \mu_a} Z_{mn}. \tag{10.8}$$

The number of particles considered to occupy the lattice in the parallel and perpendicular directions along with the holes in the lattice can be calculated using the correct parital derivative of the function

$$Z(z_\perp, z_\parallel) = z_\perp + z_\parallel + 1. \tag{10.9}$$

However in order to compute the occupation of holes it is necessary to have this in the form

$$Z(z_\perp, z_\parallel, z_h) = z_\perp + z_\parallel + z_h, \tag{10.10}$$

where $z_h = e^{-\beta(\mathcal{E}_h - \mu_h)}$, even though for a hole, the energy and chemical potential are necessarily equal to each other. Then the partial with respect to μ_{z_h} is

$$\frac{\partial}{\partial \mu_{z_h}} Z(z_\perp, z_\parallel, z_h) = \frac{\partial}{\partial \mu_{z_h}} \left(z_\perp + z_\parallel + z_h \right), \tag{10.11}$$

which yields, when $E_h - \mu_h = 0$,

$$\frac{\partial}{\partial \mu_{z_h}} Z(z_\perp, z_\parallel, z_h) = \beta z_h = \beta. \tag{10.12}$$

The occupation for holes is thus

$$\langle N_{z_h} \rangle = \frac{1}{\beta} \frac{1}{Z(z_\perp, z_\parallel, z_h)} \beta = \frac{1}{Z(z_\perp, z_\parallel, z_h)}. \tag{10.13}$$

Similar steps are taken for $z_\perp$ and $z_\parallel$ giving

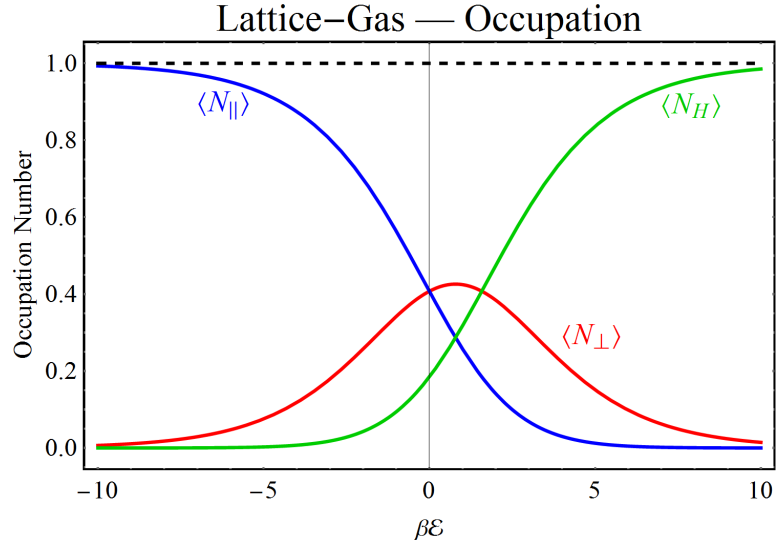$$\langle N_{z_\perp} \rangle = \frac{z_\perp}{Z(z_\perp, z_\parallel, z_h)}, \tag{10.14}$$

Fig. 23.: The occupation of species of the lattice gas model of dimers.

and

$$\langle N_{z_\parallel} \rangle = \frac{z_\parallel}{Z(z_\perp, z_\parallel, z_h)}.$$

(10.15)

These are plotted in Fig. 23 with a dashed line representing the sum of the three occupations which should, and does, total to one.

Fig. 23 demonstrates that as the binding energy becomes more negative the lattice is overwhelmingly occupied by those particles that are considered to be parallel because they have twice the binding energy of the perpendicular particles. As the energy becomes increasingly positive holes begin to dominate the lattice because the lattice repels all likely candidates for binding.

The total charge on the lattice can be calculated from the the occupation. These are particles which are binding to the outside edge of a strand of DNA which, on its own, is entirely negatively charged. Thus a hole would represent a negative charge. A parallel dimer is considered to be two positive charges lying flat on the lattice and cancelling two negative charges and would therefore appear to be neutral. A perpendicular dimer creates a net positve charge by virtue of having only one of its

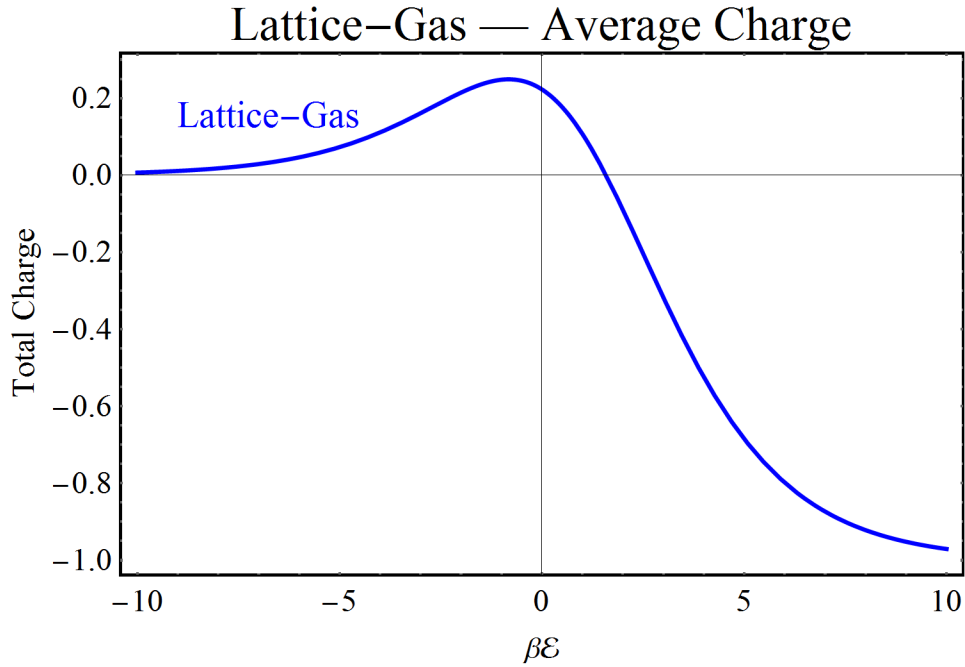## Lattice–Gas — Average Charge

Fig. 24.: The charge on the lattice as a function of the binding energy

positive charges attached to the lattice and the other hanging off the lattice into the solution. The total charge in units of the magnitude of the charge of an electron is

$$\text{Total Charge} = z_\perp + z_h. \tag{10.16}$$

Plotting gives Fig. 24, which indicates that in the lattice-gas model, as the binding energy becomes increasingly negative, the charge on the lattice trends toward neutrality.

10.2   The Fisher Model of Dimers

The results in Sec. 9.2 allow an investigation into the entropy of a lattice occupied by two species of dimers and by holes. This can be done because a monomer would occupy the same number of lattice sites as a perpendicular dimer. A picture of this situation is given in Fig. 2.

The entropy in this case is

$$S_{site} = (k_B) \left\{ \ln \left[Z(x_\parallel, z_\perp) \right] + \frac{z_\perp \beta \gamma_{z_\perp}}{2Z(x_\parallel, z_\perp)} + \frac{z_\perp^2 \beta \gamma_{z_\perp}}{2Z(x_\parallel, z_\perp)\sqrt{(z_\perp + 1)^2 + 4x_\parallel}} \right. $$
$$\left. + \frac{z_\perp \beta \gamma_{z_\perp}}{2Z(x_\parallel, z_\perp)\sqrt{(z_\perp + 1)^2 + 4x_\parallel}} + \frac{x_\parallel \beta \gamma_{x_\parallel}}{Z(x_\parallel, z_\perp)\sqrt{(z_\perp + 1)^2 + 4x_\parallel}} \right\}, $$

$$\text{(10.17)}$$

where

$$Z(x_\parallel, z_\perp) = \frac{1}{2}\left(z_\perp + 1 + \sqrt{(z_\perp + 1)^2 + 4x_\parallel}\right). \qquad \text{(10.18)}$$

To mantain consistency with the previous section, the perpendicular and parallel subscripts have been added. The activties are $z_\perp = e^{-\beta(0.5\mathcal{E}-\mu)}$ and $x_\parallel = e^{-\beta(\mathcal{E}-\mu)}$. With these substituted, the entropy is plotted in Fig. 25 with the resultfor the lattice-gas model plotted as a dashed line for comparison. Note that if $\mathcal{E}$ is the binding energy of a parallel dimer, the energy of a perpendicular dimer would be half as big since only one of the two atoms in the dimer isbound. This is the same as was used for the lattice-gas model.

While the entropy for the lattice gas model trended twoard zero as the binding energy became more negative (because it was only occupied by $z_\parallel$) the Fisher model approaches a positive finite entropy. This indicates that even as the lattice becomes incredibly attractive, it is still occuped by a mixture of species of dimer.

The occupation numbers for each of the types of particles is calculated in the same manner as they were calculated in the previous section, with one notable exception. In the lattice-gas model the formula for the average number of a particle can be taken to be the occupation number of that particle on the lattice, because each particle is actually a monomer. However, in the Fisher model, the average number of parallel dimers will not be the occupation number. Any number of parallel dimers which are selected actually occupy twice that number of lattice sites. Thus the occupation
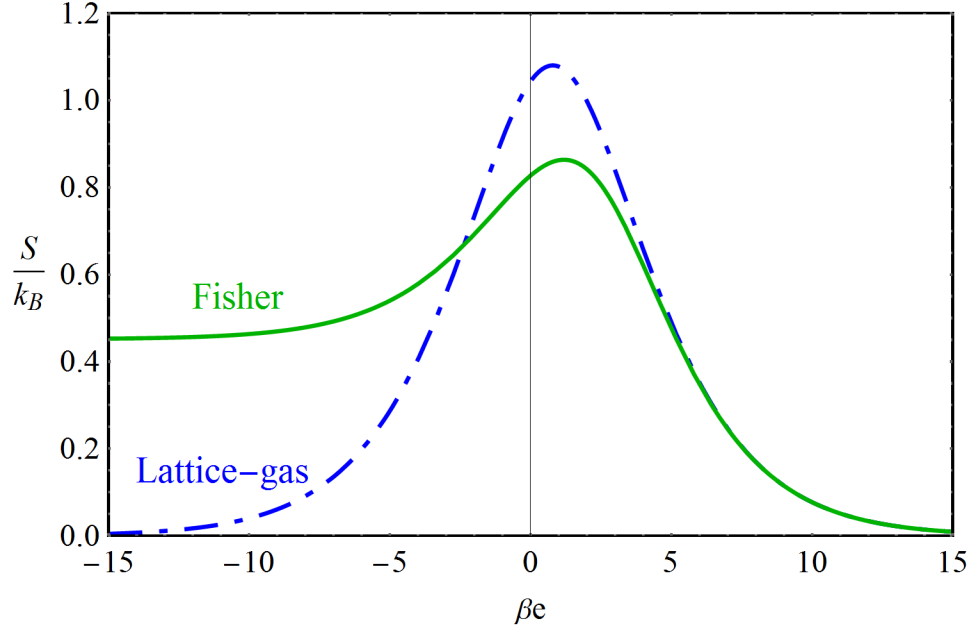
107

Fig. 25.: The plot of the two entropies given by the lattice-gas model and the Fisher model of dimers.

for parallel dimers will be twice the average number of parallel dimers. Since the derivations for each type are so similar only the average for $x_{parallel}$ is worked out below. To start

$$\langle N_{x_\|} \rangle = \frac{1}{\beta} \frac{1}{Z} \frac{\partial}{\partial \mu_{x_\|}} Z, \tag{10.19}$$

where

$$Z(x_\|, z_\perp) = \frac{1}{2} \left(z_\perp + 1 + \sqrt{(z_\perp + 1)^2 + 4x_\|} \right), \tag{10.20}$$

or, if working on $\langle N_{z_h} \rangle$, then

$$Z(x_\|, z_\perp, z_h) = \frac{1}{2} \left(z_\perp + z_h + \sqrt{(z_\perp + z_h)^2 + 4x_\|} \right). \tag{10.21}$$

Combining this information gives

$$\langle N_{x_\|} \rangle = \frac{1}{\beta} \frac{1}{Z(x_\|, z_\perp)} \frac{\partial}{\partial \mu_{x_\|}} \frac{1}{2} \left(z_\perp + 1 + \sqrt{(z_\perp + 1)^2 + 4x_\|} \right), \tag{10.22}$$
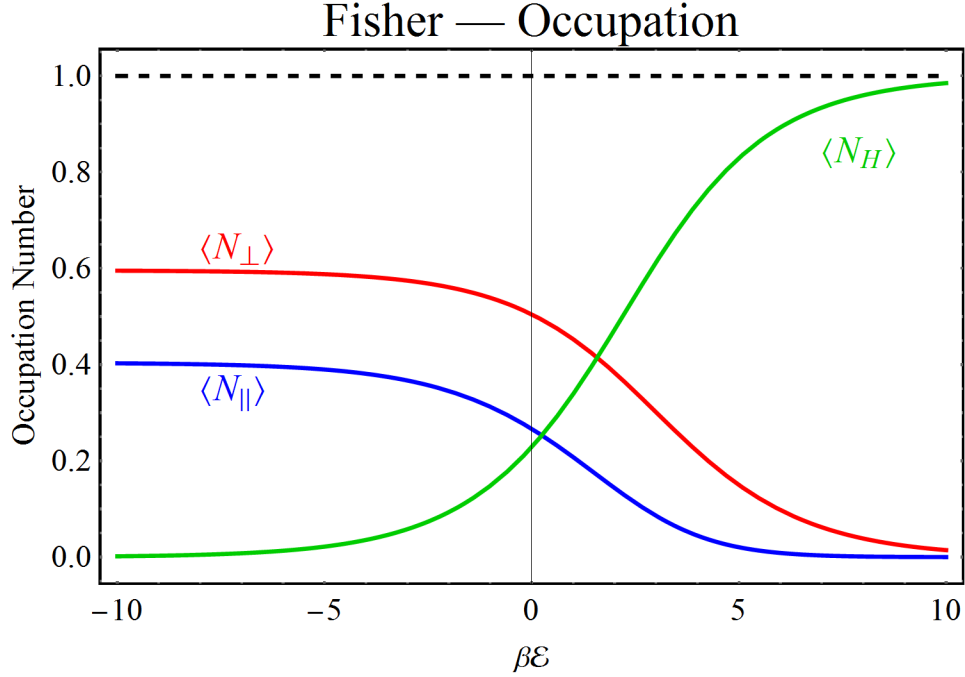
Fig. 26.: The average occupation of species in the Fisher model of dimers showing species mixing even at large negative binding energy.

which reduces to

$$\langle N_{x_\parallel} \rangle = \frac{x_\parallel \beta}{Z(x_\parallel, z_\perp, z_h)\sqrt{(z_\perp + 1)^2 + 4x_\parallel}}. \tag{10.23}$$

Then as it was stated above the parallel dimers occupy two lattice sites, and so the occupation number for this must doubled

$$2\langle N_{x_\parallel} \rangle = \frac{2x_\parallel \beta}{Z(x_\parallel, z_\perp, z_h)\sqrt{(z_\perp + 1)^2 + 4x_\parallel}}. \tag{10.24}$$

The occupation for the other types of particles is found by simply taking their averages without doubling in this same manner. The plot for the total occupation under the Fisher model along with a dashed line to check that the total occupation of the lattice is equal to one is shown in Fig. 26.

In the Fisher model, the total number of holes vanishes to zero as the binding energy becomes more and more negative. However, the lattice is still filled with a mix of parallel and perpendicular dimers. This mixing is what leads to the positive
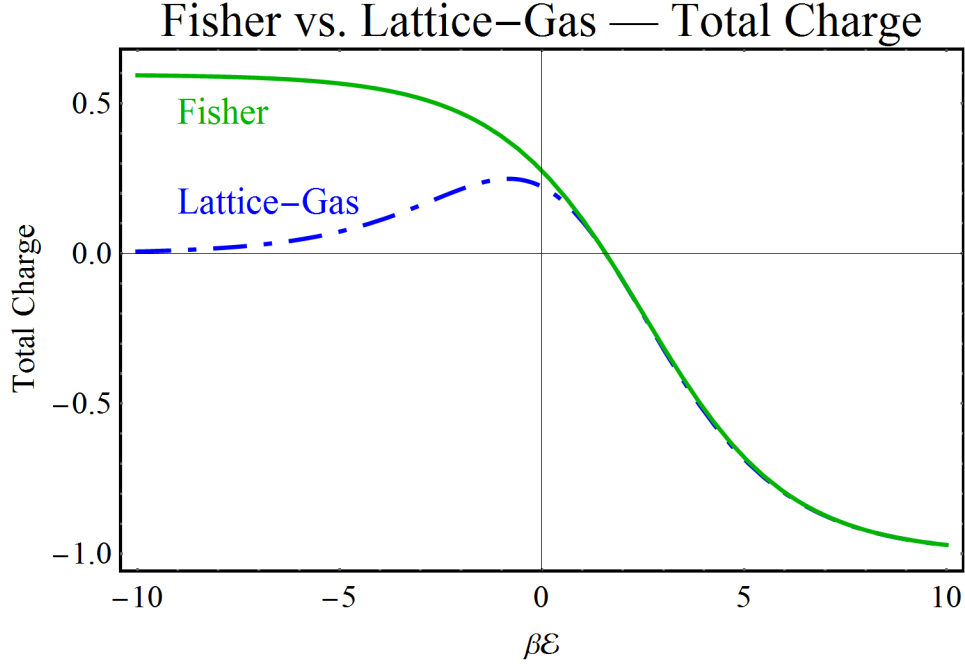
Fig. 27.: A comparison of the total charge vs.the binding energy on the lattice between the lattice-gas model and the Fisher model.

constant entropy per lattice site in Fig. 25.

The total charge of the lattice of the Fisher model can be calculated via

$$\text{Total Charge} = z_\perp + z_h, \tag{10.25}$$

which is plotted in Fig. 27.

The total charge is probably the most physically important difference between the two models. In the Fisher model, the lattice becomes positive as the binding energy trends negatively. The lattice will be occupied by mostly neutral sites caused by parallel dimers but any leftover single sites that cannot be occupied by parallel dimers will pick up a perpendicular dimer and become a positive site. This is in contrast to the lattice-gas model which, since it is actually filling the lattice with monomers, will become filled with the higher binding energy "parallel" monomers and eventually become neutral.

CHAPTER 11

CONCLUSIONS AND FUTURE DIRECTIONS

11.1 Conclusions and Results

The main body of this paper is meant as an introduction into applied statistical mechanics and as such much of it expounds the work of Fisher and of others who have come before. Despite being introductory, this paper comes to some conclusions that are physically relevant. In the lattice gas model, the charge was seen to become neutral as the lattice became more attractive. But the Fisher model has the lattice becoming positively charged. This is more in line with what is physically true for polymers in solution. The charge on the lattice should not trend toward neutral. In fact it could even have what Nguyen, Grosberg and Shklovskii termed a, "Giant Inversion of Charge," where the net charge after charge inversion is greater than the original charge on the lattice [17]. Although the result here does not demonstrate such a massive charge inversion, it is a step in the right direction from a simple lattice-gas treatment of the problem.

Additionally, the mathematics in this paper have some relevence to the Ising model, which is what gave birth to much of the work of Fisher[7, 9, 10, 12, 18]. The Ising model is one of a lattice occupied by magnetic particles that can be either spin up or spin down. The particles interact with each other and create chains and groups of oriented magnetic fields. Attempts at solutions to the Ising model drove the introduction of Pfaffians into the toolbox of the physicist.

11.2 Future Directions

Trying to directly count the lattice combinations, as was done in this paper through the Pfaffian, has been a long standing problem in combinatorics and graph theory in addition to being physically relevant. The method used in this paper of creating a generating function and converting it to a Pfaffian works for trimers with the exception of proving the sign is always positive for any lattice configuration of trimers. Proving the sign is always positive would be an interesting puzzle that, if solved, would allow these techniques to quickly find the entropy and total charge on a lattice in the presence of trimers. Additionally, steps have been made recently in working with two-dimensional lattices that have both monomers and dimers. Most notably Allegra and Fortin used Grassmanian algebra to formulate an exact fermionic solution to lattices filled with both monomers and dimers as a product of two Pfaffians[19].

Appendix A

ABBREVIATIONS

VCU   Virginia Commonwealth University

RVA   Richmond Virginia

Appendix B

MATHEMATICS

B.1   The Trace and Cyclic Interchange

In Eq. (2.24) the following relationships are given:

$$\text{Tr}\{A_i\} = 0,$$
$$\text{Tr}\{A_i A_j\} = 0,$$
$$\text{Tr}\{A_i A_j A_k\} = 0,$$
$$\text{Tr}\{A_i^2\} = Tr\{I\} = t,$$

(B.1)

where t is the dimension of the resulting matrix I. These relationships are true because the trace is invariant under cyclic interchange. If there are an odd number of operators in the trace then a new pair of operators $A_t A_t = I$ is introduced. Take the first relationship as an example

$$\text{Tr}\{A_i\} = \text{Tr}\{A_i A_t A_t\} = \text{Tr}\{-A_t A_i A_t\} = -\text{Tr}\{A_t A_i A_t\}.$$

(B.2)

If the operators are now cycled then the result of the trace will not change

$$\text{Tr}\{A_i\} = -\text{Tr}\{A_t A_i A_t\} \underset{cycle}{=} -\text{Tr}\{A_i A_t A_t\} = -\text{Tr}\{A_i\} = 0,$$

(B.3)

where the only way for the trace to equal its own negative is for it to be zero. A similar process is done if there are an even number of operators in the trace but instead of introducing an entirely new operator an operator that already exists in the

trace is chosen. In the second relationship this is

$$\mathrm{Tr}\{A_iA_j\} = \mathrm{Tr}\{A_iA_jA_jA_j\}- = \mathrm{Tr}\{A_jA_iA_jA_j\} \underset{cycle}{=} -\mathrm{Tr}\{A_iA_jA_jA_j\} = -\mathrm{Tr}\{A_iA_j\} = 0.$$

$$(\mathrm{B.4})$$

Bibliography

[1] W. M. Gelbart, R. F. Bruinsma, P. A. Pincus, and V. A. Parsegian, "DNA-inspired electrostatics," Physics Today **53**, 38–44 (2000).

[2] F. J. Solis, "Phase diagram of dilute polyelectrolytes: Collapse and redissolution by association fo counterions and co-ions," Journal of Chemical Physics **117**, 9009–9015 (2002).

[3] F. J. Solis and M. Olvera de la Cruz, "Attractive interactions between rodlike polyelectrolytes: Polarization, crystallization, and packing," Phys. Rev. E **60**, 4496–499 (1999).

[4] F. J. Solis and M. Olvera de la Cruz, "Collapse of flexible polyelectrolytes in multivalent salt solutions," Journal of Chemical Physics **1124**, 2030–2035 (2000).

[5] I. Koltover, K. Wagner, and C. R. Safinya, "DNA condensation in two dimensions," Proceedings of the National Academy of Sciences **97**, 14046–14051 (2000).

[6] M. F. Bishop and T. McMullen, "Lattice-gas model of DNA charge inversion by a positively charged polyelectrolyte," Phys. Rev. E **74**, 021906–(8) (2006).

[7] M. E. Fisher, "Statistical mechanics of dimers on a plane lattice," Phys. Rev. **124**, 1664–1672 (1961).

[8] P. W. Kasteleyn, "The statistics of dimers on a lattice," Physica **27**, 1209–1225 (1961).

[9] H. N. V. Temperley and M. E. Fisher, "Dimer problem in statistical mechanics - an exact result," Philosophical Magazine **6**, 1061–1063 (1961).

[10] C. A. Hurst and H. S. Green, "New Solution of the Ising Problem for a Rectangular Lattice," J. Chem. Phys. **33**, 1059–1062 (1960).

[11] T. Muir, *A Treatise on the Theory of Determinants* (MacMillan and Co., London, 2003).

[12] M. Kac and J. C. Ward, "A Combinatorial Solution of the Two-Dimensional Ising Model," Phys. Rev. **88**, 1332–1337 (1952).

[13] E. R. Caianiello and S. Fubini, "On the Algorithm of Dirac Spurs," Il Nuovo Cimento **9**, 1218–1226 (1952).

[14] H. B. Dwight, *Tables of Integrals and Other Mathematical Data* (The Macmillan Company, New York, 1957).

[15] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products* (Elsevier, Boston, 2007).

[16] T. T. Nguyen and B. I. Shklovskii, "Model of inversion of DNA charge by a positive polymer: Fractionalization of the polymer charge," Phys. Rev. Lett. **89**, 018101 (2002).

[17] T. Nguyen, A. Y. Grosberg, and B. Shklovkii, "Macroions in salty water with multivalent ions: Giant inversion of charge," Phys. Rev. Letters **85**, 1568–1571 (2000).

[18] L. Onsager, "Crystal statistics. i. a two-dimensional model with an order-disorder transformation," Phys. Rev. **65** (1944).

[19] N. Allegra and J.-Y. Fortin, "Grassmanian representation of the two-dimensional monomer-dimer model," Phys. Rev. E **89** (2014).