well these asymptotic expressions work even for the smallest lattices. For example, the truncated formulas (7.16), (7.17), and (8.4) give the number of dimer configurations for the plane 2×2, 2×3, 3×4, and 4×4 lattices with edges as 1.96, 2.98, 11.07, and 35.71, respectively. To the nearest whole number, these are precisely the exact values 2, 3, 11, and 36, respectively. However, for the 8×8 chessboard, the truncated expression gives to the nearest integer 12 957 925, compared with the exact value 12 988 816 given by Fisher[6] for the number of dimer configurations. Despite this numerical difference, the fractional error is only 0.24%.

ACKNOWLEDGMENTS

The author would like to thank Professor M. E. Fisher for suggesting this problem and for his interest and guidance during the course of the work. He would also like to thank the Advanced Research Projects Agency for support through the Materials Science Center at Cornell, as well as the University of the West Indies for a post-graduate award.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 12 DECEMBER 1967

Solution of the Dimer Problem by the Transfer Matrix Method

ELLIOTT H. LIEB*
Physics Department, Northeastern University, Boston, Massachusetts

(Received 25 May 1967)

It is shown how the monomer–dimer problem can be formulated in terms of a transfer matrix, and hence in terms of simple spin operators as was originally done for the Ising problem. Thus, we rederive the solution to the pure dimer problem without using Pfaffians. The solution is extremely simple once one sees how to formulate the transfer matrix.

1. INTRODUCTION

SINCE Onsager's solution[1] of the two-dimensional Ising model in 1944, there has been a great deal of activity in the general area of nearest neighbor planar lattice problems. Basically, two approaches have been used.[2] One is the "algebraic" or "transfer matrix" method (used by Onsager) which focuses attention on the manner in which two neighboring rows are connected to each other. The second is the so-called "combinatorial method" whereby one studies graphs on the lattice as a whole. This was first used by Kac and Ward[3] for the Ising problem.

The most recent and concise formulation of the combinatorial method reduces the various problems to the evaluation of a Pfaffian. This method had its first "original" success in the solution to the dimer problem by Kasteleyn,[4] and for this reason is also referred to as the dimer method.

While Pfaffians have been used to rederive the solution to the Ising problem,[5] no one has yet taken the complementary step of solving the dimer problem by the transfer matrix method. The purpose of this note is to eliminate this gap. Elsewhere,[6] it has been shown how the transfer matrix method for Ising-like problems can be reduced to a few simple steps involving only fermion creation and annihilation operators. The dimer problem is likewise simple, using the transfer matrix. We also show how the more difficult and unsolved monomer–dimer problem can be formulated this way. The analogy with the problem of the Ising model in a magnetic field is very transparent, but the monomer–dimer problem is somewhat simpler and therefore there is considerable hope that this new formulation of the problem may ultimately lead to its solution.

2. FORMULATION OF THE TRANSFER MATRIX

We have a square planar lattice of M rows and N columns and hence MN vertices. A dimer is a rigid rod just long enough to cover two neighboring

* Work supported by National Science Foundation Grant GP-6851.

[1] L. Onsager, Phys. Rev. **65**, 117 (1944).

[2] We make no attempt to give a complete bibliography because an excellent one is given in H. S. Green and C. A. Hurst, *Order-Disorder Phenomena* (Interscience Publishers, Inc., New York, 1964).

[3] M. Kac and J. C. Ward, Phys. Rev. **88**, 1332 (1952).

[4] P. W. Kasteleyn, Physica **27**, 1209 (1962); J. Math. Phys. **4**, 287 (1963); H. N. V. Temperley and M. E. Fisher, Phil. Mag. **6**, 1061 (1961); M. E. Fisher, Phys. Rev. **124**, 1664 (1961).

[5] C. A. Hurst and H. S. Green, J. Chem. Phys. **33**, 1059 (1960).

[6] T. D. Schultz, D. C. Mattis, and E. H. Lieb, Rev. Mod. Phys. **36**, 856 (1964).

vertices (either horizontally or vertically), while a monomer covers just one vertex. An allowed configuration of the lattice is one in which every vertex is covered by a monomer or a dimer such that no vertex is covered by more than one object. If h, v, and m, are, respectively, the numbers of horizontal dimers, vertical dimers, and monomers in a configuration (with $2h + 2v + m = MN$), then the partition function Z is

$$\begin{aligned} Z &= \sum \binom{\text{ALLOWED}}{\text{CONFIGURATIONS}} x^h y^v z^m \\ &= y^{\frac{1}{2}MN} \sum \binom{\text{ALLOWED}}{\text{CONFIGURATIONS}} \alpha^h \beta^m \end{aligned} \tag{2.1}$$

where x, y, and z are the appropriate "activities" and $\alpha = x/y$, $\beta = z/(y)^{\frac{1}{2}}$.

Beginning at the "bottom" of the lattice, we have a row of N vertical bonds (V bonds) followed by a row of N horizontal bonds (H bonds), and so on alternately. There are M bond rows of each type. On each V bond of the first row, we place an arrow (or spin): an up arrow (spin +) signifies the presence of a vertical dimer on that bond, while a down arrow (spin −) signifies the absence of a dimer. The next row consists of NH bonds alternating with N vertices. If an arrow points into one of these vertices (i.e., spin + on the previous V bond) that vertex is saturated and only a down arrow must be allowed to propagate out of that vertex along the next V bond. If, however, a down arrow (spin −) comes into a vertex from below, one of three things are allowed to happen: (1) an up arrow (= dimer = spin +) can propagate along the next V bond; (2) two neighboring down arrows can cóoperate to form a horizontal dimer on the intervening H bond, which means that down arrows must propagate upwards from those two vertices; (3) a down arrow can propagate upwards from the vertex in question, signifying that a monomer has been placed on that vertex.

These rules seem complicated but they can be simply formulated if one decomposes the activity on the row of vertices into two simple steps.

Step 1: Reverse all the incoming arrows. The operator that accomplishes this is simply

$$V_1 = \prod_{i=1}^{N} \sigma_i^x, \tag{2.2}$$

where σ^x is the Pauli spin matrix $\sigma^x = \left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$. The operator V_1 guarantees that an incoming up arrow (spin +) will go out as a down arrow, as required. Contrariwise, an incoming down arrow (no dimer) will propagate upwards as a V dimer (spin +). To create monomers or H dimers on the row of H bonds under consideration, it is necessary to convert some of these up arrows to down arrows and this is accomplished by the second step.

Step 2: To create a monomer at the jth vertex of the row, we multiply by σ_j^-, where $\sigma^- = \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right)$. To create a horizontal dimer at the adjacent sites j and $j+1$, we multiply by $\sigma_j^-\sigma_{j+1}^-$. Note that after multiplication by these operators, we are insured that only down arrows (i.e., no vertical dimers) propagate onto the next row of V bonds.

The operator which generates exactly m monomers on the row is $(m!)^{-1}(\sum_{i=1}^{N} \sigma_i^-)^m$ because $(\sigma_i^-)^2 = 0$. Thus, to generate an arbitrary number of monomers with the factor β^m, we multiply by the operator

$$V_2 = \exp\left(\beta \sum_{i=1}^{N} \sigma_i^-\right). \tag{2.3}$$

Likewise, to generate H dimers, we multiply by

$$V_3 = \exp\left(\alpha \sum_{i=1}^{N} \sigma_i^- \sigma_{i+1}^-\right). \tag{2.4}$$

In Eq. (2.4), we have used cyclic boundary conditions with $\sigma_{N+1}^- \equiv \sigma_1^-$. If free ends are desired, then omit the term $\sigma_1^-\sigma_N^-$.

Thus, our transfer matrix is

$$V = V_3 V_2 V_1, \tag{2.5}$$

and

$$Z = y^{\frac{1}{2}MN} \operatorname{Tr} V^M, \tag{2.6}$$

where we have used cyclic boundary conditions in the vertical direction. We note that free edge (noncyclic) boundary conditions can be used in either or both directions with no great complication, although we shall confine ourselves here to cyclic boundary conditions.

The problem is thus reduced to computing the eigenvalues of V.

3. DIAGONALIZATION OF THE TRANSFER MATRIX

It is convenient, although not essential, to eliminate V_1. To do so, we consider $V^2 = V_3V_2V_1\, V_3V_2V_1$ and make use of the fact that σ_i^x is a unitary operator with $\sigma_i^x\sigma_i^-\sigma_i^x = \sigma_i^+ = \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right)$. Thus,

$$V^2 = V_3V_2\, \bar{V}_3\bar{V}_2 \tag{3.1}$$

with

$$\begin{aligned} \bar{V}_2 &= \exp\left(\beta \sum_{i=1}^{N} \sigma_i^+\right), \\ \bar{V}_3 &= \exp\left(\alpha \sum_{i=1}^{N} \sigma_i^+ \sigma_{i+1}^+\right). \end{aligned} \tag{3.2}$$

Thus,

$$Z = y^{\frac{1}{2}MN} \sum_j \lambda_j(N)^{\frac{1}{2}M}, \tag{3.3}$$

where $\lambda_j(N)$ (for $j = 1, 2, \cdots, 2^N$) are the eigenvalues of V^2 (we have assumed M is even).

The operator V contains two kinds of operators: quadratic forms (as in V_3 and $\bar{V}_3$) and linear forms (as in V_2 and $\bar{V}_2$). It is similar to, but simpler than, the problem of the Ising model in a magnetic field, because V_2 and V_3 commute and because V_2 is the adjoint of $\bar{V}_2$ and V_3 is the adjoint of $\bar{V}_3$. Nevertheless, we still are unable to handle the linear forms and henceforth we shall set $\beta = 0$ (i.e., we consider only the pure dimer problem).

The appropriate steps to diagonalize $V_3\bar{V}_3$ are well known. The details are in Ref. 6 and it suffices to outline the steps here. First, we transform from paulions to fermions:

$$\begin{aligned} C_j &\equiv (-1)^{j-1}\left(\prod_{i=1}^{j-1}\sigma_i^z\right)\sigma_j^-, \\ C_j^\dagger &\equiv (-1)^{j-1}\left(\prod_{i=1}^{j-1}\sigma_i^z\right)\sigma_j^+, \end{aligned} \tag{3.4}$$

with $\sigma^z = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$.

This results in

$$\begin{aligned} \sigma_j^-\sigma_{j+1}^- &= -C_jC_{j+1}, \\ \sigma_j^+\sigma_{j+1}^+ &= C_j^\dagger C_{j+1}^\dagger, \end{aligned} \tag{3.5}$$

while

$$\begin{aligned} \sigma_1^-\sigma_N^- &= (-1)^{\mathcal{N}}C_NC_1, \\ \sigma_1^+\sigma_N^+ &= -(-1)^{\mathcal{N}}C_N^\dagger C_1^\dagger, \end{aligned} \tag{3.6}$$

where $\mathcal{N} = \sum_1^N C_i^\dagger C_i$ is the number operator.

Since $(-1)^{\mathcal{N}}$ commutes with a quadratic form in fermions (or paulions), it is a constant of the motion $= +1$ for states of even $\mathcal{N}$ and -1 for states of odd $\mathcal{N}$. The latter case is cyclic while the former is anticyclic.

Next we go to running waves:

$$C_j = N^{-\frac{1}{2}}e^{-i\pi/4}\sum_q e^{iqj}\eta_q, \tag{3.7}$$

with

$$q = \pm\pi/N, \pm 3\pi/N, \cdots, \pm(N-1)\pi/N \tag{3.8a}$$

for $\mathcal{N}$ even, and

$$q = 0, \pm 2\pi/N, \pm 4\pi/N, \cdots, \pm(N-2)\pi/N, \pi \tag{3.8b}$$

for $\mathcal{N}$ odd. (We have assumed that N is even.) In terms of these running waves, we have

$$V^2 = \prod_{0\le q\le\pi} A_q \tag{3.9}$$

with

$$A_q = \exp(2\alpha \sin q\eta_q\eta_{-q})\exp(2\alpha \sin q\eta_{-q}^\dagger\eta_q^\dagger). \tag{3.10}$$

Obviously, A_q and $A_{q'}$ commute with each other. An eigenvalue of V^2 is thus the product of eigenvalues of each A_q, but we must remember to use the appropriate q values [(3.8a) or (3.8b)] according to whether the total $\mathcal{N}$ value of a state is even or odd.

To diagonalize A_q (for $q \neq 0$ and $q \neq \pi$), we first list the four basis states,

$$\begin{aligned} &\Phi_0 = |0\rangle, \quad \Phi_q = \eta_q^\dagger|0\rangle, \quad \Phi_{-q} = \eta_{-q}^\dagger|0\rangle, \\ &\Phi_{-qq} = \eta_{-q}^\dagger\eta_q^\dagger|0\rangle. \end{aligned} \tag{3.11}$$

The states Φ_q and Φ_{-q} are pure odd states and are eigenvectors of A_q with eigenvalue unity. The other two states are pure even and

$$\begin{aligned} A_q\Phi_0 &= (1 + 4\alpha^2\sin^2 q)\Phi_0 + 2\alpha \sin q\Phi_{-qq}, \\ A_q\Phi_{-qq} &= 2\alpha \sin q\Phi_0 + \Phi_{-qq}. \end{aligned} \tag{3.12}$$

The two even eigenvalues of A_q are thus

$$\lambda_{\text{even}} = [\alpha \sin q \pm (1 + \alpha^2\sin^2 q)^{\frac{1}{2}}]^2. \tag{3.13}$$

For $q = 0$ or π there are only two states: $\Phi_0 = |0\rangle$ (even) and $\Phi_q = \eta_q^\dagger|0\rangle$ (odd). Both have eigenvalue unity. These values of q are effective only for $\mathcal{N}$ odd and these pairs of states may be used to insure that $\mathcal{N}$ is always odd without affecting the total eigenvalue.

Thus, from (3.3),

$$\begin{aligned} y^{-\frac{1}{2}MN}Z = 2 &\prod_{\substack{0<q<\pi \\ q\in(3.8b)}} \{[(1 + \alpha^2\sin^2 q)^{\frac{1}{2}} + \alpha \sin q]^M \\ &+ [(1 + \alpha^2\sin^2 q)^{\frac{1}{2}} - \alpha \sin q]^M + (1 + 1)\} \\ &+ \tfrac{1}{2}\prod_{\substack{0<q<\pi \\ q\in(3.8a)}} \{[(1 + \alpha^2\sin^2 q)^{\frac{1}{2}} + \alpha \sin q]^M \\ &+ [(1 + \alpha^2\sin^2 q)^{\frac{1}{2}} - \alpha \sin q]^M + (1 + 1)\} \\ &+ \tfrac{1}{2}\prod_{\substack{0<q<\pi \\ q\in(3.8a)}} \{[(1 + \alpha^2\sin^2 q)^{\frac{1}{2}} + \alpha \sin q]^M \\ &+ [(1 + \alpha^2\sin^2 q)^{\frac{1}{2}} - \alpha \sin q]^M - (1 + 1)\}. \end{aligned} \tag{3.14}$$

The first product in (3.14) gives the contribution of the $\mathcal{N}$ odd eigenvalues while the sum of the second and third products gives the contribution of the $\mathcal{N}$ even eigenvalues. [Eq. (3.14) agrees exactly with Kasteleyn's result (Ref. 4, Eq. (25)), as may be seen using Kasteleyn's identity (Ref. 4, Eq. (26)).]

For a large lattice ($MN \to \infty$), we need consider only the largest eigenvalue of V^2, which means the product on all q of Eq. (3.13), using the plus sign. In this limit, it clearly is immaterial which set we use, (3.8a) or (3.8b). Thus,

$$\begin{aligned} &\lim_{M,N\to\infty} (MN)^{-1}\ln Z \\ &\quad = \frac{1}{2\pi}\int_0^\pi dq \ln[\alpha \sin q + (1 + \alpha^2\sin^2 q)^{\frac{1}{2}}] \\ &\qquad + \tfrac{1}{2}\ln y, \end{aligned} \tag{3.15}$$

which agrees with previous results[4] for the pure dimer problem.