Kasteleyn's formula for domino tilings of rectangles

- presentation on Eric Stucky's expository paper -

Arkaraj Mukherjee

B.Math year I, Indian Statistical Institute (Bangalore)
bmat2511@isibang.ac.in

December 9, 2025

Contents

Overview

This is a presentation on Kasteleyn's formula [Kasteleyn 1961] for the number of domino tilings of a $m \times n$ rectangle heavily based on [Stucky 2015]. The key result we prove is,

Theorem 1.1 (Kasteleyn, 1961)

If $T(m, n)$ is the set of tilings of a $m \times n$ rectangle (wlog. m is even) and $H(\pi), V(\pi)$ the number of horizontal and vertical dominos in $\pi \in T(m, n)$ respectively, then for any complex x, y we have

$$\sum_{\pi \in T(m,n)} x^{H(\pi)} y^{V(\pi)} = 2^{\frac{mn}{2}} \prod_{k=1}^{m/2} \prod_{l=1}^{n} \left(x^2 \cos^2 \frac{\pi k}{m+1} + y^2 \cos^2 \frac{\pi l}{n+1} \right)^{\frac{1}{2}}$$

Setting $x = y = 1$ gives us the number of such tilings.

The proof is separated into two major parts. In the first we deal with the combinatorial aspect where we translate the problem of counting tilings into that of counting specific permutations allowing us to use matrix methods[1] especially Pfaffians, which seem to be just what we need here. For a specific $mn \times mn$ matrix D, the expression in the main theorem will be $\text{Pf}(D)$. We rely on the identity $\text{Pf}(K)^2 = \det(K)$ where K is skew symmetric of even dimension. First we will come up with a *guess* D', which only differs from D by some signs but the hardest problem of this part is constructing D itself, signs included. In the second part of the proof we will use kronecker products and other linear algebraic tools to analyse D and finally calculate its determinant which is equivalent to the main problem.

---

[1] We have already seen things like the State matrix in the same context, tilings of hexagons with lozenges using determinants etc.

Translation of the problem into counting permutations

As discussed above we will try to connect tilings with dominos to permutations on $[mn]$. We will identify the usual unit cells in the $m \times n$ rectangle with numbers from $[mn]$ using the lexicographic ordering, starting from the bottom left corner. For example, when $m = n = 2$ we identify,

$$(1, 1) \leftrightarrow 1, (1, 2) \leftrightarrow 2, (2, 1) \leftrightarrow 3, (2, 2) \leftrightarrow 4$$

Now we line all the dominos up and label both of their ends with adjacent numbers as follows: Label the ends of the first domino with 1 and 2, the second with 3 and 4, ... upto $mn/2$ dominos. Then, we can associate to each tiling a permutation σ on $[mn]$ which maps k (some end of a domino) to the lexicographic rank of its position on the $m \times n$ rectangle.

It is easy to notice that a tiling can have multiple permutations corresponding to it (this is fixed over all tilings, finding how many corresponding permutations there are is a good exercise) and also that not every permutation corresponds to a tiling (if this weren't the case then this problem would be trivial), for example take (24) on $[4]$. We seek some way to identify an unique permutation corresponding with some tilings, something *canonical*. One way to do this is to put these restrictions on the permutations: $\sigma(2k-1) < \sigma(2k)$ and $\sigma(1) < \sigma(3) < \sigma(5) < \ldots$ These hand-wavingly make sense because flipping a domino should be immaterial to the number of tilings so we fix the higher label(equivalently both) and also that switching *pairs* around just replaces two dominos which should also be immaterial - for example, if $m = n = 2$ we see that 1234 and 3412 correspond to the exact same tiling. Convince yourself that such permutations exist given any tiling, leaving uniqueness for exercise.

The Pfaffian

Definition 3.1 (Pfaffian)

For a $2n \times 2n$ skew symmetric matrix A, the pfaffian is defined as

$$\text{Pf}(A) := \frac{1}{2^n n!} \sum_{\sigma \in S_{2n}} \text{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(2i-1), \sigma(2i)}$$

However, there is a way to write the pfaffian with lesser terms. Let Π be the set of partitions of $[2n]$ into pairs. For such a partition $\pi = \{\{i_1, j_1\}, \ldots, \{i_n, j_n\}\}$ where $i_1 < i_2 < i_3 < \ldots$ and $i_k < j_k$ for all k, let σ_π be the permutation $i_1 j_1 i_2 j_2 \ldots$. With the notation above, we claim that

$$\text{Pf}(A) = \sum_{\pi \in \Pi} \text{sgn}(\sigma_\pi) \prod_{i=1}^{n} a_{\sigma_\pi(2i-1), \sigma_\pi(2i)}$$

For a proof, notice that when the pairs in the partition are permuted and the elements in each pair also permuted, these correspond to $2^n n!$ permutations and that the summand remains constant upon switching position of two elements in a pair as A is skew symmetric and it stays constant over permutation of these pairs as the sign, being a homomorphism on S_{2n}, is just multiplied by

$$\text{sgn}((2i-1, 2j-1)(2i, 2j)) = \text{sgn}((2i-1, 2j-1)) \cdot \text{sgn}((2i, 2j))$$

$$= (-1) \cdot (-1) = 1$$

in cycle notation, when we switch the i-th and j-th pairs - making the summand stay constant over these permutations and we are done.

Theorem 3.2 (spectral theorem for real skew symmetric matrices)

If A is a $2n \times 2n$ real skew symmetric matrix then the spectrum is of form $\{\pm i\lambda_k\}_{1 < k \leq n}$ and A is orthogonally similar to a block diagonal matrix H with diagonal entries being $H_1, \ldots, H_n$ with

$$H_k := \begin{pmatrix} 0 & \lambda_k \\ -\lambda_k & 0 \end{pmatrix}$$

Proof.

The structure of the spectrum is a result from basic linear algebra. As skew symmetric matrices are normal we can diagonalise them into $D := Q^\top A Q = \text{diag}(i\lambda_1, -i\lambda_1, \ldots)$ with Q orthogonal and thus it is sufficient to show that D is orthogonally similar to H. Suppose wlog $\lambda_1 \neq 0$. Let $v = u + iv$ be a complex vector with u, v being real, such that $Av = i\lambda v$ or, $Au + iAv = i\lambda u - \lambda v$ comparing the real and imaginary parts: $Au = -\lambda v, Av = \lambda u$.

continued.

It can be shown that u, v are orthogonal and have the same norm, thus we normalise them to orthonormal vectors u', v'. Let $W := \mathsf{Span}\{u', v'\}$, then clearly $AW \subseteq W$, let $\mathcal{B} = \{u', v'\}$ be its basis and we have

$$[A|_W]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 0 & -\lambda \\ \lambda & 0 \end{bmatrix}$$

It can be shown that $AW^{\perp} \subseteq W^{\perp}$ and we can proceed doing this inductively on $W^{\perp}$. After we are done, a suitable change of basis gives us H and the change of basis matrix has all columns being real orthonormal vectors and is orthogonal. $\qquad\square$

Theorem 3.3

For any $2n \times 2n$ skew symmetric matrix A and arbitrary matrix B we have, $\mathrm{Pf}(BAB^{\top}) = \det(B)\,\mathrm{Pf}(A)$.

For our use case however B is invertible and this version of the theorem can be proved via considering B to be a product of elementary matrices and using some of the determinant-like properties which can be found on wikipedia. Now we state the most important theorem we will use regarding pfaffians.

Theorem 3.4 (T. Muir. 1960)

If A is $2n \times 2n$ and skew symmetric then $\mathrm{Pf}(A)^2 = \det(A)$

Proof.

It is known that $P \in \mathbb{R}[x_1, \ldots, x_m]$ and $P(\mathbb{R}^m) = \{0\} \implies P \equiv 0$ (the proof is left as an exercise. Hint: Try to induct on the degree of P). Thus it suffices to prove this theorem for all real matrices. We have seen that we can write $A = QHQ^\top$ with H the same as above, for some Q. Then, $\mathrm{Pf}(A)^2 = \mathrm{Pf}(H)^2 \det(Q)^2 = \det(A)$ with the last equality left as another exercise. $\qquad\square$

Constructing the matrix

We first try to get a glimpse at the possible structure of D, only then rectifying our guess to the final matrix D which will give us $\text{Pf}(D) = |T(m, n)|$. The Pfaffian goes through all the permutations which <u>could</u> correspond to some tiling, but we want a way to filter out the *bad* permutations. A way to do this is by setting the entries $D_{i,j}$ (with $i < j$, i.e. in the upper triangular half) where i, j can not be labels on a singular domino in some canonical permutation, to zero. Then the only non zero entries could be :

1. horizontal dominos $\rightarrow$ clearly the larger even label must lie exactly to the right of the smaller odd label, unless the odd label is on the rightmost edge, in which case it can not correspond to a domino and we set it to zero.

2. vertical dominos $\rightarrow$ here, the even label is similarly always directly above the smaller odd one, unless this odd one is on the topmost edge.

Thus combining these with the definition of permutations we are interested in, we take the following guess at D, namely D': the (i,j)-th entry ($i < j$ still, entries with $i > j$ are determined by skew symmetry) is nonzero if

- i is odd, $j = i + 1$ and $m \nmid i$ (not on the right edge basically.)
- $j = i + m$ (directly above) and $i \leq mn - m$ (not on the topmost row).

Now if we make the entries which correspond to horizontal dominos x and the vertical ones y respectively, then

$$\prod_{i=1}^{mn/2} D'_{\sigma(2i-1),\sigma(2i)} = \begin{cases} x^{H(T)} y^{V(T)} & \text{if } \sigma \text{ corresponds to some tiling T} \\ 0 & \text{otherwise} \end{cases}$$

thus, using the notation developed prior to this we have

$$\text{Pf}(D') = \sum_{\pi \in T(m,n)} \text{sgn}(\sigma_\pi) x^{H(\pi)} y^{V(\pi)}$$

This was a very good guess but it doesn't work as a sign is multiplied to every summand. If we want to alter D' and make it work, then the most natural course of action is to analyse $\text{sgn}(\sigma_\pi)$ for tilings π. It is proved in [Stucky 2015] that the sign of the permutation corresponding to a tiling is the number of dominos inside odd columns, thus we can nullify the sign upfront by assigning $-y$ to the entries corresponding to these dominos, instead of y. The proof is totally elementary but best explained with visuals which I do not feel the need to reproduce here so interested readers can see it for themselves in [Stucky 2015]. Compiling everything we know about D till now, heres a full description of the nonzero entries:

- $D_{i,i+1} = x$ and $D_{i+1,i} = -x$ for $1 \leq i \leq mn$ such that $m \nmid i$.
- $D_{2i,2i+m} = y$ and $D_{2i+m,2i} = -y$ for $1 \leq i \leq m(n-1)/2$
- $D_{2i-1,2i-1+m} = -y$ and $D_{2i-1+m,2i-1} = y$ for
  $2 \leq 2i \leq m(n-1) + 1 \iff 1 \leq i \leq m(n-1)/2$

With this construction we see that D is skew symmetric and

$$\text{Pf}(D) = \sum_{\pi \in T(m,n)} x^{H(\pi)} y^{V(\pi)}$$

By looking at the entries we can see that, in block form

$$D = \begin{bmatrix} xH_m & yV_m & & & & \\ -yV_m & xH_m & yV_m & & & \\ & -yV_m & xH_m & & & \\ & & & \ddots & & \\ & & & & xH_m & yV_m \\ & & & & -yV_m & xH_m \end{bmatrix}$$

where V_k and is diagonal with $(V_k)_{ii} = (-1)^i$ and H_k has all 1s on the super and -1s on the sub diagonal respectively, every other entry being zero. V_k correspond to vertical dominos and H_k to horizontal.

Representing the matrix using Kronecker products

Definition 5.1 (kronecker product or, tensor product[1])

If A is $p \times q$ and B is $m \times n$, we define $(A \otimes B)_{ij} := a_{ij} \cdot B$ in block form. This is a $mp \times nq$ matrix.

If A, B are $m \times m$ and C, D are $n \times n$ (it is enough to consider only square matrices for our case) then,

- $(A + B) \otimes C = A \otimes C + B \otimes C$
- $C \otimes (A + B) = C \otimes A + C \otimes B$
- $(A \otimes C)(B \otimes D) = (AB) \otimes (CD)$
- for invertible M, N we have $(M \otimes N)^{-1} = M^{-1} \otimes N^{-1}$

proofs are left as an exercise.

---

[1]There are connections aplenty between this and the conventional Tensor products but here this is used interchangeably with the kronecker product

Using these, notice that we can write $D = I_n \otimes xH_m + H_n \otimes yV_m$. As H_k are skew symmetric, they are normal and hence diagonalisable. Suppose we have $U_k^{-1} H_k U_k = \Lambda_k$, a diagonal matrix. To get more information about the determinant of D, its natural to try to diagonalise it or atleast block diagonalise it. We try to diagonalise the H_k in our expression by considering $\tilde{\Lambda} := (U_n \otimes U_m)^{-1} D (U_n \otimes U_m)$, where U_m hits the H_m on the right term and U_n the H_n on the left. By using the properties mentioned above we find that

$$\tilde{\Lambda} = (U_n \otimes U_m)^{-1} D (U_n \otimes U_m) = I_n \otimes x\Lambda_m + \Lambda_n \otimes yU_m^{-1} V_m U_m$$

this is clearly in block diagonal form as I_n, Λ_n are both diagonal and of the same dimension. We will now try to figure out what $U_m^{-1} V_m U_m$ is.

Evaluating $U_m^{-1} V_m U_m$

Using cofactor expansion on the determinant we see that
$\chi_{H_k}(x) = x \cdot \chi_{H_{k-1}}(x) + \chi_{H_{k-2}}(x)$ with $\chi_{H_2} = x^2 + 1$ and $\chi_{H_1} := x$
to make the recurrence work. We can solve this linear recurrence
to get that,

$$\chi_{H_k}(x) = \frac{(x + \sqrt{x^2 + 4})^{k+1} - (x - \sqrt{x^2 + 4})^{k+1}}{2^{k+1}\sqrt{x^2 + 4}}$$

From this we can find the roots (thus the eigenvalues) to be
$\lambda_j^{(k)} = 2i \cos(j\pi/(k+1))$ for $j = 1, \ldots, k$. We can solve the
system $\lambda_j^{(k)} \cdot v = H_k v$ for eigenvectors v as well, this reduces to
finding another term in a linear recurrence and after some algebra
we get an eigenvector $v_j^{(k)} = (i^m \sin(mj\pi/(k+1)))_{1 \le m \le k}^{\top}$ for the
eigenvalue $\lambda_j^{(k)}$. Now as all the eigenvalues were distinct, the
corresponding eigenvectors are known to be linearly independent
and we can form a change of basis matrix U_m with $(U_m)_{*,j} := v_j^{(k)}$.
Then we see that $\Lambda_k = U_m^{-1} H_k U_m$ so we found what we wanted.

Calculating $V_m U_m$ is easy, but working with U_m^{-1} probably not so much. But we can notice that,

$$(V_m U_m)_{k,l} = (-1)^k i^k \sin\left(\frac{\pi k l}{m+1}\right) = (U_m)_{k,m+1-l}$$

with some trigonometry, what this tells us is that $V_m U_m = U_m J_m$ where J_m is the anti-diagonal identity matrix (what we get by flipping I_m along the horizontal axis.) Thus, simplifying everything we get,

$$\tilde{\Lambda} = I_n \otimes x\Lambda_m + \Lambda_n \otimes y J_m$$

This is block diagonal with $m \times m$ blocks. The block with its bottom right entry being at the (km, km)-th position is thus

$$x\Lambda_m + y\lambda_k^{(n)} J_m$$

Manipulating the blocks

Using this we see that the k-th block is

$$
\begin{bmatrix}
x\lambda_1^{(m)} & & & & & & & y\lambda_k^{(n)} \\
& x\lambda_2^{(m)} & & & & & y\lambda_k^{(n)} & \\
& & \ddots & & & \iddots & & \\
& & & x\lambda_{m/2}^{(m)} & y\lambda_k^{(n)} & & & \\
& & & y\lambda_k^{(n)} & x\lambda_{m/2+1}^{(m)} & & & \\
& & \iddots & & & \ddots & & \\
& y\lambda_k^{(n)} & & & & & x\lambda_{m-1}^{(m)} & \\
y\lambda_k^{(n)} & & & & & & & x\lambda_m^{(m)}
\end{bmatrix}
$$

As m is even, both the diagonals are *supposed to miss each other.*

Firstly the determinant of $\tilde{\Lambda}$ is the product of those of these blocks, this is standard. Focus on the k-th block now. Swap rows 2 and m, then swap columns 2 and m hence making the top-left and bottom-right 2×2 submatrix

$$\begin{bmatrix} x\lambda_1^{(m)} & y\lambda_k^{(n)} \\ y\lambda_k^{(n)} & x\lambda_m^{(m)} \end{bmatrix}, \quad \begin{bmatrix} x\lambda_{m-1}^{(m)} & y\lambda_k^{(n)} \\ y\lambda_k^{(n)} & x\lambda_2^{(m)} \end{bmatrix}$$

respectively. We can keep repeating this swapping on the smaller and smaller blocks in the middle until the matrix is a diagonal block matrix with 2×2 blocks. The blocks left will be

$$\begin{bmatrix} x\lambda_j^{(m)} & y\lambda_k^{(n)} \\ y\lambda_k^{(n)} & x\lambda_{m+1-j}^{(m)} \end{bmatrix} \text{ for } j = 1, 2, \ldots, m/2$$

And as, for every column swap we did exactly one row swap, the sign of the determinant of this $m \times m$ block does not change.

Putting it all together

Now we can find the determinant of matrix $\tilde{\Lambda}$ as,

$$\det \tilde{\Lambda} = \prod_{k=1}^{n} \prod_{j=1}^{m/2} \det \left(\begin{bmatrix} x\lambda_j^{(m)} & y\lambda_k^{(n)} \\ y\lambda_k^{(n)} & x\lambda_{m+1-j}^{(m)} \end{bmatrix} \right)$$

$$= \prod_{k=1}^{n} \prod_{j=1}^{m/2} \left(x^2 \lambda_j^{(m)} \lambda_{m+1-j}^{(m)} - y^2 \lambda_k^{(n)} \lambda_k^{(n)} \right)$$

$$= \prod_{k=1}^{n} \prod_{j=1}^{m/2} \left(4x^2 \cos^2 \frac{\pi j}{m+1} + 4y^2 \cos^2 \frac{\pi k}{n+1} \right)$$

as $\cos(\pi - \theta) = -\cos\theta$. Now, using Muir's result we can link all of this back to the Pfaffian :

$$\mathrm{Pf}(D)^2 = \det(D) = \det\left((U_n \otimes U_m) \tilde{\Lambda} (U_n \otimes U_m)^{-1} \right) = \det \tilde{\Lambda}$$

With what we have done so far, we have

$$
\sum_{\pi \in T(m,n)} x^{H(\pi)} y^{V(\pi)} = \mathrm{Pf}(D)
$$

$$
= \sqrt{\det \tilde{\Lambda}}
$$

$$
= 2^{mn/2} \prod_{k=1}^{n} \prod_{j=1}^{m/2} \left(x^2 \cos^2 \frac{\pi j}{m+1} + y^2 \cos^2 \frac{\pi k}{n+1} \right)^{\frac{1}{2}}
$$

and we are finally done!

References

📄 Kasteleyn, P. W. (Dec. 1961). "The statistics of dimers on a lattice: I. The number of dimer arrangements on a quadratic lattice". In: *Physica* 27, pp. 1209–1225.

📄 Stucky, Eric (2015). "An Exposition of Kasteleyn's Solution of the Dimer Model". Senior Thesis. Harvey Mudd College. URL: https://scholarship.claremont.edu/hmc_theses/89.

Questions?