

A quick note on Kasteleyn and Probabilities

Peru d'Ornellas

June 6, 2025

These are just a few notes so I have it written down somewhere exactly why and how the inverse Kasteleyn matrix computes probabilities. It's a remarkably simple and elegant piece of mathematics.

1 The Inverse Kasteleyn Matrix

Let's start by restating Kasteleyn's theorem, which we shall not prove. For a graph Λ , we may define a Kasteleyn-oriented skew-adjacency matrix, K . Then the number of valid dimerisations is given by

$$N = |\text{pf } K|. \quad (1.1)$$

Now let's look at the inverse Kasteleyn matrix. In order to derive this, we must first state two properties of Pfaffians, which can both be found on the Wiki page for the Pfaffian. Firstly, we state the recursive definition of a Pfaffian for a general $2n \times 2n$ matrix, expanded along a row indexed with i ,

$$\text{pf } K = \sum_{j \neq i}^{2n} (-1)^{i+j+1+\theta(i-j)} K_{ij} \text{pf } K_{\tilde{i}\tilde{j}}, \quad (1.2)$$

where the matrix $K_{\tilde{i}\tilde{j}}$ is the matrix K with rows and columns i and j removed. Taking a derivative with respect to K_{ij} , we arrive at the following identity,

$$\frac{d}{dK_{jk}} \text{pf } K = (-1)^{i+j+1+\theta(i-j)} \text{pf } K_{\tilde{i}\tilde{j}}. \quad (1.3)$$

Next, let us write down a derivative identity for the Pfaffian, for a matrix with some arbitrary parameter $K(x)$

$$\frac{1}{\text{pf } K} \frac{d}{dx} \text{pf } K = \frac{1}{2} \text{Tr} \left(K^{-1} \frac{dK}{dx} \right). \quad (1.4)$$

Now, let us suppose that our derivative is just in the ij element of K . In this case, the K derivative on the right-hand side becomes

$$\frac{dK}{dK_{ij}} = |i\rangle\langle j| - |j\rangle\langle i| \quad (1.5)$$

The second term comes from the condition that K must remain antisymmetric. Thus, the expression 1.4 becomes

$$\frac{1}{\text{pf } K} \frac{d}{dK_{ij}} \text{pf } K = K_{ji}^{-1}. \quad (1.6)$$

We now insert eq. (1.4) into eq. (1.3), arriving at the following

$$K_{ji}^{-1} = \frac{(-1)^{i+j+1+\theta(i-j)} \text{pf } K_{\tilde{ij}}}{\text{pf } K}. \quad (1.7)$$

Finally, taking the modulus of both sides, we can get rid of that factor of -1 and arrive at the final well-known identity

$$|K_{ji}^{-1}| = \left| \frac{\text{pf } K_{\tilde{ij}}}{\text{pf } K} \right|. \quad (1.8)$$

Now we consult Kasteleyn's theorem and realise that this is a rather remarkable result! If ij represents an edge on our lattice, then we can be sure that $|\text{pf } K_{\tilde{ij}}|$ counts the total number of dimerisations on the lattice with edge ij removed, whereas the denominator counts the number of dimerisations on the original lattice. Thus, we have calculated the probability that edge ij is a dimer!

$$p(e_{ij}) = |K_{ji}^{-1}|. \quad (1.9)$$

1.1 Working on the Torus

So far, everything we have done has been in open boundary conditions. Things become a little more complex when one works on the torus. This is because Kasteleyn's theorem looks rather different on the torus due to the presence of topologically non-trivial loops, which wind the torus in the x and y direction. Thus, we must define four Kasteleyn matrices, indexed by K^{ab} , with $a, b \in \{0, 1\}$. Each of these indexes denotes whether the boundary conditions are periodic or anti-periodic in that respective direction. Equivalently, they denote the flux threaded through these two non-contractible loops. Again, we shall only state Kasteleyn's theorem for this case, which says that

$$N = \frac{1}{2} \left| \text{pf } K^{00} - \text{pf } K^{01} - \text{pf } K^{10} - \text{pf } K^{11} \right|. \quad (1.10)$$

This time, let's start by writing down the formula for the probability of a single edge hosting a dimer,

$$p(e_{ij}) = \frac{N_{\tilde{ij}}}{N} \quad (1.11)$$

$$= \left| \frac{\omega_{ab} \text{pf } K_{\tilde{ij}}^{ab}}{\omega_{cd} \text{pf } K^{cd}} \right|, \quad (1.12)$$

where we use Einstein summation convention and have defined the sign operator as a shorthand for the sign structure of the torus Kasteleyn theorem,

$$\omega_{ab} = \begin{cases} +1 & \text{if } a, b = 0, \\ -1 & \text{otherwise.} \end{cases} \quad (1.13)$$

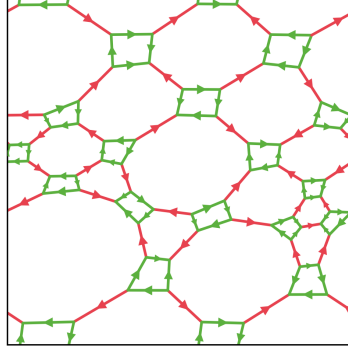
Now, we insert eq. (1.7), leading to the following expression for the overall probability,

$$p(e_{ij}) = \left| \frac{\omega_{ab} K_{ji}^{ab-1} \text{pf } K^{ab}}{\omega_{cd} \text{pf } K^{cd}} \right|, \quad (1.14)$$

This is a bit more complex than the expression for the regular probability but still totally manageable.

2 Some Neat Tricks on a Semi-regular Lattice

Now suppose we are working on an amorphous graph which still has the following regular structure, something like this:



where we represent the Kasteleyn orientation with the arrows, and all loops have an odd number of clockwise arrows. Thus, we can decompose the Kasteleyn ‘Hamiltonian’ into two parts: a term that just lives on each square, K_0 , and a term that connects two squares, V , highlighted in green and red respectively,

$$K = K_0 + vV, \quad (2.1)$$

where v parametrises the relative strengths of each bond type. Clearly we are aiming for $v = 1$. Now let us define the resolvent,

$$R(\lambda) = (\lambda - K)^{-1}, \quad (2.2)$$

which can be decomposed in terms of the ‘unperturbed’ resolvent, $R_0(\lambda) = (\lambda - K_0)^{-1}$, using the following series,

$$R = R_0 + R_0 V R_0 + R_0 V R_0 V R_0 + \dots \quad (2.3)$$

Obviously, the resolvent is just a fancy word for the matrix inverse which, as we discussed in the previous section, encodes the local dimer probabilities. So basically what we have here is a nice series expansion for the local dimer probabilities, and more!

Now, let’s calculate this expansion. The first thing we want to do is find R_0 , which is pretty trivial because H_0 is just a collection of unconnected squares, where each square has three clockwise arrows and one counterclockwise oneⁱ. Thus, the Hamiltonian H_0 is a bunch of copies of the following matrix,

$$H_0 \sim h = \begin{pmatrix} & 1 & & 1 \\ -1 & & 1 & \\ & -1 & & 1 \\ -1 & & -1 & \end{pmatrix}. \quad (2.4)$$

This is a rather remarkable operator because it satisfies the property

$$h^2 = -2\mathbb{1}, \quad (2.5)$$

ⁱOr the other way round, but these two configurations are gauge equivalent, just swap all arrows, which is always a gauge transform when we’re bipartite.

i.e. it is almost idempotent, $h^{-1} = -h/2$. Note that only the square case has this property, the same cannot be said of the hexagonal lattice where we can define an h_6 matrixⁱⁱ. To find the whole resolvent, let's make an ansatz for the inverse of $\lambda - h$, as the following expression parameterised by a and b ,

$$(\lambda - h)^{-1} = a + bh, \quad (2.6)$$

which is satisfied by

$$a = \frac{\lambda}{\lambda^2 + 2}, \quad (2.7)$$

$$b = \frac{1}{\lambda^2 + 2}. \quad (2.8)$$

Thus, the resolvent is given by

$$R_0(\lambda) = \frac{\lambda + h}{\lambda^2 + 2}. \quad (2.9)$$

That was easy! And it's all rather encouraging, since the factor of $1/(\lambda^2 + 2)$ is going to be useful in making sure our infinite sum doesn't blow up!

Hilbert's Identity—One useful identity for calculating stuff like this is the following,

$$R(\lambda) - R(\xi) = (\lambda - \xi)R(\lambda)R(\xi), \quad (2.10)$$

which we can massage to give the following expression

$$R(\lambda) = R(0)[\mathbb{1} - \lambda R(0)]^{-1} \quad (2.11)$$

2.1 The Resolvent at $\lambda = 0$

Let's focus on the resolvent at $\lambda = 0$, which is the simplest case, since $R_0(0) = h/2$. Now that we have an exact expression for R_0 , we can start to evaluate the full-system resolvent, R , using eq. (2.3). Specifically, we are interested in calculating the matrix element of R on the bond (ij) , which can be written as

$$R_{ij} = \frac{1}{2}h_{ij} + \frac{1}{4}h_{i\mu}V_{\mu\nu}h_{\nu j} + \frac{1}{8}h_{i\mu}V_{\mu\nu}h_{\nu\alpha}V_{\alpha\beta}h_{\beta j} + \dots \quad (2.12)$$

This is useful because as we have seen, the matrix element of $R(\lambda = 0)$ is precisely the probability of a given site hosting a dimerⁱⁱⁱ. At each order in the series, the terms can be interpreted as a weighted sum over all paths from i to j .

Naively this looks like it hasn't really helped, path counting on a disordered lattice is generally pretty hard. However, not all paths are counted in this expression! Our paths have to always be made up of alternating R and V elements, which do different things! Each R step can be viewed as entering and leaving a single green square in the diagram, whereas every V element jumps from one square to the next using a red bond. Furthermore, the R steps only have non-zero matrix elements for two of the four paths across a square, either turning left or turning right, any path that crosses over a square, or hits a square and turns back on itself is forbidden. It turns out that these constraints MASSIVELY reduce the number of paths we need to count in our expansion.

ⁱⁱMore on that later, h_6 and h_8 still have some lovely properties.

ⁱⁱⁱObviously in periodic boundaries this is complicated slightly, but let's not worry about that for now.

2.1.1 Counting All Paths

2.1.2 The Sign of a Path

3 Lieb's Theorem

Remarkably, this calculation appears to provide a clue towards a proof of Lieb's theorem on our special square-full bipartite lattice—and even might lead us to a general proof! To see why, let us start by stating what we aim to prove:

Theorem 1 Consider a signed adjacency matrix M , defined on a bipartite graph Λ with $|\Lambda| = 2n$, as

$$M_{jk} = \begin{cases} \pm 1 & \text{if } \langle jk \rangle, \\ 0 & \text{otherwise,} \end{cases} \quad (3.1)$$

where $M_{jk} = -M_{kj}$ and we have free choice over the which direction is positive for each edge. Since M is skew-symmetric, the eigenvalues of M come in complex pairs $\pm i\varepsilon_j$. We define the ground state energy as

$$E_{\text{g.s.}} = \sum_j^n \varepsilon_j. \quad (3.2)$$

Lieb's theorem states that the ground state energy is extremised when the signs of M are chosen such that M is a Kasteleyn orientation.

To work towards a proof, let us write the ground state energy as a contour integral in the resolvent,

$$E_{\text{g.s.}}(M) = \frac{1}{2\pi} \int_{\mathcal{C}} d\lambda \text{Tr}[MR(\lambda)], \quad (3.3)$$

Where the path \mathcal{C} encircles only the upper or lower half plan in the Argand diagram. To see why this works, write the expression in the eigenbasis of M , $\{|\pm\varepsilon_j\rangle\}$, where the eigenvalues are $\pm i\varepsilon_j$

$$E_{\text{g.s.}}(M) = \frac{1}{2\pi} \int_{\mathcal{C}} \sum_j \left(\frac{i\varepsilon_j}{\lambda - i\varepsilon_j} + \frac{-i\varepsilon_j}{\lambda + i\varepsilon_j} \right). \quad (3.4)$$

By encircling only the negative imaginary poles in this expression, we use Cauchy's residue theorem to show that the calculation evaluates the sum of negative eigenvalues,

$$\frac{1}{2\pi} \int_{\mathcal{C}} d\lambda \text{Tr}[MR(\lambda)] = \sum_j \varepsilon_j. \quad (3.5)$$

Now, let us use the expansion we derived above for the resolvent, again expanding the matrix M into two parts, M_0 and V , which live on- and between-squares respectively.

$$E_{\text{g.s.}}(M) = \frac{1}{2\pi} \int_{\mathcal{C}} d\lambda \text{Tr}[(M_0 + V)(R_0 + R_0 V R_0 + \dots)] \quad (3.6)$$

Next, we make the first assumption (which we shall work to prove is the right assumption later). That is, assume that all squares already have flux -1 , which means that M_0 is already described by the h that we showed above. Then we have the following identity,

$$hR_0(\lambda) = \lambda \frac{h - 2/\lambda}{\lambda^2 + 2} \quad (3.7)$$