JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 3 MARCH 1969

Negative Finding for the Three-Dimensional Dimer Problem*

J. M. HAMMERSLEY, A. FEUERVERGER, A. IZENMAN, AND K. MAKANI
Statistical Laboratory, University of California, Berkeley

(Received 30 August 1968)

The dimer problem can be solved if one can evaluate the permanent of $\mathbf{P} = (p_{ij})$, the incidence matrix of the lattice. All known methods of solving the two-dimensional case consist (explicitly or implicitly) in finding another matrix $\mathbf{Q} = (q_{ij})$, such that $p_{ij} = |q_{ij}|$ and per $\mathbf{P} = |\det \mathbf{Q}|$, and then computing the determinant of $\mathbf{Q}$. We show that in the three-dimensional case no such matrix $\mathbf{Q}$ exists for any choice of elements q_{ij}, whether real or complex numbers, or quaternions. A stronger negative result of an asymptotic character seems to be true, but this rests upon a plausible but unproved conjecture.

INTRODUCTION AND STATEMENT OF RESULTS

Let a, b, c be positive integers with $N = abc$ even, and define the lattice L to be the set of N points in three-dimensional Euclidean space with integer coordinates (x, y, z) such that $1 \leq x \leq a$, $1 \leq y \leq b$, $1 \leq z \leq c$. A dimer is a pair of points of L which are unit distance apart; and a dimer configuration is a partitioning of L into $\frac{1}{2}N$ disjoint dimers. Let f denote the number of dimer configurations on L. It can be proved[1] that $N^{-1} \log f$ tends to a limit (denoted by λ) as $a \to \infty$, $b \to \infty$, $c \to \infty$ independently. (For brevity, we hereafter write $N \to \infty$ to signify $a, b, c \to \infty$.) The dimer problem is to determine f as a function of a, b, c and hence (or otherwise) to calculate λ.

Number the points of L from 1 to N in a fixed arbitrary way, and write $p_{ij} = 1$ or 0, according as the ith and jth points of L are or are not unit distance apart. The $N \times N$ matrix $\mathbf{P} = (p_{ij})$ is called the incidence matrix of L; and it can be shown[2] that $f^2 = \text{per}\,\mathbf{P}$, the permanent of $\mathbf{P}$. Thus a solution of the dimer problem is equivalent to an evaluation of this permanent. Unlike determinants, to which they bear a superficial algebraic resemblance, permanents do not enjoy any practicable algorithms for their evaluation when N is large. However, most of the elements of $\mathbf{P}$ are zero, and this has suggested the possibility of finding another matrix $\mathbf{Q}$, such that

$$p_{ij} = |q_{ij}| \quad \text{and} \quad \text{per}\,\mathbf{P} = |\det \mathbf{Q}|, \tag{1}$$

and so calculating f via det $\mathbf{Q}$. Here the q_{ij} are real or complex numbers or quaternions; and, if q is a real or complex number, $|q|$ denotes its modulus in the ordinary sense; while, if q is a quaternion, its modulus $|q|$ is the positive square root of its norm. (The possibility of using quaternions in this context has not, so far as we know, been mentioned in the literature, but from private conversations we know that this idea has occurred independently to several colleagues; for, indeed, there are anticommuting features in the dimer problem which lend appeal to use of quaternions as a tool.) We discuss below two methods of defining the determinant of a matrix of quaternions.

In the two-dimensional case (i.e., when $a = 1$ and only $b \to \infty$ and $c \to \infty$) all known methods of solving the dimer problem depend, explicitly or implicitly, on finding a solution of (1), and a variety of such real and complex solutions are known. Here we prove that no solutions exist in the three-dimensional case: specifically, we show that, when $a \geq 2$, $b \geq 4$, and $c \geq 4$, then

$$p_{ij} = |q_{ij}| \Rightarrow \text{per}\,\mathbf{P} > |\det \mathbf{Q}|, \tag{2}$$

for any choice of real or complex or quaternion $\mathbf{Q}$. This does not, however, completely dispose of (1) as a device for computing λ; for, despite (2), it might still be true that

$$\limsup_{N\to\infty} \tfrac{1}{2}N^{-1} \log |\det \mathbf{Q}| = \lim_{N\to\infty} \tfrac{1}{2}N^{-1} \log \text{per}\,\mathbf{P} = \lambda, \tag{3}$$

when $p_{ij} = |q_{ij}|$. We believe that (3) is actually false; but the best we can do in this direction is to deduce the falsity of (3) from the following plausible but unproved conjecture. Define a *block* of L to be a set of 32 points of L whose coordinates (x, y, z) satisfy $\xi \leq x < \xi + 2$, $\eta \leq y < \eta + 4$, $\zeta \leq z < \zeta + 4$ for some integers ξ, η, ζ. Thus L contains $(a - 1)(b - 3)(c - 3)$ different blocks when $a \geq 2$, $b \geq 4$, $c \geq 4$. Given a dimer configuration on L, we say that a particular block is *smooth* if there is no dimer of the configuration with one of its points in this block and the other point not in the block. A dimer configuration is called *rough* if no block of L is smooth. Let g denote the number of rough configurations on L. We conjecture

* This investigation was partially supported by research grant No. GM 10525-05 from the National Institutes of Health, Public Health Service.

[1] J. M. Hammersley, "Existence Theorems and Monte Carlo Methods for the Monomer-Dimer Problem" in *Research Papers in Statistics: Festschrift for J. Neyman* (John Wiley & Sons, Inc., New York, 1966), pp. 125–146.

[2] J. M. Hammersley, Proc. Cambridge Phil. Soc. **64**, 455 (1968).

that

$$\liminf_{N\to\infty} N^{-1} \log g < \lambda. \tag{4}$$

However, we do not give here the proof that (3) is false if (4) is true.

DETERMINANTS OF QUATERNION MATRICES

The literature contains two slightly different definitions of the determinant of a matrix **Q** with quaternion elements. They are due to Moore[3] and Dieudonné,[4] and we denote them by $\det_M \mathbf{Q}$ and $\det_D \mathbf{Q}$, respectively. We recall that a quaternion can be written $q = \alpha + \beta i + \gamma j + \delta k$, where α, β, γ, δ are real numbers and i, j, k are indeterminates satisfying $i^2 = j^2 = k^2 = ijk = -1$; that the conjugate of q is $\bar{q} = \alpha - \beta i - \gamma j - \delta k$; that the norm of q is $N(q) = q\bar{q} = \bar{q}q = \alpha^2 + \beta^2 + \gamma^2 + \delta^2$; and we write $|q|$ for the modulus of q, i.e., the positive square root of $N(q)$. Then $|q_1q_2| = |q_2q_1| = |q_1|\,|q_2|$ for any quaternions q_1, q_2. We write $\mathbf{Q}^*$ for the transposed conjugate of a quaternion matrix **Q**; we call **Q** Hermitian if $\mathbf{Q} = \mathbf{Q}^*$, i.e., if $q_{ij} = \bar{q}_{ji}$. All matrices mentioned below are square matrices with quaternion elements, unless the contrary is explicitly stated.

Dieudonné's paper deals with a slightly more general case than we need. Reduced to the quaternion case in hand, and stripped of its abstract terminology, it boils down to the following. If **Q** is a diagonal matrix, $\det_D \mathbf{Q}$ is defined to be $|q|$, where q is the product of the diagonal elements of **Q**. For general **Q**, the value of $\det_D \mathbf{Q}$ is (by definition) unchanged if, to any row of **Q** we add a constant multiple of any other row, it being understood that the constant multiplier (which is a quaternion) acts as a left-hand multiplier of the row. Similarly the value is unchanged for similar operations on columns, the constant multiplier now being a right-hand multiplier of the column. The value of $\det_D \mathbf{Q}$ also is unchanged by any permutation of rows or of columns of **Q**. As with ordinary determinants, these row and column operations let us reduce a general **Q** to diagonal form, and so to determine the value of $\det_D \mathbf{Q}$. Dieudonné shows that the foregoing requirements are self-consistent and uniquely determine $\det_D \mathbf{Q}$, and that $\det_D (\mathbf{Q}_1\mathbf{Q}_2) = \det_D \mathbf{Q}_1 \det_D \mathbf{Q}_2$ for any two matrices $\mathbf{Q}_1$, $\mathbf{Q}_2$.

Moore's definition applies only to the case when **Q** is Hermitian. Suppose **Q** has N rows and N columns, and write Z for the set $\{1, 2, \cdots, N\}$. Let z be some given nonempty subset of Z, and suppose that z has s elements. Define

$$q(z, i_1) = \sum (-1)^{s-1} q_{i_1 i_2} q_{i_2 i_3} \cdots q_{i_{s-1} i_s} q_{i_s i_1}, \tag{5}$$

where i_1 is a selected element of z, and the sum in (5) is taken over all permutations of the remaining unselected elements $i_2, i_3, \cdots, i_s$ of z. Thus there are $(s-1)!$ summands in (5). The Hermitian character of **Q** ensures, as Moore proves, that $q(z, i_1)$ is a real number (i.e., a quaternion with $\beta = \gamma = \delta = 0$) and that $q(z, i_1)$ is independent of the choice of i_1 in z. We may thus write $q(z)$ in place of $q(z, i_1)$ and regard the sum in (5) as being taken over the $(s-1)!$ different cycles which can be formed from the elements of z. Next, let T be a partition of Z into disjoint nonempty subsets $z_1^T, z_2^T, \cdots, z_t^T$ (whose union is Z, of course); and define

$$\det_M \mathbf{Q} = \sum_T q(z_1^T) q(z_2^T) \cdots q(z_t^T), \tag{6}$$

where the sum in (6) is over all possible distinct partitions of Z. [As is usual in a partition, the order of the parts and the order of the elements in each part is immaterial; but the order of the parts does not affect the definition (6), because the $q(z^T)$ are all real and therefore commute; and the order of the elements in each part does not affect the definition (5), because $q(z, i_1)$ is independent of the selection i_1 and the summation in (5) is over all $(s-1)!$ permutations of the remaining unselected elements.]

Moore also proves that, if $\mathbf{Q}_0$ is an arbitrary quaternion matrix and if **Q** is Hermitian, then $\mathbf{Q}_0\mathbf{Q}_0^*$ and $\mathbf{Q}_0\mathbf{Q}\mathbf{Q}_0^*$ are both Hermitian and

$$\det_M (\mathbf{Q}_0\mathbf{Q}\mathbf{Q}_0^*) = \det_M (\mathbf{Q}_0\mathbf{Q}_0^*) \det_M \mathbf{Q}.$$

In particular, if we take the diagonal elements of $\mathbf{Q}_0$ to be all 1, and all the nondiagonal elements, except just one of them, to be zero, then it is easy to verify from (5) that $\det_M (\mathbf{Q}_0\mathbf{Q}_0^*) = 1$. However, by choosing a succession of such $\mathbf{Q}_0$'s to premultiply and postmultiply **Q** in the fashion of $\mathbf{Q}_0\mathbf{Q}\mathbf{Q}_0^*$ we can reduce **Q** to diagonal form, just as with ordinary Hermitian transformations.[5] Since clearly $\det_D \mathbf{Q}_0 = \det_D \mathbf{Q}_0^* = 1$, we can prove in this way that

$$\det_D \mathbf{Q} = |\det_M \mathbf{Q}| \tag{7}$$

for any Hermitian matrix **Q**. It follows that the notation $|\det \mathbf{Q}|$ may be used without ambiguity for either $|\det_D \mathbf{Q}|$ or $|\det_M \mathbf{Q}|$ when **Q** is Hermitian. (It is easy to see from simple examples that $\det_D \mathbf{Q} = \det_M \mathbf{Q}$ is not always true for Hermitian matrices **Q**.)

[3] E. H. Moore, *General Analysis, Part I* (Memoirs series, Vol. 1, The American Philosophical Society, Philadelphia, Pa., 1935).

[4] J. Dieudonné, Bull. Soc. Math. de France **71**, 27 (1943).

[5] H. W. Turnbull and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices* (Blackie & Son Ltd., Glasgow, 1932), p. 85.

PROOF OF EQUATION (2)

We assume throughout that $a \geq 2$, $b \geq 4$, $c \geq 4$, and show that

$$p_{ij} = |q_{ij}| \quad \text{and} \quad \text{per}\,\mathbf{P} \leq |\det \mathbf{Q}| \tag{8}$$

leads to a contradiction. The case when the q_{ij} are real or complex numbers is a particular case of quaternions q_{ij} with $\gamma = \delta = 0$ for all elements, since then the modulus of the ordinary determinant of $\mathbf{Q}$ coincides with $\det_D \mathbf{Q}$, as the above definition of the latter shows. Hence we may suppose that $\mathbf{Q}$ is a quaternion matrix satisfying (8). Color the points of L black and white after the fashion of a chessboard, i.e., all points of L at unit distance from a white point shall be black and vice versa. We say that the ith row of $\mathbf{Q}$ is black or white according as the ith point of L is black or white. Let T be the permutation of rows of $\mathbf{Q}$ which places all the black rows before all the white rows, while leaving the relative order of the black rows among themselves unchanged and similarly preserving the relative order of the white rows. Apply this permutation to the rows of $\mathbf{Q}$ and the *same* permutation to the columns of $\mathbf{Q}$. Since each dimer contains one black and one white point wherever it may be on L, $\mathbf{Q}$ is transformed to the form

$$\mathbf{Q}' = \begin{pmatrix} \mathbf{0} & \mathbf{Q}_1 \\ \mathbf{Q}_2 & \mathbf{0} \end{pmatrix}$$

where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are $\frac{1}{2}N \times \frac{1}{2}N$ quaternion matrices. Then from (8) we have

$$f^2 = \text{per}\,\mathbf{P} \leq |\det \mathbf{Q}| = \det_D \mathbf{Q} = \det_D \mathbf{Q}' = \det_D \mathbf{Q}_1 \det_D \mathbf{Q}_2. \tag{9}$$

Hence there exists $\mathbf{Q}_0$, equal to one or another of $\mathbf{Q}_1$ or $\mathbf{Q}_2$, such that $f \leq \det_D \mathbf{Q}_0$. But $\det_D \mathbf{Q}_0 = \det_D \mathbf{Q}_0^*$ from the definitions. Hence,

$$\text{per}\,\mathbf{P} = f^2 \leq \det_D \mathbf{Q}_0 \det_D \mathbf{Q}_0^* = \det_D \begin{pmatrix} \mathbf{0} & \mathbf{Q}_0 \\ \mathbf{Q}_0^* & \mathbf{0} \end{pmatrix} = \det_D \mathbf{Q}'' = |\det_D \mathbf{Q}''|, \tag{10}$$

where $\mathbf{Q}''$ is the matrix obtained by applying the inverse permutation T^{-1} to both the rows and columns of $\left(\begin{smallmatrix} \mathbf{0} & \mathbf{Q}_0 \\ \mathbf{Q}_0^* & \mathbf{0} \end{smallmatrix}\right)$. We have $p_{ij} = |q''_{ij}|$; and $\mathbf{Q}''$ is Hermitian. Thus if any solution $\mathbf{Q}$ of (8) exists, there is at least one Hermitian solution of (8). Hereafter we suppose that $\mathbf{Q}$ is such a Hermitian solution of (8); and accordingly we may now interpret $\det \mathbf{Q}$ as $\det_M \mathbf{Q}$.

Let $P_1, P_2, \cdots, P_N$ denote the points of L in the fixed enumeration used for specifying the incidence matrix $\mathbf{P}$. We define an (oriented) polygon on L as a cyclic sequence of distinct points of L, say $(P_{j_1}P_{j_2}\cdots P_{j_s})$. We further say that a polygon is a *nonzero polygon* if its sides $P_{j_1}P_{j_2}, P_{j_2}P_{j_3}, \cdots, P_{j_{s-1}}P_{j_s}, P_{j_s}P_{j_1}$ are all of unit length. The cubic character of L guarantees that a nonzero polygon must have an even number of sides. We include two-sided polygons (i.e., ones with only a pair of sides $P_{j_1}P_{j_2}, P_{j_2}P_{j_1}$) in our discussion; indeed, nonzero two-sided polygons play an important role, and we call them *degenerate polygons*.

Let π be any permutation of $Z = \{1, 2, \cdots, N\}$. This permutation can be written, in the usual way, as a product of disjoint cycles $\sigma_1^\pi \sigma_2^\pi \cdots \sigma_t^\pi$ (including 1-cycles if they occur). This product is unique apart from the order of its terms. A cycle in the product, say $\sigma = (j_1 j_2 \cdots j_s)$, corresponds naturally to a polygon $(P_{j_1}P_{j_2}\cdots P_{j_s})$; hence, there is a one-to-one correspondence between a permutation π and a partition of L into disjoint oriented polygons. However, each permutation π is in one-to-one correspondence with a product in the expansion of $\text{per}\,\mathbf{P} = \sum_\pi p_{1\pi(1)} p_{2\pi(2)} \cdots p_{N\pi(N)}$. Moreover, the nonzero products in this expansion correspond to the partitions of L into polygons which are all nonzero polygons.

Again, in any cycle $\sigma = (j_1 j_2 \cdots j_s)$ we can select a particular element i_1, say the numerically smallest element in σ, and then write $\sigma = (i_1 i_2 \cdots i_s)$, where $i_1 i_2 \cdots i_s$ is obtained from $j_1 j_2 \cdots j_s$ by cyclic permutation. Thus a cycle corresponds to a product in the sum (5); and a permutation $\pi = \sigma_1^\pi \sigma_2^\pi \cdots \sigma_t^\pi$ corresponds to a term (a product of N quaternions) in the sum obtained by substituting (5) into (6). This correspondence is one to one and again maps the nonzero terms in the expansion of $\det_M \mathbf{Q}$ onto the partitions of L into nonzero polygons.

Thus the number of nonzero products in the expansions of $\text{per}\,\mathbf{P}$ and $\det_M \mathbf{Q}$ is f^2 in both cases. Since each nonzero product in the expansion of $\det_M \mathbf{Q}$ is a product of N unit quaternions, such a product is a unit quaternion. It now follows from (8) that the modulus of a sum of f^2 unit quaternions can only be not less than $f^2 = \text{per}\,\mathbf{P}$ if all these unit quaternions are equal. Hence every nonzero term in the expansion of $\det_M \mathbf{Q}$ must equal $(-1)^{N/2}$, because this is the value of one such particular product obtained when all the polygons are degenerate, whereupon each quantity in (5) takes the form

$$-q_{i_1 i_2} q_{i_2 i_1} = -q_{i_1 i_2} \bar{q}_{i_1 i_2} = -1.$$

We say that a polygon on L is *admissible* if it is a nonzero polygon, and if there exists a polygon-partition of L, containing this polygon and having all its other polygons degenerate. Consider any admissible polygon with $2r$ sides, and suppose that

$q_1, q_2, \cdots, q_{2r}$ are the values of $q_{i_1 i_2}, \cdots, q_{i_{2r} i_1}$ encountered in following the cycle σ around this polygon. There is a polygon partition with $\frac{1}{2}N - r$ degenerate polygons besides the given admissible polygon. Hence the corresponding summand in (6) yields

$$(-q_1 q_2 \cdots q_{2r})(-1)^{\frac{1}{2}N-r} = (-1)^{\frac{1}{2}N}, \tag{11}$$

i.e.,

$$q_1 q_2 \cdots q_{2r} = (-1)^{r-1} \tag{12}$$

for any admissible polygon. In particular,

$$q_1 q_2 q_3 q_4 = -1 \tag{13}$$

for an admissible square; and

$$q_1 q_2 q_3 q_4 q_5 q_6 = +1 \tag{14}$$

for an admissible hexagon (not necessarily a planar hexagon).

Suppose temporarily that $a = 2$, $b = 4$, $c = 4$. We show that an admissible hexagon, whose opposite sides are opposite sides of a cube, lies near the center of L. This follows from the diagram in Fig. 1. Here the points of L are denoted by crosses or circles according to their x coordinate. It is easy to see from similar diagrams that any square, forming a face of the cube in the above diagram, is also admissible. Moreover,

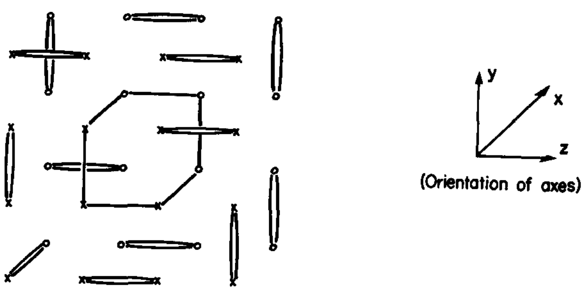

FIG. 1. Admissible hexagon in a smooth block.

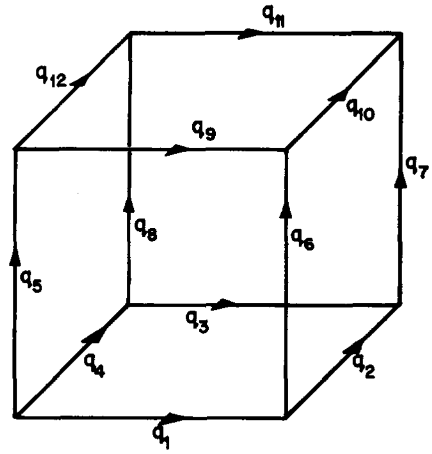

FIG. 2. Numbering and orientation of quaternions on a cube.

this diagram can be embedded in a larger polygon partition with $a \geq 2$, $b \geq 4$, $c \geq 4$ by pairing the points outside this $2 \times 4 \times 4$ configuration in an obvious fashion. So the existence of this admissible hexagon also follows for $a \geq 2$, $b \geq 4$, $c \geq 4$.

Now consider the cube carrying this admissible hexagon, and let the q_{ij} on its sides, with respect to the marked orientations, be $q_1, q_2, \cdots, q_{12}$ as shown in Fig. 2. From (14) we have

$$q_1 q_2 q_7 \bar{q}_{11} \bar{q}_{12} \bar{q}_5 = +1 \tag{15}$$

and from (13) we have

$$q_1 q_2 \bar{q}_3 \bar{q}_4 = q_3 q_7 \bar{q}_{11} \bar{q}_8 = q_8 \bar{q}_{12} \bar{q}_5 q_4 = -1. \tag{16}$$

Hence,

$$\begin{aligned} +1 &= q_1 q_2 q_7 \bar{q}_{11} \bar{q}_{12} \bar{q}_5 \\ &= q_1 q_2 (\bar{q}_3 q_3) q_7 \bar{q}_{11} (\bar{q}_8 q_8) \bar{q}_{12} \bar{q}_5 (q_4 \bar{q}_4) \\ &= q_1 q_2 \bar{q}_3 (q_3 q_7 \bar{q}_{11} \bar{q}_8)(q_8 \bar{q}_{12} \bar{q}_5 q_4) \bar{q}_4 \\ &= q_1 q_2 \bar{q}_3 \bar{q}_4 = -1. \end{aligned} \tag{17}$$

This contradiction denies (8) and completes the proof of (2).