MATH. SCAND. 12 (1963), 25–35

ON THE PERMANENT OF A BISTOCHASTIC MATRIX

HELGE TVERBERG

1.

For any (n,n)-matrix $A=\{a_{ij}\}$, the permanent, $\mathrm{per}(A)$, is defined by

$$\mathrm{per}(A) = \sum a_{1j_1}a_{2j_2}\ldots a_{nj_n}, \tag{1}$$

where the sum is extended over all permutations $(j_1,\ldots,j_n)$ of $(1,\ldots,n)$. A is said to be bistochastic if the conditions

$$a_{ij} \geqq 0, \quad \sum_{j=1}^{n} a_{ij} = 1, \quad \sum_{i=1}^{n} a_{ij} = 1, \qquad i,j = 1,\ldots,n, \tag{2}$$

are fulfilled.

Let Ω_n denote the set of all bistochastic matrices of order n. It is often identified with a set of points in n^2-dimensional space, by means of the correspondence

$$A \leftrightarrow (a_{11},a_{12},\ldots,a_{nn}).$$

As (2) implies $a_{ij} \leqq 1$, $i,j=1,\ldots,n$, it is seen that Ω_n is a closed and bounded subset of n^2-space, and there thus exists, by continuity, $A_{0,n} \in \Omega_n$ such that

$$\mathrm{per}(A_{0,n}) = P_n = \inf\{\mathrm{per}(A) \mid A \in \Omega_n\}.$$

König [1] proved that $A \in \Omega_n \Rightarrow \mathrm{per}(A) > 0$. Thus $P_n = \mathrm{per}(A_{0,n}) > 0$. Van der Waerden [4] raised the problem, which has not yet been solved, of determining P_n. Clearly $P_1 = 1!/1^1$; it is easy to see that $P_2 = 2!/2^2$; and it was proved by Marcus and Newman [2] that $P_3 = 3!/3^3$. The conjecture in [4] was that $P_n = n!/n^n = \mathrm{per}(Y_n)$, where $Y_n = \{n^{-1}\}$ is the matrix whose entries are all $1/n$. It was proved in [2] that $P_n \geqq (n^2-n+1)^{1-n}$ and it was also proved that if P_n is assumed by a matrix A having only positive elements, then $A = Y_n$. In [3], Marcus and Newman proved that if $A \in \Omega_n$ is symmetric and positive semi-definite, then $A \neq Y_n \Rightarrow \mathrm{per}(A) > n!/n^n$. The contents of the present paper are as follows.

We first transform the problem into a discrete one, concerning k-matrices. A k-matrix is an (n,n)-matrix, having non-negative integers as

Received June 22, 1962.

elements, for which the row- and column-sums all have the value kn, where k is an integer. In Section 2 we prove a theorem that asserts the possibility of dividing any k-matrix in k almost equal 1-matrices, the differences being only those which result from the fact that not all integers are divisible by k. By means of this theorem we formulate in Section 3 a conjecture which is a little stronger than the one given in [4]. This conjecture has the interesting feature that there exists a very explicit decision procedure for it. In Section 4 we describe a combinatorial approach to the discrete version of our problem, stressing the interpretation of sums of products of non-negative integers as cardinal numbers of unions of product-sets. In Section 5 we introduce the function $\operatorname{per}_l(\boldsymbol{A})$ which is, by definition, the sum of the l'th order subpermanents of $\boldsymbol{A}$. Van der Waerdens conjecture on $\operatorname{per}_n(\boldsymbol{A})$ extends in a natural way to $\operatorname{per}_l(\boldsymbol{A})$, and so does the combinatorial approach to his conjecture. The purpose of Section 5 is to try out this approach on the case $l=2$, in order to learn something about what ought to be done when $l>2$. In Section 6 the result obtained in Section 5, and the corresponding result for $l=3$, are proved in another way.

The reader may be interested in having a proof of König's result, as this is the real basis of the problem treated here. Such a proof comes out as a by-product of our proof of the theorem on k-matrices.

2.

If $\boldsymbol{A} \in \Omega_n$ is given, we can, by making a small "correction", obtain $\boldsymbol{A}' \in \Omega_n$, such that $\boldsymbol{A}'$ has only rational elements and is as close to $\boldsymbol{A}$ as we wish. Assume that $a_{nn} > 0$. We correct the irrational a_{ij} for which $i<n$, $j<n$ by subtracting positive numbers. The elements on the right and lower edge (except a_{nn}) must then be corrected by adding non-negative numbers so as to get the correct value of the row- and column-sums, and by making the first corrections small enough, a'_{nn} will be >0 and $\boldsymbol{A}'$ will be as close to $\boldsymbol{A}$ as we wish. By continuity we find that it is sufficient to consider rational matrices if we want to prove that $\operatorname{per}(\boldsymbol{A}) \geqq \operatorname{per}(\boldsymbol{Y}_n)$ for $\boldsymbol{A} \in \Omega_n$. As $\operatorname{per}(\boldsymbol{A})$ is a homogeneous polynomial in the a_{ij}, this inequality is equivalent to $\operatorname{per}(kn\boldsymbol{A}) \geqq \operatorname{per}(kn\boldsymbol{Y}_n)$ for any positive integer k. If $\boldsymbol{A}$ is rational, we let k be a common denominator of the fractions occuring, and we thus see that it is sufficient to consider k-matrices.

Let us call a 1-matrix $\boldsymbol{A}^{(1)}$ an admissible component of a k-matrix $\boldsymbol{A}$ if the conditions

$$(3) \qquad [k^{-1}a_{ij}] \leqq a_{ij}^{(1)} < k^{-1}a_{ij}+1, \qquad i,j = 1,\ldots,n ,$$

are fulfilled. If $a_{ij} \equiv 0 \pmod{k}$, the only possible value for $a_{ij}^{(1)}$ is then $k^{-1}a_{ij}$, while otherwise there are two possible values. We shall make some use of the following

THEOREM. *A k-matrix can always be written as a sum of k admissible components.*

It is convenient to prove first the

LEMMA. *A k-matrix has at least one admissible component.*

PROOF. If all elements of A are $\equiv 0 \pmod{k}$, the matrix $\{k^{-1}a_{ij}\}$ is the only admissible component of A. Let $N(A)$ denote the number of elements of A which are $\not\equiv 0 \pmod{k}$. Then the case $N(A)=0$ is just covered, and we can proceed by induction on $N(A)$. If $N(A)>0$, we can find a sequence of elements $a_{ij} \not\equiv 0 \pmod{k}$, the corresponding sequence of pairs of indices being as follows:

$$(i_1,j_1),\ (i_1,j_2),\ (i_2,j_2),\ (i_2,j_3),\ (i_3,j_3),\ldots .$$

As no row, or column, of A can have only one element $\not\equiv 0 \pmod{k}$, we can assume that $i_{m+1} \neq i_m$, $j_{m+1} \neq j_m$. We conclude that for some $s<t$, $u<v$, $i_s=i_t$ and $j_u=j_v$. Choosing the differences $t-s$ and $v-u$ as small as possible and assuming that $t-s \leqq v-u$, $s=1$ (these restrictions are inessential), we conclude further that in the sequence

$$a_{i_1j_2}, a_{i_2j_2}, a_{i_2j_3}, \ldots, a_{i_tj_t}, \tag{4}$$

no element occurs twice. Let r be the least and R the greatest of the residues $\bmod k$ of the elements in (4), and let $\varrho=\inf(r,k-R)$. Then $0<\varrho<k$. If $\varrho=r$ we choose an element from (4) which is $\equiv \varrho \pmod{k}$, subtract ϱ from that element, add ϱ to its neighbours in (4), subtract ϱ from its second neighbours etc. If $\varrho \neq r$ we choose an element $\equiv -\varrho \pmod{k}$, add ϱ to that element, subtract ϱ from its neighbours in (4) etc. In both cases, leaving unchanged the elements which do not belong to (4), we arrive at a new k-matrix A'. At least one element of A' is $\equiv 0 \pmod{k}$ while the corresponding element of A is $\not\equiv 0 \pmod{k}$, and, furthermore, $a_{ij}' \not\equiv 0 \pmod{k} \Rightarrow a_{ij} \not\equiv 0 \,(\mathrm{mid}\, k)$. Thus $N(A')<N(A)$ and by the hypothesis of induction, A' has an admissible component $A^{(1)}$. If $a_{ij} \equiv 0 \pmod{k}$, then $a_{ij}'=a_{ij}$, and $a_{ij}^{(1)}=k^{-1}a_{ij}'=k^{-1}a_{ij}$. If $a_{ij}=ak+b$, $0<b<k$, and $a_{ij}' \neq a_{ij}$, then

$$ak = a_{ij}-b \leqq a_{ij}-\varrho \leqq a_{ij}' \leqq a_{ij}+\varrho \leqq a_{ij}+k-R \leqq a_{ij}+k-b = (a+1)k\,,$$

and accordingly $a \leqq a_{ij}^{(1)} \leqq a+1$. Thus $A^{(1)}$ is an admissible component of A.

The theorem, which is obviously true for $k=1$, is now proved by induction on k. Let $k>1$, and let $A^{(1)}$ be an admissible component of A. Put $A''=A-A^{(1)}$. We find, using (3), that the elements of A'' are non-negative, and hence A'' is a $(k-1)$-matrix which, by the hypothesis of induction, can be written as a sum $A^{(2)}+\ldots+A^{(k)}$ of admissible components of A''. We must finally check that $A^{(2)}$, say, is also an admissible component of A. Let $a_{ij}=ak+b$, $0 \leqq b < k$. If $b>0$, then $a \leqq a_{ij}^{(1)} \leqq a+1$ and

$$\begin{aligned} a(k-1) &\leqq a(k-1)+b-1 \\ &= ak+b-(a+1) \leqq a_{ij}'' \leqq ak+b-a \\ &= a(k-1)+b \leqq (a+1)(k-1)\,. \end{aligned}$$

Thus $a \leqq a_{ij}^{(2)} \leqq a+1$. If $b=0$, then $a_{ij}^{(1)}=a$, $a_{ij}''=(k-1)a$ and $a_{ij}^{(2)}=a$. Thus $A^{(2)}$ is really an admissible component of A, and the proof of the theorem is complete.

Now, let A be a bistochastic matrix. We see that no row (or column) of A can have exactly one element which lies strictly between 0 and 1. Thus, if A has got at least one element different from 0 and 1, we can repeat the reasoning that led to the sequence (4), and obtain a sequence like (4), having only positive elements. If ϱ is the least one of these elements, we can subtract and add ϱ, alternatingly, and obtain a new bistochastic matrix A'. A' has all the zeros of A, and in addition at least one more $(=\varrho-\varrho)$. If $\operatorname{per}(A)=0$, then $\operatorname{per}(A')=0$. If A' has at least one element different from 0 and 1, we can repeat the procedure, and obtain A''. If $\operatorname{per}(A)=0$, then $\operatorname{per}(A'')=0$. Continuing like that, we must finally arrive at a bistochastic matrix with n^2-n zeros and n ones, a so-called permutation matrix. But the permanent of such a matrix is 1. The assumption $\operatorname{per}(A)=0$ thus leads to a contradiction and we have proved König's result, as promised in the introduction.

In view of the close connection between König's and our own theorem (referred to as I and II respectively), that has just been revealed, it is tempting to try do deduce II from I. In fact, by I, we can conclude that, given a k-matrix, there exists a permutation matrix P_1, such that $A-P_1 \geqq 0$ in all places. As $A-P_1$ is $(kn-1)$ times a bistochastic matrix, I shows again the existence of a permutation matrix P_2 such that $A-P_1-P_2 \geqq 0$ in all places. Continuing like that we get $A=P_1+\ldots+P_{kn}$. But I can not see how the P's can be put together n by n so as to give k admissible components of A.

3.

Let P and Q be polynomials in $x_1, x_2, \ldots, x_k$. Then $P \geqq Q$ is sometimes defined as meaning that each coefficient of P is $\geqq$ the corresponding coefficient of Q. Let $\boldsymbol{A} = \boldsymbol{A}^{(1)} + \boldsymbol{A}^{(2)} + \ldots + \boldsymbol{A}^{(k)}$ be an admissible decomposition of the k-matrix $\boldsymbol{A}$. Letting $\geqq$ have the meaning described above, we make the following

CONJECTURE. $\operatorname{per}(x_1\boldsymbol{A}^{(1)} + \ldots + x_k\boldsymbol{A}^{(k)}) \geqq \operatorname{per}\big((x_1 + \ldots + x_k)n\boldsymbol{Y}_n\big)$.

As before, $\boldsymbol{Y}_n = \{n^{-1}\}$. It is seen that this conjecture is stronger than that of van der Waerden; put $x_1 = \ldots = x_k = 1$. The conjecture may be found a little unmotivated at this point. It is based on the considerations of the next two sections, but it is placed here, as it is only dependent on our theorem for its formulation.

Let n be given and assume that the conjecture has been found to hold for $k \leqq n$. Let $k > n$, and let us consider the two coefficients of a monomial M.

As the polynomials in question are of degree n, not more than n variables can occur in M, and it is no restriction to assume that $M = x_1^{i_1} \ldots x_n^{i_n}$. The two coefficients of M are not affected if we put $x_{n+1} = \ldots = x_k = 0$, and so we can look at them in the polynomials $\operatorname{per}(x_1\boldsymbol{A}^{(1)} + \ldots + x_n\boldsymbol{A}_n^{(n)})$ and $\operatorname{per}\big((x_1 + x_2 + \ldots + x_n)n\boldsymbol{Y}_n\big)$. But these polynomials are those of the conjecture, in the case $k = n$, $\boldsymbol{A} = \boldsymbol{A}^{(1)} + \ldots + \boldsymbol{A}^{(n)}$ (note that $\boldsymbol{A}^{(1)} + \ldots + \boldsymbol{A}^{(n)}$ is an admissible decomposition of $\boldsymbol{A}^{(1)} + \ldots + \boldsymbol{A}^{(n)}$. By our assumption, the coefficient of M to the left is $\geqq$ the one to the right, and we are through.

Thus the conjecture needs only to be verified for $k \leqq n$. It would be easy to give an explicit bound for the necessary amount of calculations, but we shall not do so, as our observation is probably only of theoretical interest.

4.

Let S be a collection of kn^2 objects, let $\boldsymbol{A} = \{a_{ij}\}$ be a k-matrix and let $D(S, \boldsymbol{A})$ be a distribution of S in n^2 boxes B_{ij} such that the box B_{ij} contains a_{ij} objects. Consider a subset σ of S, containing n objects, belonging to $B_{i_1j_1}, \ldots, B_{i_nj_n}$. If $(i_1, \ldots, i_n)$ and $(j_1, \ldots, j_n)$ are permutations of $(1, \ldots, n)$, we say that σ is a proper n-tuple. Let $C_n\big(D(S, \boldsymbol{A})\big)$ denote the set whose elements are the proper n-tuples. As one gets $a_{1i_1}a_{2i_2} \ldots a_{ni_n}$ proper n-tuples by picking one object from each of the boxes $B_{1i_1}, \ldots, B_{ni_n}$ (provided $(i_1, \ldots, i_n)$ is a permutation of $(1, \ldots, n)$), it is seen that the cardinal number of $C_n\big(D(S, \boldsymbol{A})\big)$, $\big|C_n\big(D(S, \boldsymbol{A})\big)\big|$, is given by $\operatorname{per}(\boldsymbol{A})$. We want to prove that $\operatorname{per}(kn\boldsymbol{Y}_n) \leqq \operatorname{per}(\boldsymbol{A})$, that is, $|C_n(D'(S',kn\boldsymbol{Y}_n))| \leqq |C_n(D''(S'',\boldsymbol{A}))|$, when D' and D'' are any distributions associated with $kn\boldsymbol{Y}_n$ and $\boldsymbol{A}$. D' and D'' can be chosen in many ways, and we may hope that by making the proper choice, we can prove the inequality by constructing an injective function

$$\varphi: \quad C_n(D'(S',kn\boldsymbol{Y}_n)) \to C_n(D''(S'',\boldsymbol{A}))$$

(injective means that $x \neq y \Rightarrow \varphi(x) \neq \varphi(y)$). It is natural to let $S' = S'' = S$, as both of these sets are subjected to the sole condition of having kn^2 elements, and we shall, in what follows, describe a way of making D'' dependent on D'.

Let us start with the case $k=1$. We let S_0 consist of objects α_{ij}, $i,j=1,\ldots,n$, and $D(S_0,n\boldsymbol{Y}_n)$ is obtained by placing the object α_{ij} in B_{ij}. If $\boldsymbol{A}$ is a 1-matrix, and $a_{ij}=0$, we move α_{ij} from B_{ij} to some B_{ij_1}, where j_1 is chosen such that $a_{ij_1} > 1$. As $\sum_{j=1}^{n} a_{ij} = n$, we can arrange these moves so that if $a_{ij} > 1$, exactly $a_{ij}-1$ objects are moved into B_{ij}. Thus, after having done the moves, we have arrived at a distribution $D(S_0,\boldsymbol{A})$. With it, we associate a certain permutation f: $S_0 \to S_0$, having the following properties: If α is placed in the same box under $D(S_0,\boldsymbol{A})$ as under $D(S_0,n\boldsymbol{Y}_n)$, $f(\alpha)=\alpha$. If α is moved from B_{ij} to B_{ij_1}, $f(\alpha)$ is moved from $B_{i_1j_1}$ to $B_{i_1j_2}$ for some i_1,j_2. It is convenient in the last case to think of $f(\alpha)$ as the object which must be removed from the j_1'th column, in order that α may be moved into that column. The possibility of constructing an f as described is granted by the fact that the column sums in $\boldsymbol{A}$ are the same as those in $n\boldsymbol{Y}_n$ (namely n).

We now assume that for each 1-matrix $\boldsymbol{A}$ a definite distribution $D(S_0,\boldsymbol{A})$ and a corresponding permutation f: $S_0 \to S_0$ have been constructed as described above. Choose for each k-matrix $\boldsymbol{A}$ a definite representation $\boldsymbol{A}^{(1)} + \ldots + \boldsymbol{A}^{(k)}$ by admissible components. Make k copies $S_1,\ldots,S_k$ of the set S_0 by putting $S_s = \{\alpha_{ij}{}^s\}$. It is convenient to think of the superindices as "colours" put on the objects α_{ij}. The distribution $D(S_1 \cup \ldots \cup S_k, kn\boldsymbol{Y}_n) = D(S,kn\boldsymbol{Y}_n)$ is now obtained by placing $\alpha_{ij}{}^s$ in B_{ij}. If the k-matrix $\boldsymbol{A}$ is given, and $\boldsymbol{A}^{(1)} + \ldots + \boldsymbol{A}^{(k)}$ is the chosen representation of $\boldsymbol{A}$, we define $D(S,\boldsymbol{A})$ by the rule: If α is placed in B_{ij} by $D(S_0,\boldsymbol{A}^{(s)})$, then α^s is placed in B_{ij} by $D(S,\boldsymbol{A})$. A permutation f: $S \to S$ is defined by: $f(\alpha^s) = (f_s(\alpha))^s$. Here f_s denotes the permutation on S_0 which is associated with $D(S_0,\boldsymbol{A}^{(s)})$.

In the next section we shall need the following properties of f: $S \to S$: The objects x and $f(x)$ always have the same colour. The equation $f(x)=x$ holds if and only if x is not moved by the construction of $D(S,\boldsymbol{A})$ from $D(S,kn\boldsymbol{Y}_n)$. We shall also need a certain property of this construction, which is most simply stated as: A box does not at the same time receive *and* give away objects. This is a consequence of (3). If an object $\alpha_{ij}{}^s$ is moved from B_{ij} to B_{ij_1} under the transition from $D(S,kn\boldsymbol{Y}_n)$ to $D(S,\boldsymbol{A})$, then $a_{ij}{}^{(s)}=0$. But then, by (3), $a_{ij}<k$, and $a_{ij}{}^{(t)}\leqq 1$ for $t=1,\ldots,k$. But this again means that no object is moved to B_{ij} under any of the k transitions $D(S_0,n\boldsymbol{Y}_n)\to D(S_0,\boldsymbol{A}^{(t)})$, and thus no object is moved to B_{ij} under the transition $D(S,kn\boldsymbol{Y}_n)\to D(S,\boldsymbol{A})$.

As we have now constructed everything we need, we shall simplify the notation in what follows. Thus α is no longer an object in the auxiliary set S_0, but an element of $S=S_1\cup\ldots\cup S_k$ and $C_n(D(S,\boldsymbol{A}))$ is replaced by $C_n(\boldsymbol{A})$. We shall also make use of two names of α, namely α and α' according to whether we have $D(kn\boldsymbol{Y}_n)$ or $D(\boldsymbol{A})$ in mind. Thus "α' belongs to the fifth column" means "By $D(\boldsymbol{A})$, α is placed in a box B_{i5}". The meaning of "α' and α are always in the same row, while α' and $f(\alpha)$ are always in the same column" should then be clear. These two facts are also among the ones that are important to keep in mind for the next section.

Consider now $(\alpha,\beta,\ldots)\in C_n(kn\boldsymbol{Y}_n)$. If $(\alpha',\beta',\ldots)\in C_n(\boldsymbol{A})$, it feels natural to put $\varphi(\alpha,\beta,\ldots)=(\alpha',\beta',\ldots)$. But what is to be done if, say, α' and β' are in the same column, and thus $(\alpha',\beta',\ldots)\notin C_n(\boldsymbol{A})$? Then the permutation f seems to be our hope of rescue, because it points to other candidates for $\varphi(\alpha,\beta,\ldots)$. The next section is devoted to trying out this point of view on a simple analogue of our problem.

5.

If $\boldsymbol{A}\in\Omega_n$ and $1\leqq l\leqq n$, then $\mathrm{per}_l(\boldsymbol{A})$, the sum of the l'th order subpermanents of $\boldsymbol{A}$, is positive. In fact we know that, say, $a_{11}a_{22}\ldots a_{nn}>0$, and this implies that at least $\binom{n}{l}$ of the l'th order subpermanents of $\boldsymbol{A}$ are >0. In the previous sections we have discussed $\mathrm{per}(\boldsymbol{A})=\mathrm{per}_n(\boldsymbol{A})$, but we might just as well have treated $\mathrm{per}_l(\boldsymbol{A})$ for some fixed value of l. Especially we are led to the definition of $C_l(\boldsymbol{A})$ when $\boldsymbol{A}$ is a k-matrix, and to the question of finding an injective mapping $\varphi\colon C_l(kn\boldsymbol{Y}_n)\to C_l(\boldsymbol{A})$. (Each element of $C_l(\boldsymbol{A})$ is a subset of S consisting of l objects, no two of which are placed in the same row or in the same column by $D(\boldsymbol{A})$). We shall now treat the case $l=2$, and we start by recommending the reader to make his own illustrations while following the essentially easy argument.

Let $(\alpha,\beta)\in C_2(kn\boldsymbol{Y}_n)$ and consider the sequence

$$(5)\qquad (\alpha,\beta),\ (\alpha',\beta'),\ (f(\alpha),f(\beta)),\ (f(\alpha)',f(\beta)'),\ (f^2(\alpha),f^2(\beta)),\ldots$$

Consider the statements "$f^r(\alpha)'$ and $f^r(\beta)'$ are in different columns" and "$f^r(\alpha)$ and $f^r(\beta)$ are in the same row", each of them concerning half of the couples occurring in (5). The first one is true for $r=uv-1$, where $u(v)$ is the length of the cycle to which $\alpha(\beta)$ belongs in the permutation f. This is so because $f^{uv-1}(\alpha)'$ $(f^{uv-1}(\beta)')$ belongs to the same column as $f^{uv}(\alpha)$ $(f^{uv}(\beta))$, but $f^{uv}(\alpha)=\alpha$ $(f^{uv}(\beta)=\beta)$, and α and β are in different columns, as $(\alpha,\beta)\in C_2(kn\mathbf{Y}_n)$. Thus there is a first couple in (5) for which the appropriate statement is true.

Consider first the case when this "first couple" is of the type $(f^r(\alpha)', f^r(\beta)')$, $r\geqq 0$. Then $f^r(\alpha)$ and $f^r(\beta)$ are in different rows, as $(f^r(\alpha), f^r(\beta))$ precedes $(f^r(\alpha)', f^r(\beta)')$ in (5). This implies that $f^r(\alpha)'$ and $f^r(\beta)'$ are in different rows, and since they are also in different columns, we can put $\varphi(\alpha,\beta)=(f^r(\alpha)', f^r(\beta)')$.

If we look at an image element (γ',δ') obtained by this partial definition, we can find $(\alpha,\beta)=\varphi^{-1}(\gamma',\delta')$ by going backwards in (5), determining the least $t\geqq 0$ for which $f^{-t}(\gamma)$ and $f^{-t}(\delta)$ are in different columns. Thus φ is injective so far. We note that $f^{-1}(\gamma)'$ and $f^{-1}(\delta)'$ are never in the same box, while, as we shall see shortly, for the image elements (γ',δ') obtained in the case when the "first couple" is of the type $(f^r(\alpha), f^r(\beta))$, it will be true that $f^{-1}(\gamma)'$ and $f^{-1}(\delta)'$ are always in the same box.

This second case is conveniently divided in two subcases. The first one occurs when, say, $f(\alpha)=\alpha$, but $f(\beta)\neq\beta$. Then (5) is (α,β), (α',β'), $(\alpha, f(\beta))$, $(\alpha', f(\beta)'),\ldots$, and, as (α',β') is supposed not to be the "first couple", α' (and α) and β' are in the same column. This column is then different from the one to which $f(\beta)'$ belongs (as $f(\beta)\neq\beta$), and accordingly α' and $f(\beta)'$ are in different columns. As $(\alpha', f(\beta)')$ is supposed not to be the "first couple", $(\alpha, f(\beta))$ must be it, and so α and $f(\beta)$ are in the same row (and in the same box). An object ε is now defined by: ε belongs to the same box as β' and has the same colour as α. We put $\varphi(\alpha,\beta)=(\varepsilon', f(\beta)')$. As β and $f(\beta)$ are in different rows, ε' and $f(\beta)'$ will also be. As to the columns, it is to be noted that $f(\varepsilon)=\varepsilon$ because β has moved into that box to which ε belongs. Thus $f(\beta)$ moves out from the same column in which ε remains; thus ε' and $f(\beta)'$ are in different columns.

If (γ',δ') is arrived at by the definition just given, then $f^{-1}(\gamma)'$ and $f^{-1}(\delta)'$ are in the same box. One of the objects, say δ, is moving, while the other is not. If $\varphi(\alpha,\beta)=(\gamma',\delta')$, then $\beta=f^{-1}(\delta)$ while α is the object which belongs to the same box as δ and has the same colour as γ (we have denoted by β the moving object in the couple $\varphi^{-1}(\gamma',\delta')$). Thus φ is still injective.

The second subcase is when $f(\alpha)\neq\alpha$, $f(\beta)\neq\beta$, (now all possibilities are exhausted, as $f(\alpha)=\alpha$, $f(\beta)=\beta$ implies that (α',β') is the "first couple").

Then $f^{r-1}(\alpha)'$ and $f^{r-1}(\beta)'$ are in the same column, but in different rows. We choose an arbitrary one of the two objects α and β, say α, and put $\varphi(\alpha,\beta)=\big(f^r(\alpha)',\varepsilon'\big)$, where ε is the object which is in the same box as $f^{r-1}(\alpha)'$ and has the same colour as β. It is clear that $f^r(\alpha)'$ and ε' are in different rows, and they are also in different columns by the same reasoning as that given in the preceding case. The reconstruction of the "first couple" is also made in a similar way as in that case, and as the "first couple" in (5) determines (α,β) uniquely, we are through with the verification of the injectivity of φ as soon as we note that also in this second subcase $f^{-1}\big(f^r(\alpha)\big)'$ and $f^{-1}(\varepsilon)'$ are in the same box.

Thus we have obtained what we wanted; the combinatorial approach works nicely in the case $l=2$. The construction of φ may seem to be a complicated process, but then it must be taken into account that our problem probably is a difficult one. (This belief is based on statements by Erdös and van der Waerden, on the consideration of Marcus' and Newman's work on the problem and on the author's own experience.)

It is interesting to notice that if (α,β) consists of, say, one blue and one green object, then we never looked at anything but blue and green objects, and, especially, $\varphi(\alpha,\beta)$ consists of one blue and one green object. If $A^{(1)}+\ldots+A^{(k)}$ is the decomposition of A used for constructing $D(A)$, this observation can be expressed by the inequality

$$\mathrm{per}_2\big((x_1+\ldots+x_k)nY_n\big) \leqq \mathrm{per}_2(x_1A^{(1)}+\ldots+x_kA^{(k)})$$

in the sense of Section 3. The coefficient of x_ix_j to the right is the same as the number of couples in $C_2(A)$ consisting of one object of colour no. i and one of colour no. j. A similar interpretation is valid for the coefficients to the left, and for the coefficients of x_i^2 etc. The conjecture given in Section 3 is an extrapolation from the inequality above, based on the hope that the considerations in this section will turn out to be the clue to the solution of our problem also when $l>2$.

6.

We shall now prove

$$(6) \qquad A \in \Omega_n \;\&\; A \neq Y_n \;\&\; 1<l<4 \;\Rightarrow\; \mathrm{per}_l(A) > \mathrm{per}_l(Y_n)\,.$$

By definition, we have

$$(7) \qquad 3!\,\mathrm{per}_3(A) = \sum a_{pq}a_{rs}a_{tu}, \qquad p \neq r \neq t \neq p,\; q \neq s \neq u \neq q\,.$$

Keeping p, q, r, s fixed and summing the last factor over t and u, we get

$$\sum a_{tu} = \sum (1-a_{tq}-a_{ts}) = n-2-(1-a_{pq}-a_{rq})-(1-a_{ps}-a_{rs})$$
$$= n-4+a_{pq}+a_{rq}+a_{ps}+a_{rs}\,.$$

Thus (7) splits into five sums, of which the first one is $2(n-4)\,\mathrm{per}_2(A)$. We have

$$2\,\mathrm{per}_2(A) = \sum_{p\neq r,\, q\neq s} a_{pq}a_{rs} = \sum_{p\neq r} a_{pq}(1-a_{rq})$$
$$= \sum a_{pq}(n-1-(1-a_{pq}))$$
$$= (n-2)\sum a_{pq}+\sum a_{pq}^2 = (n-2)n+\sigma_2\,.$$

The second sum is

$$\sum_{p\neq r,\, q\neq s} a_{pq}^2\, a_{rs} = \sum_{p\neq r} a_{pq}^2(1-a_{rq})$$
$$= \sum a_{pq}^2(n-2+a_{pq}) = (n-2)\sigma_2+\sigma_3\,.$$

The third one is

$$\sum_{p\neq r,\, q\neq s} a_{pq}a_{rs}a_{rq} = \sum_{p\neq r} a_{pq}a_{rq}(1-a_{rq})$$
$$= \sum (1-a_{rq})a_{rq}(1-a_{rq}) = n-2\sigma_2+\sigma_3\,.$$

By symmetry, the fourth (fifth) sum equals the third (second) one. Thus

$$(8)\qquad 2\,\mathrm{per}_2(A) = (n-2)n+\sigma_2\,,$$

$$(9)\qquad 6\,\mathrm{per}_3(A) = n^3-6n^2+10n+3(n-4)\sigma_2+4\sigma_3\,.$$

Now if $r>1$, the sum $x_1^r+x_2^r+\ldots+x_n^r$ has a single minimum in the domain $x_1+x_2+\ldots+x_n=1$, $(x_1,x_2,\ldots,x_n)\geqq(0,0,\ldots,0)$. The minimum is assumed as the domain is closed and bounded, and it is assumed for $x_1=x_2=\ldots=x_n=1/n$. This is an immediate consequence of the fact that

$$0\leqq x\leqq c \ \&\ x\neq \tfrac{1}{2}c \ \Rightarrow\ x^r+(c-x)^r > 2(\tfrac{1}{2}c)^r\,.$$

We then see that (8) and (9) imply (6) if $n\geqq 4$, because σ_2 and σ_3 each is a sum of n sums of the type just treated. If $n=3$, it suffices to remark that one finds, using Lagrange multipliers, that $4(x_1^3+x_2^3+x_3^2)-3(x_1^2+x_2^2+x_3^2)$ has its only minimum in the domain $x_1+x_2+x_3=1$, $(x_1,x_2,x_3)\geqq(0,0,0)$, for $x_1=x_2=x_3=\frac{1}{3}$. Thus we have verified van der Waerden's conjecture in the case $n=3$, by a method completely different from the one used by Marcus and Newman in [2].

One can find formulae similar to (8) and (9) for $\mathrm{per}_l(A)$, $l>3$, but the simplicity seems to disappear, mainly because other terms than pure power sums occur.

REFERENCES

1. D. König, *Über Graphen und ihre Anwendung auf Determinantentheorie und Mengenlehre*, Math. Ann. 77 (1916), 453–465.
2. M. Marcus and M. Newman, *On the minimum of the permanent of a doubly stochastic matrix*, Duke Math. J. 26 (1959), 61–72.
3. M. Marcus and M. Newman, *Inequalities for the permanent function*, Ann. of Math. 75 (1962), 47–62.
4. B. L. van der Waerden, *Aufgabe* 45, Jber. Deutsch. Math. Verein. 35 (1926), 117.

UNIVERSITY OF BERGEN, NORWAY