COMBINATORICS AND DIAGONALS OF MATRICES

A THESIS

Submitted for the award of the Degree

of

DOCTOR OF PHILOSOPHY

IN

STATISTICS

by

K. BALASUBRAMANIAN

Department of Statistics

Loyola College, Madras

To

MATH / STAT DIVISION

INDIAN STATISTICAL INSTITUTE

CALCUTTA-700 035

DECEMBER 1980

COMBINATORICS AND DIAGONALS OF MATRICES

A THESIS
Submitted for the award of the Degree
of
DOCTOR OF PHILOSOPHY
IN
STATISTICS

by

K. BALASUBRAMANIAN
Department of Statistics
Loyola College, Madras

To
MATH / STAT DIVISION
INDIAN STATISTICAL INSTITUTE
CALCUTTA-700 035
DECEMBER 1980

CERTIFICATE

This is to certify that the thesis entitled 'COMBINATORICS AND DIAGONALS OF MATRICES' which is being submitted by Shri K. BALASUBRAMANIAN to the Indian Statistical Institute, Calcutta, for the award of the Degree of Doctor of Philosophy, is a record of bonafide research work carried out by him under my supervision and guidance. The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Dr. K. R. PARTHASARATHY
Professor and Head,
Department of Mathematics,
Indian Institute of Technology,
Madras-600036.

MADRAS 600 036

DECEMBER, 1980.

hint seemed to have been practically ignored. But when formalized as an algebra, Muir algebra seems to be a very versatile tool to deal with many aspects of combinatorics. With the help of this algebra all the existing formulas for the evaluation of permanents are derived and even better formulas are derived. Moreover a general procedure, by which any number of such formulas can be derived, is given. An interesting result on permanent of integral matrices is derived. Then all the formulas of inclusion-exclusion are derived through Muir algebra. Finally the generating function for the partition function is got using Muir algebra. It can be safely predicted that Muir algebra will turn out to be a very important tool in combinatorics.

Chapter four is the direct result of the attempts to resolve vander Waerden conjecture on doubly stochastic matrices. A new function $h_r(A)$ is introduced and all the results are expressed interms of this function in an elegant manner. Some inequalities that look like van der Waerden's are proved. Tverberg's conjecture and Djokovic conjecture (though not resolved) are analysed. On doubly stochastic matrices the problem posed by Friedland and Minc is partially solved. Some inequalities on t permanents are got through the use of multinomial distribution and covariance. Then the reasons for the failure of Ryser's conjecture (through Junkat's counter example) are analysed and this results in the definition of column dominating matrices and row dominating

C O N T E N T S

PAGES

PREFACE

This thesis mainly deals with combinatorial aspects of diagonals of matrices. Of course, there are also results which are not combinatorial in nature; but these are merely by-products. Chapter-0 gives a very short summary of the contents of the thesis. The results are of two kinds. (1) completely new and (2) old results through new methods.

This thesis, wholly or partly, has not been submitted to any other University or Institute for a degree.

I express hereby my deepest sense of gratitude to Dr. K.R. PARTHASARATHY, Head of the Department of Mathematics, Indian Institute of Technology, Madras, under whose benign guidance this work was completed, for his encouragement and all possible help.

Dr. K.N. VENKATARAMAN, Head of the Department of Statistics, University of Madras, has been a constant source of encouragement, showing a personal interest in me. I thank him wholeheartedly.

Rev. Fr. KURIAKOSE, Principal, Loyola College, Madras had an affectionate and fatherly interest in my research work. I take this opportunity to thank him profusely for his kindness and genuine interest.

Lastly I thank Mr. B. MAHALINGAM, Department of Statistics, University of Madras, for his excellent typing and Mr. M. BALASUBRAMANIAM Department of Statistics, University of Madras, for his able assistance in cyclostyling the thesis.

DEPARTMENT OF STATISTICS
LOYOLA COLLEGE
MADRAS 600 034.

K. BALASUBRAMANIAN

-o0o-

*****

matrices. Thus a new class of doubly stochastic matrices, on which van der Waerden conjecture is true, results. Structure of this class is analysed.

The fifth and last chapter deals with the use of permanent function to get a complete invariant polynomial for graphs. Some new results on graphs emerge. Though a complete invariant polynomial is only suggested (and not proved) it has been verified through the use of the computer in I.I.T., Madras that the polynomial is indeed a complete invariant for graphs with atmost 7 vertices. This polynomial is modified and two more equivalent polynomials which are better behaved are suggested to be complete invariant polynomials. If this conjecture proves to be correct, then, there is every possibility that Ulam's conjecture will also be settled in the same manner as there are close links. Thus a detailed study of this polynomial is likely to be of immense use in graph theory.

It is appropriate here to mention that whenever a name without a reference or a result without a reference is given in the thesis the reference is ALWAYS to 'Permanents - Henryk Minc Vol.6 of Encyclopedia of Mathematics and its Applications - Addison Wesley (1978).

COMBINATORICS AND DIAGONALS OF MATRICES

C H A P T E R - 0

This thesis mainly deals with combinatorial aspects of diagonals of matrices. While it is true that certain quantitative aspects of matrices that cannot be properly brought under combinatorics (which may profitably go under linear algebra) are certainly dealt with in the thesis, these should be taken more as a detour or unavoidable accompaniments, the main thread being only combinatorics. Perhaps, as under combinatorics, certain quantitative techniques cannot be avoided just as certain parts of member theory require complex variable techniques, the by-products cannot be avoided.

The entire work arose out of an attempt to solve some combinatorial problems and Vander Waerden conjecture on doubly stochastic matrices. While the basic questions still remain unresolved the attempt itself produced highly fruitful results.

The first chapter deals with diagonals of matrices. This is an important chapter as it resolves affirmatively two conjectures of Tzu Hsia Wang and disproves a third by means of counter examples. In fact the first two conjectures are considerably extended and several proofs are given. A result in assignment problem in operations Research comes as a by-product in disproving the third conjecture.

The first chapter also introduces the concept of disjoint diagonals and deals with latinsquares and orthogonal latin squares in a unified manner with the help of this concept. A conjecture is proposed on the form of orthogonal latin squares. The same concept also tackles projective planes and application of difference sets to projective planes.

The second chapter is mainly on diagonal products in matrices. Naturally determinants and permanents find due place here. Permanents and determinants with respect to group complexes are introduced and in this setting the usual properties acquire new dimensions. Then one of the most productive of concepts, namely, the concept of inner products in tensor products to define permanents, introduced by Marvin Marcus and Morris Newman is used and generalizations of many results and some new results on permanents are got.

Then, in the same chapter, an important conjecture of Marvin Marcus on the rank of positive matrices is affirmatively settled with help of some beautiful lemmas. But what strike more are some of the consequences of this conjecture. This conjecture may turn out to be extremely useful in combinatorics if these consequences are any indication.

The third chapter deals with Muir algebra. Muir hinted at an algebra to deal with permanents long time back. But this

CHAPTER - I

THE DIAGONALS IN MATRICES

1.0 INTRODUCTION

Matrices are ubiquitous in pure and applied mathematics. They represent linear transformations and hence find extensive applications in vector spaces. The determinant associated with square matrices are extremely useful in almost all branches of mathematics. But when we deal with combinatorial structures we need functions of matrices other than determinants. A modification of determinants known as permanents find increasing use in combinatorics. But if we carefully analyse determinants and permanents one aspect that strikes our mind forcefully is the concept of a diagonal. Rows and columns of matrices are well known. Less known are the diagonals. But most of the useful functions of matrices, directly or indirectly, can be expressed through the diagonals.

After defining diagonals of a square matrix we initially consider diagonals of doubly stochastic matrices. Here Wang [17] proposed three conjectures on diagonal sums of doubly stochastic matrices. In this chapter two conjectures are affirmatively settled and the third is disproved in a general way and also through counter-examples. In fact the proofs consider much wider problems. Then through the concept of disjoint diagonals latin squares are analysed and some interesting results come out regarding completion of a set of mutually orthogonal latin squares of a special class. This naturally leads to the consideration of finite projective planes and we get a nice characterization.

1.1 NOTATIONS AND TERMINOLOGIES

Let $A = (a_{ij})$ be an $n \times n$ matrix over some set X. Let S_n be the full symmetric group of degree n acting on the set $N = \{1, 2, 3, \ldots, n\}$. If $s \in S_n$, then $s(i) \in N$ for every $i \in N$ and $i \neq j$ implies $s(i) \neq s(j)$. If $A = (a_{ij})$, then the collection $\{a_{i,s(i)} \mid i \in N\}$ is called the DIAGONAL s of A. When no confusion arises the 'positions' $\{(i, s(i)) \mid i \in N\}$ can be defined as the diagonal s.

When we deal with rectangular matrices we may still use the concept of diagonals. But usually we restrict our attention to $m \times n$ matrices with $m \leq n$. This is not a severe restriction as either A or its transpose A^T satisfies this restriction. Let $S_{m,n}$ be the set of all one-one mappings from $M = (1, 2, 3, \ldots, m)$ into $N = (1, 2, 3, \ldots, n)$. Clearly then $S_n = S_{n,n}$. If $A = (a_{ij})$ is an $m \times n$ matrix with $m \leq n$, then for $s \in S_{m,n}$, the collection $\{a_{i,s(i)} \mid i \in M\}$ is called the diagonal s of A.

If X is a commutative ring we define <u>diagonal sums</u> and <u>diagonal products</u> as follows. For $s \in S_{m,n}$ the diagonal sum of A is $\sum_{i \in M} a_{i,s(i)}$ and the diagonal product is $\prod_{i \in M} a_{i,s(i)}$.

The diagonals, diagonal sums and diagonal products are extremely useful in combinatorics. This chapter deals mainly with these.

Suppose $m \leq n$. Let $ND_{r,n}$ denote the set of all non-decreasing r-sequences of elements from $(1,2, \ldots, n)$. Let $IN_{r,n}$ denote the set of all strictly increasing r-sequences of elements from $(1, 2, \ldots, n)$. Of course we assume that $r \leq n$. If $A = (a_{ij})$ is an $m \times n$ matrix, then for $s \in ND_{r,m}$ and $t \in ND_{s,n}$ then $A[s \mid t]$ represents the $r \times s$ submatrix of A with (i,j)th element equal to $a_{s(i),t(j)}$ where $s(i)$ is the ith component of the sequence s and $t(j)$ is the jth component of the sequence t. Even when $s \in IN_{r,m}$ or $t \in IN_{s,n}$ we use the notation $A[s \mid t]$ in the same manner.

Suppose $s \in IN_{r,m}$, $t \in IN_{s,n}$. Then we define $A(s \mid t)$ as $A[s^c \mid t^c]$ where s^c is obtained from s by deleting the elements of s from $(1,2, \ldots, m)$ and t^c is obtained from t by deleting the elements of t from $(1,2, \ldots, n)$. For obvious reasons we may call $A[s \mid t]$ and $A(s \mid t)$ as complements of each other. Note that $A(i \mid j)$ is the matrix obtained from A by omitting its ith row jth column. For example if A is an $n \times n$ matrix on a commutative ring X, then the expansion of the determinant of A in terms of ith row will take the form det $A = \sum_{j \in N} (-1)^{i+j} a_{ij} \det A(i | j)$.

We use the symbol R for the real field and C for the complex field. If $A = (a_{ij})$ is an n x n matrix on R, then we call A, a <u>stochastic matrix</u> if $a_{ij} \geq 0$ for all i, j and $\sum_{j \in N} a_{ij} = 1$ for every i. If A and A^T are both stochastic, then A is said to be <u>doubly stochastic</u>. The set of all n x n doubly stochastic matrices will be represented by D_n. The set of all m x n matrices with entries in a set X will be represented by $M_{m,n}(X)$.

Suppose $s \in S_n$. Let δ_{ij} represent the usual Kronecker delta. Then $P_s = (\delta_{i,s(j)}) \in M_{n,n}(R)$ is called the permutation matrix associated with the permutation s. If $X = \{0, 1\}$, then the elements of $M_{m,n}(X)$ will be called (0,1)-matrices. Note that P_s is a (0, 1)-matrix. Let Π_n denote the set of all n x n permutation matrices.

1.21 DIAGONAL SUMS

Wang [19] gave three conjectures. We settle all the three conjectures here.

Suppose $A = (a_{ij}) \in M_{n,n}(R)$. Then h(A), called the maximal diagonal sum of A, is defined as follows.

$$h(A) = \max_{s \in S_n} \sum_{i \in N} a_{i,s(i)}$$

clearly we can also define $h(A)$ in the following manner.

1.22... $$h(A) = \max_{P \in \Pi_n} \operatorname{tr}(PA) = \max_{P \in \Pi_n} \operatorname{tr}(AP)$$

where $\operatorname{tr}(A)$, the trace of A, is $\sum_{i=1}^{n} a_i$

1.23 WANG'S CONJECTURE I : If $A, B \in D_n$, then

$$h(A) + h(B) - h(AB) \leq n.$$

We will prove a stronger result. Let $U = [0, 1]$, the closed unit interval of real numbers from 0 to 1.

THEOREM 1.24 : If $A, B \in M_{n,n}(U)$, then

$$h(A) + h(B) - h(AB) \leq n.$$

PROOF : We will prove an interesting lemma first.

LEMMA 1.25 : If $A, B \in M_{n,n}(U)$, then $\operatorname{tr}(A) + \operatorname{tr}(B) - \operatorname{tr}(AB) \leq n$.

If $A = (a_{ij})$ and $B = (b_{ij})$, then

$$\operatorname{tr}(A) + \operatorname{tr}(B) - \operatorname{tr}(AB) = \operatorname{tr}(A+B-AB) = \sum_{i \in N} (a_{ii} + b_{ii} - \sum_{j \in N} a_{ij} b_{ji})$$

$$\leq \sum_{i \in N} (a_{ii} + b_{ii} - a_{ii} b_{ii}) = \sum_{i \in N} [(a_{ii}-1)(1-b_{ii})+1]$$

$$\leq \sum_{i \in N} 1 = n \quad \text{for } (a_{ii}-1)(1-b_{ii}) \leq 0.$$

Now, from the definition 1.22, there must exist $P_1, P_2 \in \Pi_n$ such that $h(A) = \operatorname{tr}(P_1A)$, $h(B) = \operatorname{tr}(BP_2)$

By lemma 1.25, $tr(P_1A) + tr(BP_2) - tr(P_1ABP_2) \leqslant n$

But $tr(P_1ABP_2) = tr(P_2P_1AB) \leqslant h(AB)$ for $P_1 P_2 \in \Pi_n$.

1.26... Thus $h(A) + h(B) - h(AB) \leqslant h(A) + h(B) - tr(P_1ABP_2) \leqslant n$.

This proves Theorem 1.24 and consequently Wang's conjecture I.

1.27 THE CASE OF EQUALITY : It is interesting to see when the equality holds in Theorem 1.24. Assuming equality we have, from 1.26, $h(A) + h(B) - h(AB) = tr(P_1A) + tr(BP_2) - tr(P_1ABP_2) = n$

Taking $P_1A = (c_{ij})$ and $BP_2 = (d_{ij})$ we have

$$\sum_{i \in N} \left[c_{ii} + d_{ii} - \sum_{j \in N} c_{ij}\, d_{ji} \right] = n.$$

But $c_{ii} + d_{ii} - \sum_{j \in N} c_{ij}\, d_{ji} \leqslant c_{ii} + d_{ii} - c_{ii}\, d_{ii} \leqslant 1$.

Hence $c_{ii} + d_{ii} - \sum_{j \in N} c_{ij}\, d_{ji} = c_{ii} + d_{ii} - c_{ii}\, d_{ii} = 1$ for each $i \in N$.

But $c_{ii} + d_{ii} - c_{ii}\, d_{ii} = 1$ if and only if atleast one of c_{ii} or d_{ii} is 1 for each $i \in N$. Also $\sum_{j \in N} c_{ij}\, d_{ji} = c_{ii}\, d_{ii}$ if and only if $\sum_{\substack{j \in N \\ j \neq i}} c_{ij}\, d_{ji} = 0$ for each $i \in N$ or equivalently $c_{ij}\, d_{ji} = 0$ for $i \neq j$.

Hence the equality holds if and only if the following conditions are satisfied.

1.28... $c_{ii} = 1$ or $d_{ii} = 1$ (or both) for each $i \in N$

1.29... $c_{ij}\, d_{ji} = 0$ for each $i \neq j$.

PARTICULAR CASE : If $A, B \in D_n$, then $P_1A, BP_2 \in D_n$. Hence the condition 1.28 automatically implies the condition 1.29. Hence 1.28 is a necessary and sufficient condition for the equality in Wang's conjecture I.

1.3 GENERALIZATION OF THEOREM 1.24

Let G be any subgroup of S_n. For $A \in M_{n,n}(R)$ define $h_G(A)$, the maximal diagonal sum of A restricted to the subgroup G, as follows.

1.31... $$h_G(A) = \max_{s \in G} \sum_{i \in N} a_{i,s(i)} = \max_{s \in G} tr(P_S\ A) = \max_{s \in G} tr(AP_S)$$

Then we have the following generalization of Theorem 1.24.

THEOREM 1.32 : If $A, B \in M_{n,n}(U)$, then for any subgroup G of S_n, $h_G(A) + h_G(B) - h_G(AB) \leq n$.

PROOF : Let $\Pi_n(G)$ be the set of all permutation matrices P_S for $s \in G$. It is easily seen that $P_sP_t = P_{st}$ for s, $t \in G$ and hence $\Pi_n(G)$ with product as the binary operation is isomorphic image of G. There exist $P_1, P_2 \in \Pi_n(G)$ such that $h_G(A) = tr\ (P_1\ A)$ and $h_G(B) = tr(BP_2)$. By Lemma 1.25 we get

$$tr(P_1A) + tr(BP_2) - tr(P_1ABP_2) \leq n$$

But $tr(P_1ABP_2) = tr\ (P_2P_1AB) \leq h_G(AB)$ for $P_2P_1 \in \Pi_n(G)$.
Thus $h_G(A) + h_G(B) - h_G(AB) \leq n$.

1.4 DISJOINT DIAGONALS

Suppose s, $t \in S_n$. Let $A \in M_{n,n}(R)$. We say that the 'diagonals' s and t of A are disjoint if the sets $\{(i,s(i)) \mid i \in N\}$ and $\{(i,t(i)) \mid i \in N\}$ are disjoint. Note that the disjointness of diagonals s and t of A has nothing to do with the elements of A. It is simply a property of the permutations s and t.

LEMMA 1.41 : Diagonals s and t are disjoint if and only if

$$\operatorname{tr}(P_s P_t^T) = 0$$

PROOF : $$(P_s P_t^T)_{ij} = \sum_{r \in N} (P_s)_{ir} (P_t^T)_{rj} = \sum_{r \in N} (P_s)_{ir} (P_t)_{jr}$$

$$= \sum_{r \in N} \delta_{i,s(r)} \delta_{j,t(r)} = \sum_{r \in N} \delta_{s(r),t(r)}$$

= number of elements common to $\{(r,s(r)) \mid r \in \}$ and $\{(r,t(r)) \mid r \in N\}$.

Lemma 1.41 now follows readily.

1.42 DEFINITION : Suppose $A = (a_{ij}) \in M_{n,n}(R)$. We say that s is a ZERO DIAGONAL of A if $a_{i,s(i)} = 0$ for each $i \in N$. Now we are ready to state Wang's conjucture II.

1.43 WANG'S CONJECTURE II : Let $A \in D_n$, and let $t_1, t_2, \ldots, t_m$ be m mutually disjoint zero diagonals of A, $1 \leq m \leq n-1$. If every diagonal disjoint from each t_j, $j = 1, 2, \ldots, m$ has a constant sum (this constant sum must be $n/(n-m)$) then all entries off the m zero diagonals are equal to $1/(n-m)$.

PROOF : We know that every doubly stochastic matrix is in the convex hull of permutation matrices. This is a well known result due to Birkhoff and Von Neumann. Thus we have

1.44... $A = \sum_{i=1}^{r} a_i P_i$ where $P_i \in \Pi_n$ for $i = 1, 2, \ldots, r$, $a_i > 0$ for $i = 1, 2, \ldots, r$ and $\sum_{i=1}^{r} a_i = 1$.

Clearly each P_i is disjoint from each P_{t_j}. Thus by Lemma 1.41 $tr(P_i P_{t_j}^T) = 0$ for $i = 1, 2, \ldots, r$ and $j = 1, 2, \ldots, m$. By the hypothesis of the theorem

1.45... $tr(AP_i^T) = n/(n-m)$ for $i = 1, 2, \ldots, r$.

From 1.44 and 1.45 we get $n/(n-m) = tr \sum_{i=1}^{r} a_i P_i P_j^T = \sum_{i=1}^{r} a_i e_{ij}$ where $e_{ij} = tr(P_i P_j^T)$, for $j = 1, 2, \ldots, r$.

Let $L = (a_1, a_2, \ldots, a_r)$ and $E = (e_{ij})$, an $r \times r$ matrix and $J = (1, 1, \ldots, 1)$ with r elements. Then we have

1.46... $L J^T = 1, \quad LE = \frac{n}{(n-m)} J$

From 1.46 we get the following.

$$1.47\ldots \quad LEL^T = \frac{n}{(n-m)} \; JL^T = \frac{n}{n-m} = \sum_{i=1}^{r} \sum_{j=1}^{r} a_i \, a_j \, e_{ij}$$

Also $tr(AA^T) = \sum_{i=1}^{r} \sum_{j=1}^{r} a_i \, a_j \, e_{ij}$ from 1.44. Thus we get,

$$1.48\ldots \quad tr(AA^T) = \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}^2 = n \,/\, (n-m).$$

If $a_{ij_1^i}, a_{ij_2^i}, \ldots, a_{ij_{n-m}^i}$ are the (n-m) elements of the ith row of A <u>not</u> in any diagonal t_i, $i = 1, 2, \ldots, m$, then clearly

$$\sum_{r=1}^{n-m} a_{ij_r^i} = 1 \, .$$

Hence the mean of such a_{ij}'s for each i is $\frac{1}{(n-m)}$ and hence so must be the overall mean. On the otherhand, by 1.48, the overall mean of the squares of such a_{ij}'s $1/(n-m)^2$. This implies that such a_{ij} must each be equal to $1/(n-m)$.

This completes the proof of Wang's conjecture II. We will give a second proof which is much more elegant. For this purpose we need a lemma.

LEMMA 1.49 : Let $a_i \geq 0$, $i = 1, 2, \ldots, n$ and $\sum_{i=1}^{n} a_i > 0$. Then for any real numbers $x_1, x_2, \ldots, x_n$

$$\sum_{i=1}^{n} a_i \, x_i^2 = (\sum_{i=1}^{n} a_i x_i)^2 \,/\, \sum_{i=1}^{n} a_i \quad \text{implies}$$

$$a_i x_i = a_i \frac{\sum_{i=1}^{n} a_i x_i}{\sum_{i=1}^{n} a_i} \qquad \text{for } i = 1, 2, \ldots, n.$$

PROOF : By Cauchy's inequality $(\sum a_i x_i)^2 = (\sum \sqrt{a_i}\ \ x_i\sqrt{a_i}\ )^2$ $\leq (\sum a_i)\ (\sum a_i x_i^2)$ equality holding if and only if for some K, $\sqrt{a_i}\ x_i = K\sqrt{a_i}$ for $i = 1, 2, \ldots, n$ or equivalently $a_i x_i = Ka_i$. Obviously then $K = \sum a_i x_i \ / \sum a_i$ and hence $a_i x_i = a_i \dfrac{\sum a_i x_i}{\sum a_i}$ for $i = 1, 2, \ldots, n$.

Another interesting result we need is the following.

1.5 ... Suppose A is an m x n matrix. Let S denote the set of all positions in A, i.e., $S = \{(i,j) \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$. For $T \subseteq S$ define the incidence matrix $E_T = (e_{ij})$ of order m x n by setting $e_{ij} = 1$ if $(i,j) \in T$ and 0 otherwise.

Then $\sum_{(i,j) \in T} a_{ij} = tr(E_T^T\ A)$ where E_T^T is the transpose of E_T. This result is quite obvious.

II PROOF OF WANG'S CONJECTURE II : Let $A = \sum_{i=1}^{r} a_i\ P_i$ where $P_i \in \Pi_n$ for $i = 1, 2, \ldots, r$ and $a_i > 0$, $\sum_{i=1}^{r} a_i = 1$.

Let us assume ONLY that the diagonal sums of A corresponding to these r diagonals P_i are equal to $n/(n-m)$.

1.51... $tr(P_i^T\ A) = n/(n-m)$ for $i = 1, 2, \ldots, r$.

Let T be the set of all positions in A off the m zero diagonals and let E_T be its incidence matrix. Then, 1.52... $\sum_i \sum_j e_{ij} = n(n-m)$ Also $n = \sum_i \sum_j a_{ij} = \sum_i \sum_j e_{ij} a_{ij}$

Now $tr(A^T A) = tr\left(\sum_{i=1}^{r} a_i P_i^T A\right) = \sum_{i=1}^{r} a_i \frac{n}{(n-m)} = n/(n-m)$ by 1.51

But $tr(A^T A) = \sum_i \sum_j a_{ij}^2 = \sum_i \sum_j e_{ij} a_{ij}^2$ whence

$$\sum_i \sum_j e_{ij} a_{ij}^2 = n/(n-m) = n^2/n(n-m) = \left(\sum_i \sum_j e_{ij} a_{ij}\right)^2 \Big/ \sum_i \sum_j e_{ij}$$

Hence by lemma 1.49, $e_{ij} a_{ij} = e_{ij} \dfrac{\sum\sum e_{ij} a_{ij}}{\sum\sum e_{ij}} = \dfrac{e_{ij}}{(n-m)}$

and this proves Wang's Conjecture II.

The II Proof lends itself to an extension of Wang's conjecture as follows.

THEOREM 1.53 : Suppose A [redacted] k x n row stochastic matrix. Let $A = \sum_{i=1}^{r} a_i B_i$ where $a_i > 0$, $a_i = 1$ and B_i is a k x n row stochastic (0,1) matri[redacted] for i = 1, 2, ..., r. Suppose in each row of A there are m positions, $0 \leq m \leq n-1$, where A has zero entry. If the sum of all entries in A corresponding to entries 1 in B_i is $K/(n-m)$ for i = 1, 2, ..., r then every entry off the mk zero positions is equal to $1/(n-m)$.

PROOF : For a row stochastic matrix, [redacted] A a representation of the form $\sum a_i B_i$ always exists. In fact row stochastic matrices are in the convex hull of row stochastic (0,1) matrices.

Let T be the set of all k(n-m) positions off the km zero positions and let E_T be the incidence matrix of T. Then,

1.54... $\sum_i \sum_j e_{ij} = k(n-m)$, $tr(B_i^T A) = k/(n-m)$ for $i = 1, 2, \dots, r$.

1.55... $$k = \sum_i \sum_j a_{ij} = \sum_i \sum_j e_{ij}\, a_{ij}$$

Now $\sum_i \sum_j e_{ij}\, a_{ij}^2 = \sum_i \sum_j a_{ij}^2 = tr(A^T A) = tr \sum_i a_i\, B_i^T A = \sum_i a_i tr(B_i^T A)$

$$= \sum a_i \frac{k}{n-m} = k \,/\, (n-m) \text{ using 1.55.}$$

Thus $\sum_i \sum_j e_{ij}\, a_{ij}^2 = k \,/\, (n-m) = k^2 \,/\, k(n-m) = (\sum_i \sum_j e_{ij} a_{ij})^2 \,/\, \sum_i \sum_j e_{ij}$

Applying Lemma 1.49 we get $e_{ij}\, a_{ij} = e_{ij} \dfrac{\sum\sum a_{ij} e_{ij}}{\sum\sum e_{ij}} = \dfrac{e_{ij}}{(n-m)}$

and the proof is complete.

Note that, in this proof, the basic result seems to be unconnected with matrices. A more general result applied to matrices yields theorem 1.53 and Wang's conjecture II. This is an abstract generalization given the name constancy of functions restricted to a subset.

1.56 CONSTANCY OF FUNCTIONS RESTRICTED TO A SUBSET

Suppose S is a finite nonempty set. Let F be the algebra of all real functions on S under usual addition and scalar multiplication and multiplication defined by

$$(f.g)(x) = f(x).\, g(x) \text{ for each } x \in S \text{ and } f,\, g \in F.$$

Define f^2 as $f.f$ for every $f \in F$. For $A \subset S$ define $X_A \in F$ by $X_A(x) = 1$ if $x \in A$ and 0 otherwise. Thus X_A is merely the characteristic function of the set A. Let $|A|$ represent the cardinality of A. If A and B are two sets let $A-B$ represent the relative complement of B in A.

THEOREM 1.57 : Suppose $f = \sum_{i=1}^{t} r_i X_{A_i}$ where A_i's are distinct (not necessarily disjoint) subsets of S and r_i's are non-zero real numbers. Let $B \subset S$ be such that $X_B.f = 0$, the zero function of F. Let $\sum_{x \in A_i} f(x) = k_i$ for $i = 1, 2, \ldots, t$. If $\sum_{i=1}^{t} r_i k_i = \left(\sum_{x \in S-B} f(x)\right)^2/|S-B|$ then f restricted to $S-B$ is a constant function.

PROOF : Clearly $|S - B| > 0$.

$$\sum_{x \in S} (X_{S-B}.\, f^2)(x) = \sum_{x \in S} f^2(x) = \sum_{x \in S} \left(f(x).\sum_{i=1}^{t} r_i X_{A_i}(x)\right)$$

$$= \sum_{i=1}^{t} r_i \sum_{x \in S} (X_{A_i}.\, f)(x)$$

$$= \sum_{i=1}^{t} r_i \sum_{x \in A_i} f(x) = \sum_{i=1}^{t} r_i k_i$$

Also $\sum_{x \in S} (X_{S-B}.\, f)(x) / \sum_{x \in S} X_{S-B}(x) = \left(\sum_{x \in S-B} f(x)\right)/|S-B|$

Hence $\sum_{x \in S} X_{S-B}(x).f^2(x) = \left(\sum_{x \in S} X_{S-B}(x)\, f(x)\right)^2 / \left(\sum_{x \in S} X_{S-B}(x)\right).$

Thus by Lemma 1.49 we get

$$X_{S-B}(x)\ f(x) = X_{S-B}\ (x) \left\{ \frac{\sum_{x \in S} (X_{S-B} \cdot f)(x)}{\sum_{x \in S} X_{S-B}\ (x)} \right\}$$

This means that the function f, restricted to S-B is a constant function.

CORDLLARY 1.58 : Suppose $|A_i| = n$ for $i = 1, 2, \ldots, t$ and $f(x) = \sum_{i=1}^{t} r_i X_{A_i}(x)$ where $r_i > 0$, $\sum r_i = 1$. Let $\sum_{x \in A_1} f(x) = k$, for $i = 1, 2, \ldots, t$. If $B \subset S$ be such that $f(x) = 0$ for each $x \in B$ and $k = n^2 \Big| \, |S-B|$, then f restricted to S-B is a constant function.

From Corollary 1.58 we can easily prove Wang's Conjecture II and its generalization to now stochastic matrices. We shall give only the proof for Wang's conjecture II, its generalization being similarly proved.

WANG'S CONJECTURE II : PROOF : Let $S = \left\{(i, j) \,\middle|\, i, j = 1, 2, \ldots, n\right\}$. Let A be an n x n doubly stochastic matrix. Let B be the set of all points in S belonging to the m zero diagonals. Clearly $|B| = nm$. Thus $a_{ij} = 0$ if $(i, j) \in B$. Let $A = \sum_{i=1}^{t} r_i P_i$ where $r_i > 0$ and $\sum r_i = 1$. A may be considered to be a function from S to R. Let A_i be the subset of S representing the positions occupied by 1'S in P_1. Then $|A_i| = n$ for $i = 1, 2, \ldots, t$. Clearly P_i's are disjoint from the m zero diagonals. A, considered as a function from S to R can be written as $A = \sum_{i=1}^{r} r_i X_{A_i}$. Let us assume ONLY that the sum of the elements of A on each A_i is $n \mid (n-m)$.

Thus $\sum_{x \in S} A(x) = \frac{n}{(n-m)} = \sum_{x \in S} r_i X_{A_i}(x)$ for $i = 1, 2, \ldots, t$.

But $n \mid (n-m) = n^2 \mid n(n-m) = n^2 \mid |S-B|$. By Corollary 1.58, we get $X_{S-B}(x) A(x) = k$, a constant. But $\sum_{x \in S} X_{S-B}(x) A(x) = \sum_{x \in S} A(x) = n$. Now $k = X_{S-B}(x) A(x) = (\sum_{x \in S} X_{S-B}(x) A(x)) \mid (\sum_{x \in S} X_{S-B}(x)) = n \mid n(n-m) = 1/(n-m)$. Hence $X_{S-B}(x) A(x) = 1/(n-m)$. This completes the proof.

1.59 WANG'S CONJECTURE III : Wang [17], discussing the important properties of the function h, the maximum diagonal sum, restricted to D_n observed that it behaves like the usual rank function. Thus he conjectured the following.

$$\text{If } A, B, C \in D_n, \text{ then } h(AB) + h(BC) - h(ABC) \leq h(B).$$

This is the familiar Frobenius inequality if h is replaced by the rank function r.

This conjecture turns out to be false. Rather than disproving the conjecture by giving counter examples obtained by trial and error we develop here a theory from which an infinite number of counter examples can be generated. To this end we will prove the following theorem quite useful in its own right. A diagonal s is said to be positive in A if all the elements of A in s are positive.

THEOREM 1.60 : If $A \in D_n$, $B \in M_{n,n}(R)$, then $tr(AB) = h(AB) = h(B)$ if and only if s is positive diagonal of A implies $tr(P_s^T B) = h(B)$.

PROOF : Suppose $tr(AB) = h(AB) = h(B)$. Let $A = \sum_{i=1}^{t} r_i P_i$ where $r_i > 0$ for $i = 1, 2, \ldots, t$ and $\sum_{t} r_i = 1$ & $P_i \in \Pi_n$ for each i. Then $h(AB) = tr(AB) = \sum_{i=1}^{t} r_i \; tr(P_i B) \leqslant \sum_{i=1}^{t} r_i \; h(B) = h(B)$. Hence $tr(P_i B) = h(B)$ for all i. Moreover if s is a positive diagonal of A, then there is a representation of A in the form $\sum r_i P_i$ with $P_1 = P_s^T$ [See Ryser [14]]. Hence we conclude that $tr(P_s^T B) = h(B)$ whenever 'P_s' is a positive diagonal of A.

Suppose s is a positive diagonal of A implies $tr(P_s^T B) = h(B)$. If $A = \sum_{i=1}^{t} r_i P_i$ then it follows that $tr\,(P_i B) = h(B)$ for $i = 1, 2, \ldots, t$. Hence we get

1.61... $tr(AB) = \sum r_i \; tr(P_i B) = \sum r_i \; h(B) = h(B)$. Also,

1.62... $tr(AB) \leqslant h(AB)$

Thus $h(AB) = \max_{P \in \Pi_n} tr(PAB) = \max_{P \in \Pi_n} \sum_{i=1}^{t} r_i \; tr(PP_i B)$

$\max_{P \in \Pi_n} \sum_{i=1}^{t} r_i \; h(B) = h(B)$ for $tr(PP_i B) \leqslant h(B)$.

From 1.61 and 1.62 it follows that

$h(B) = tr(AB) \leq h(AB) \leq h(B)$. Thus $h(AB) = tr(AB) = h(B)$.

COROLLARY 1.63 : Suppose $A \in D_n$. Let $A = \sum_{i=1}^{t} r_i P_i$. Let $B \in M_{n,n}(R)$. If $tr(P_i B) = h(B)$ for $i = 1, 2, \ldots, t$ then $tr(P_s^T B) = h(B)$ where s is any positive diagonal of A.

In other words if B has a number of maximum diagonals then any diagonal formed solely out of the positions occupied by those maximum diagonals is also a maximum diagonal. This is a familiar result in an optimal allocation problem namely the assignment problem which is usually proved by using duality theorem in linear programming.

Let us now see how Wang's Conjecture III can be disproved. Suppose we choose A, $B \in D_n$ such that $h(AB) = h(B)$, then, if the conjecture is to be true, it is necessary that $h(BC) - h(ABC) \leq 0$. But in general, for A, B, $C \in D_n$ we have $h(BC) - h(ABC) \geq 0$. Thus "almost all" $C \in D_n$ will then disprove the conjecture.

To choose A, $B \in D_n$ such that $h(AB) = h(B)$ we use the previous theorem. We choose an arbitrary $B \in D_n$ such that it has atleast two maximum diagonals, say, s and t. Then we may choose $A = r P_s + (1-r) P_t$ for $0 < r < 1$.

Then $tr(AB) = h(B) = h(AB)$. Then we may choose "almost any" $C \in D_n$. Let us see a numerical example.

Let $B = \frac{1}{5}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$ This clearly has two maximum diagonals s and t where $P_s = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ and $P_t = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

Take $A = \frac{1}{3} P_s + \frac{2}{3} P_t = \frac{1}{3}\begin{bmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix}$, $AB = \frac{1}{15}\begin{bmatrix} 6 & 5 & 4 \\ 4 & 6 & 5 \\ 5 & 4 & 6 \end{bmatrix}$

$tr(AB) = h(AB) = 6/5 = h(B)$. Taking $C = A$ we have

$$BC = AB = \frac{1}{15}\begin{bmatrix} 6 & 5 & 4 \\ 4 & 6 & 5 \\ 5 & 4 & 6 \end{bmatrix}, \quad ABC = \frac{1}{45}\begin{bmatrix} 14 & 14 & 17 \\ 17 & 14 & 14 \\ 14 & 17 & 14 \end{bmatrix}$$

Thus $h(ABC) = 17/15 < 6/5 = h(BC)$. This counter example disproves Wang's Conjecture III.

NOTE : s is defined by $s(1) = 2$, $s(2) = 3$, $s(3) = 1$ and t is defined by $t(1) = 3$, $t(2) = 1$, $t(3) = 2$.

$$P_s = (\delta_{i,s(j)}) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and } P_t = (\delta_{i,t(j)}) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

1.64 DISJOINT DIAGONALS, LATIN SQUARES AND PROJECTIVE PLANES :

A latin square on an n-set X is an element $A \in M_{n,n}(X)$ such that every row and every column of A contains all the elements of X. Though latin squares are, by their very definition, matrices very little matrix theoretic operations are generally used to deal with them. We are going to analyse latin squares by the concept of disjoint diagonals and permutation matrices. Let $L_n(X)$ represent the set of all latin squares on the n-set X. The elements of X may be conveniently taken as $(x_1, x_2, \ldots, x_n)$ or even merely as $(1,2,\ldots,n)$.

Consider the positions occupied by the symbol x_i on a latin square $A \in L_n(X)$. This is clearly a diagonal, say, s_i. As no cell in A has more than one symbol from X clearly s_i's are mutually disjoint. Conversely n mutually disjoint diagonals s_i, $i = 1, 2, \ldots, n$ give a latin square. Thus a latin square L can be taken as $L = (s_1, s_2, \ldots, s_n)$ where s_i is the diagonal occupied by x_i in L. Consequently

1.65... $tr(P^T_{s_i} P_{s_j}) = 0$ for $i, j = 1, 2, \ldots, n$ & $i \neq j$.

Equivalently L can be defined by the permutation matrices P_{s_i}'s . Let us write, for convenience, P_{s_i} as P_i. Thus any latin square $L = (P_1, P_2, \ldots, P_n)$ where $P_i \in \Pi_n$ and $tr(P^T_i P_j) = 0$ for $i, j = 1, 2, \ldots, n$ & $i \neq j$.

1.66 ORTHOGONAL LATIN SQUARES : Suppose $L_1, L_2 \in L_n(X)$. We say that L_1 and L_2 are orthogonal if, when we superimpose L_1 on L_2 all the elements in the cartesian product of X with itself appear in the n^2-cells. This means precisely that each diagonal s_i of L_1 meets each diagonal t_j of L_2 in exactly one point. Conversely if the diagonals of L_1 and L_2 are such that the above condition is satisfied, then L_1 and L_2 are orthogonal.

Suppose $L_1 = (P_1, P_2, \ldots, P_n)$ & $L_2 = (Q_1, Q_2, \ldots, Q_n)$ where $P_i, Q_i \in \Pi_n$ for each i. Then L_1 and L_2 are orthogonal if and only if

1.67... $\quad tr\,(P_i^T\, Q_j) = 1$ for $i, j = 1, 2, \ldots, n$.

THEOREM 1.68 : If $L_1 = (P_1, P_2, \ldots, P_n)$ and $L_2 = (Q_1, Q_2, \ldots, Q_n)$ are orthogonal, then for any $P \in \Pi_n$, $PL_1 = (PP_1, PP_2, \ldots, PP_n)$ and $PL_2 = (PQ_1, PQ_2, \ldots, PQ_n)$ are also orthogonal.

PROOF : First of all we will prove that PL_1 and PL_2 are latin squares.

1.69... $\quad tr((PP_i)^T(PP_j)) = tr(P_i^T P^T P P_j) = tr(P_i^T\, P_j) = 0$ if $i \neq j$.

Here we have used the fact that $P^T P = I$ as any permutation matrix is trivially an orthogonal matrix.

1.69 shows that PL_1 is a latin square. In a similar manner PL_2 is also a latin square.

1.70... $\quad Tr(PP_i)^T\,(PQ_j)) = tr(P_i^T\, P^T\, PQ_j) = tr\,(P_i^T\, Q_j) = 1$ for $i, j = 1, 2, \ldots, n$.

1.70 shows that PL_1 and PL_2 are orthogonal latin squares.

COROLLARY 1.70 : If L_1 and L_2 are orthogonal latin squares so are $PL_1\,Q$ and $PL_2\,Q$ for $P, Q \in \Pi_n$.

PROOF : $\operatorname{tr}((PP_iQ)^T (PP_jQ)) = \operatorname{tr}(Q^T P_i^T P^T P P_j Q) = \operatorname{tr}(Q^T P_i^T P_j Q)$

$$= \operatorname{tr}(QQ^T P_i^T P_j) = \operatorname{tr}(P_i^T P_j) = 0 \text{ if } i \neq j.$$

$$\operatorname{tr}((PP_iQ)^T (PQ_jQ)) = \operatorname{tr}(Q^T P_i^T P^T P Q_j Q) = \operatorname{tr}(Q^T P_i^T Q_j Q)$$

$$= \operatorname{tr}(QQ^T P_i^T Q_j) = \operatorname{tr}(P_i^T Q_j) = 1 \text{ for } i,\ j = 1,2,\ldots,n.$$

INTERPRETATION : Theorem 1.68 also implies that if $s \in S_n$, then $(P_{s(1)}, P_{s(2)}, \ldots, P_{s(n)})$ is also a latin square and Corollary 1.70 implies that $(P_{s(1)}, P_{s(2)}, \ldots, P_{s(n)})$ and $(Q_{t(1)}, Q_{t(2)}, \ldots, Q_{t(n)})$ are orthogonal latin squares for $s,\ t \in S_n$ assuming that $(P_1, \ldots, P_n)$ and $(Q_1, Q_2, \ldots, Q_n)$ are orthogonal latin squares. The last result means that the orthogonality of two latin squares are unaffected even if the symbols in them are independently permuted.

THEOREM 1.71 : Suppose L and P_sL are orthogonal latin squares. Then $s \in S_n$, when expressed in cycles, has exactly one cycle of length one. In otherwords s fixes just one symbol.

PROOF : If $L = (P_1, P_2, \ldots, P_n)$ then orthogonality of L and P_sL implies $\operatorname{tr}(P_i^T(P_sP_j)) = 1$ for $i,\ j = 1,\ 2,\ \ldots,\ n$.

Taking $i = j$, $\operatorname{tr}(P_i^T P_s P_i) = \operatorname{tr}(P_i P_i^T P_s) = \operatorname{tr}(P_s) = 1$.

But $\operatorname{tr}(P_s) = 1$ is true if and only if s fixes exactly one symbol.

INTERPRETATION : If L is a latin square and we want to find another latin square M orthogonal to L merely by permuting the rows of L, then we have to fix <u>exactly</u> one row of L. Of course this does not mean that given a latin square we can form another orthogonal to it by permuting the rows. Fixation of one row is just necessary.

THEOREM 1.72 : Suppose $L, \rho_s L, \rho_s^2 L, \ldots, \rho_s^{r-1} L$ are pairwise orthogonal (mutually orthogonal), then $s \in S_n$ expressed as a product of disjoint cycles contains just one cycle of length one and all other cycles are of length atleast r.

PROOF : Suppose L is orthogonal to $\rho_s^t L$ for $t = 1, 2, \ldots, r-1$. Since $\rho_s^t = \rho_{s^t}$ it follows that s^t leaves exactly one symbol unaltered for $t = 1, 2, \ldots, r-1$. Thus any other cycle must have a length of atleast r (note that if a cycle C is of length m then C^m is identity).

COROLLARY 1.73 : Suppose $L, \rho_s L, \rho_s^2 L, \ldots, \rho_s^{r-1} L$ are mutually orthogonal and $r > \frac{n-1}{2}$, then s has a cycle of length one and another cycle of length $(n-1)$.

PROOF : Suppose s has a cycle of length t, where, $r \leq t \leq (n-2)$. Then s must have another cycle of length almost $(n-t-1)$; but $r > \frac{n-1}{2}$ implies $\frac{n-1}{2} < t \leq (n-2)$. Hence $1 \leq n-t-1 < n - \frac{n-1}{2} - 1 = \frac{n-1}{2} < r$. But this is impossible according to Theorem 1.72. The only possibility is $t = n-1$ and thus proves the corollary.

THEOREM 1.74 : Suppose $L, \rho_s L, \rho_s^2 L, \ldots, \rho_s^{r-1} L$ are mutually orthogonal latin squares with $r > \frac{n-1}{2}$ then $L, \rho_s L, \rho_s^2 L, \ldots, \rho_s^{n-2} L$ form a complete set of $(n-1)$ mutually orthogonal latin squares.

PROOF : By Corollary 1.73, s consists of a single cycle of length one, and another cycle of length (n-1). Thus $P_s^{n-1} = I$, the unit matrix. Let $L = P_1, P_2, \ldots, P_n$. Orthogonality of $P_s^u L$ and $P_s^v L$ for $0 \leq u < v \leq (r-1)$ shows that $tr\,[(P_s^u P_i)^T (P_s^v P_j)] = 1$ for every i,j. Hence $tr\,[P_i^T P_s^{v-u} P_j] = 1$. But clearly we can also interchange u and v and get $tr\,[P_i^T P_s^{u-v} P_j] = 1$.

But the set of values of u-v and v-u for $0 \leq u < v \leq r-1$ where $r > \frac{n-1}{2}$ is clearly (1, 2, ...,(n-2)) modulo (n-1). Thus we get $tr\,[P_i^T P_s^t P_j] = 1$ for each i, j and t = 1, 2, ..., (n-2).

Now consider $P_s^u L$ and $P_s^v L$ for $0 \leq u < v \leq (n-2)$. $tr\,[(P_s^u P_i)^T (P_s^v P_j)] = tr\,[P_i^T P_s^{v-u} P_j] = 1$ for every i, j for $1 \leq v-u \leq n-2$. Thus $P_s^u L$ and $P_s^v L$ are orthogonal. Hence the theorem.

INTERPRETATION : If $L, PL, P^2L, \ldots, P^{r-1} L$ are mutually orthogonal with $r > \frac{n-1}{2}$, then this set can be extended to a complete set of (n-1) mutually orthogonal latin squares. In this connection we conjecture the following.

CONJECTURE 1.74 : If for an n there are r mutually orthogonal latin squares of order n, then we can find L and P of order n such that $L, PL, P^2L, \ldots, P^{r-1} L$ are mutually orthogonal.

Note that when n is a power of a prime number it is a well known result that (n-1) mutually orthogonal latin squares of the form $L, PL, P^2L, \ldots, P^{n-2}L$ exist. In fact P permutes cyclically the last (n-1) rows of L [Mann (18)]. Thus the conjecture is certainly true when n is a power of a prime.

1.75 LATIN SQUARES AND KRONECKER PRODUCTS : Suppose $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ is an $r \times s$ matrix where a_{ij}'s and b_{ij}'s are in a commutative ring. The Kronecker product of A and B is defined by

$$1.76\ldots \quad A \times B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \text{ in partitioned form.}$$

clearly $A \times B$ is an $mr \times ns$ matrix. Important properties of the Kronecker product are the following.

1.77... $(A \times B)(C \times D) = AC \times BD$ provided both sides are meaningful.

1.78... $\mathrm{tr}(A \times B) = \mathrm{tr}\, A.\ \mathrm{tr}\, B$ if A and B square matrices.

1.79... $(A \times B)^T = A^T \times B^T$.

1.80... $(A \times B)^{-1} = A^{-1} \times B^{-1}$ if A and B are invertible square matrices.

LEMMA 1.81 : If $P \in \Pi_m$ and $Q \in \Pi_n$, then $P \times Q \in \Pi_{mn}$.

PROOF : Clearly $P \times Q$ is a (0, 1) matrix. P and Q being orthogonal matrices we get $(P \times Q)^T (P \times Q) = (P^T \times Q^T)(P \times Q) = (P^TP) \times (Q^TQ)$ $= I_m \times I_n = I_{mn}$. Thus $P \times Q$ is an orthogonal (0, 1) matrix and hence must be a permutation matrix.

THEOREM 1.82 : If $L_1 = (P_1, P_2, \ldots, P_m)$ and $L_2 = (Q_1, Q_2, \ldots, Q_n)$ are latin squares then $L_1 \times L_2$, defined by $L_1 \times L_2 = \left\{ P_i \times Q_j \,\middle|\, \begin{matrix} i=1,2,\ldots,m \\ j=1,2,\ldots,n \end{matrix} \right\}$, is a latin square of order mn.

PROOF : By Lemma 1.81 clearly $P_i \times Q_j \in \Pi_{mn}$.

Also $\mathrm{tr}\left[(P_{i_1} \times Q_{j_1})^T (P_{i_2} \times Q_{j_2})\right] = \mathrm{tr}\left[(P^T_{i_1} P_{i_2}) \times (Q^T_{j_1} Q_{j_2})\right]$

$$= \mathrm{tr}(P^T_{i_1} P_{i_2}) . \mathrm{tr}(Q^T_{j_1} Q_{j_2}) = 0 \text{ if either } i_1 \neq i_2 \text{ or } j_1 \neq j_2.$$

Thus if $(i_1, j_1) \neq (i_2, j_2)$, then $\mathrm{tr}\left[(P_{i_1} \times Q_{j_1})^T (P_{i_2} \times Q_{j_2})\right] = 0$.

Hence $L_1 \times L_2$ is a latin square.

THEOREM 1.83 : If $L_1, L_2, \ldots, L_k$ are mutually orthogonal latin squares of order m and $M_1, M_2, \ldots, M_l$ are mutually orthogonal latin squares of order n and if $r = \min(k, l)$, then $L_1 \times M_1, L_2 \times M_2, \ldots, L_r \times M_r$ are mutually orthogonal latin squares of order mn.

PROOF : By Theorem 1.82, certainly $L_1 \times M_1, L_2 \times M_2, \ldots, L_r \times M_r$ are latin squares. Consider $L_u \times M_u$ and $L_v \times M_v$ for $1 \leq u < v \leq r$.

Let $L_u = (P_{11}, P_{12}, \ldots, P_{1m})$, $L_v = (P_{21}, P_{22}, \ldots, P_{2n})$

$M_u = (Q_{11}, Q_{12}, \ldots, Q_{1m})$, $M_v = (Q_{21}, Q_{22}, \ldots, Q_{2n})$.

Then $\mathrm{tr}\left[(P_{1i_1} \times Q_{1i_2})^T (P_{2j_1} \times Q_{2j_2})\right] = \mathrm{tr}\left[(P^T_{1i_1} P_{2j_1}) \times (Q^T_{1j_2} Q_{2j_2})\right]$

$$= \mathrm{tr}(P^T_{1i_1} P_{2j_1}) . \mathrm{tr}(Q_{1i_2} Q_{2j_2}) = 1 \times 1 = 1 \text{ for every } i_1, i_2, j_1, j_2.$$

Thus $L_u \times M_u$ and $L_v \times M_v$ are orthogonal for $1 \leq u < v \leq r$ and this completes the proof.

INTERPRETATIONS : Theorem 1.83 implies that if there are k mutually orthogonal latin squares of order m and l mutually orthogonal latin squares of order n, then for r = min (k,l) certainly there are r mutually orthogonal latin squires of order mn. In fact Theorem 1.83 gives the construction of these r latin squares.

1.84 APPLICATION OF DISJOINT DIAGONALS TO FINITE PROJECTIVE PLANES:

A finite projective plane, PG(2, s), consists of $m=s^2+s+1$ points arranged in (s+1) lines, each containing (s+1) points such that any two lines have just one point in common and any two points lie in just one line. Existence of a PG(2, s) is equivalent to the existence of an $m \times m$ (0, 1) matrix N such that $N^T N = NN^T = sI + J$ where I is the unit matrix of order m and J is the $m \times m$ matrix with each entry unity.

THEOREM 1.85 : A PG(2, s) exists if and only if there exist (s+1) disjoint permutation matrices $P_1, P_2, \ldots, P_{s+1}$ of order $m \times m$ such that I together with $P_i^T P_j$, $i \neq j$ $i, j = 1, 2, \ldots, s + 1$ form m disjoint permutation matrices.

PROOF : N is (0, 1) matrix with each row sum and each column sum (s + 1). Hence there exist (s+1) disjoint permutation matrices $P_1, P_2, \ldots, P_{s+1}$ such that $N = \sum_{i=1}^{s+1} P_i$.

$$\text{Thus } N^T N = \sum_i \sum_j P_i^T P_j = s\,I + J.$$

For $i = j$, $P_i^T P_j = I$. Thus $N^T N = (s+1)I + \sum_{i \neq j}\sum P_i^T P_j = sI+J$.

Thus $\sum_{i \neq j}\sum P_i^T P_j = J - I$. $tr \sum_{i \neq j}\sum P_i^T P_j = tr(J-I) = 0$.

Thus $i \neq j$ implies $tr(P_i^T P_j) = 0$ and hence P_i's are disjoint.

Also J-I is a (0, 1) matrix and hence the permutation matrices $P_i^T P_j$ for $i \neq j$ must be disjoint and disjoint from I. Conversely suppose I, $P_i^T P_j$ for $i \neq j$ form m disjoint permutation matrices. Together the 1's in these matrices must occupy m^2 cells and hence $I + \sum_{i \neq j}\sum P_i^T P_j = J$.

Take $N = \sum_{i=1}^{s+1} P_i$. Then $N^T N = (s+1)\,I + \sum_{i \neq j}\sum P_i^T P_j = sI+J$.

Thus PG(2, s) exists.

In fact we don't require the P_i's to be disjoint in the converse and we get the following theorem.

THEOREM 1.86 : A PG(2, s) exists if and only if (s+1) permutation matrices $P_1, P_2, \ldots, P_{s+1}$ of order m exist such that $P_i^T P_j$'s for $i \neq j$ are disjoint.

PROOF : "Only if" part of the Theorem is same as that of Theorem 1.85. In "If" part disjointness of $P_i^T P_j$'s imply $tr[(P_i^T P_j)^T (P_k^T P_l)] = 0$

if $i \neq j$, $k \neq l$ and $(i, j) \neq (k, l)$. Taking $i = k$ but $j \neq l$, we get $tr [P_j^T P_k P_k^T P_l] = tr [P_j^T P_l] = 0$. Thus $tr [P_j^T P_l] = 0$ for $j \neq l$. Hence P_i's are disjoint. Now, Theorem 1.85 completes the proof.

1.87 DIFFERENCE SETS AND PG(2, s) : Suppose $d_1, d_2, \ldots, d_{s+1}$ are $(s+1)$ integers such that $d_i - d_j$ (for $i \neq j$) form the numbers $1, 2, \ldots, (m-1)$ modulo m ($= s^2 + s + 1$). Then $d_1, d_2, \ldots, d_{s+1}$ is said to be a difference set of order $(s+1)$. Finding a set of necessary and sufficient conditions on s for the existence of such a difference set is still an unsolved problem. The following theorem is well known; but we give a simple proof.

THEOREM 1.88 : Existence of a difference set $(d_1, d_2, \ldots, d_{s+1})$ implies the existence of PG(2,s).

PROOF : Let P be the permutation matrix of order $m(= s^2+s+1)$ with 1's in $(1,2)$, $(2,3)$, ..., $(m-1,m)$ and $(m,1)$ positions. Actually $P = P_\sigma$ where σ is the cycle $(m, m-1, m-2, \ldots, 3, 2, 1)$. Obviously $P, P^2, P^3, \ldots, P^m(= I)$ are disjoint. Take $P_1 = P^{d_1}$, $P_2 = P^{d_2}$, ..., $P_{s+1} = P^{d_{s+1}}$ where $(d_1, d_2, \ldots, d_{s+1})$ is a difference set. Then $P_i^T P_j = P^{d_j - d_i}$ and by the definition of a difference set $(P_i^T P_j)$'s form the matrices $P, P^2, \ldots, P^{m-1}$ in some order. Thus these matrices are disjoint and Theorem 1.86 shows that PG(2, s) exists.

Note that we actually get $N = P^{d_1} + P^{d_2} + \ldots + P^{d_{s+1}}$ and thus the proof is constructive. Clearly we could have taken P as P_σ for any m-cycle σ on $(1, 2, \ldots, m)$.

C H A P T E R - II

DIAGONAL PRODUCTS OF MATRICES

2.0 INTRODUCTION :

In this chapter we define a diagonal product for a square matrix and through this concept develop systematically the determinants and permanents. These are extended to group complexes and interesting results are derived. Then we consider the application of tensor product of vectors to permanents and quite a number of surprising generalizations to the existing results on permanents are derived. Finally we state and prove Marcus's conjecture on diagonal products and rank of positive matrices and a number of very interesting results are derived from the conjucture.

2.1 DETERMINANTS AND PERMANENTS WITH RESPECT TO GROUP COMPLEXES :

Suppose $A \in M_{n,n}$ (CR) where CR is any commutative ring. We define determinant of A by

$$\det (A) = \sum_{\in S_n} \epsilon(s) \prod_{r=1}^{n} a_{r,s(r)} \quad \text{where } A = (a_{ij}).$$

Here $\epsilon(s) = 1$ if s is an even permutation and -1 otherwise.

Permanent of A is defined by $\operatorname{Per}(A) = \sum_{s \in S_n} \prod^{n} a_{r,s(r)}$.

Suppose G is a group complex of S_n. In otherwords G is a non-empty subset of S_n. Then we define determinant of A and permanent of A with respect to the complex G as follows.

2.2... $$\det_G (A) = \sum_{s \in G} \epsilon(s) \prod_{r=1}^{n} a_{r,s(r)}$$

2.3... $$\mathrm{per}_G (A) = \sum_{s \in G} \prod_{r=1}^{n} a_{r,s(r)} .$$

Suppose $s \in S_n$. We define $P_s \in \Pi_n$ by $P_s = (\delta_{i,s(j)})$. With this definition, the following results follow easily.

2.4... $P_s P_t = P_{st}$ for $s, t \in S_n$.

2.5... $P_s^{-1} = P_s^T = P_{s^{-1}}$ for $s \in S_n$.

2.6... $P_s A = (a_{s^{-1}(i),j})$ and $A P_s = (a_{i,s(j)})$ if $A = (a_{ij})$.

From 2.2 and 2.3 it is clear that $\det_G(A)$ and $\mathrm{Per}_G(A)$ are multi-linear functions of rows (columns) of A.

THEOREM 2.7 : Suppose $G \subset S_n$ and G^{-1} denotes the set of all inverses of elements of G. Then,

(a) $\det_G (A) = \det_{G^{-1}} (A^T)$

(b) $\mathrm{per}_G (A) = \mathrm{per}_{G^{-1}} (A^T)$

PROOF :
$$\det_G(A) = \sum_{s \in G} \epsilon(s) \prod_{r=1}^{n} a_{r,s(r)}$$
$$= \sum_{s \in G} \epsilon(s) \prod_{r=1}^{n} a_{s^{-1}(r)r}$$
$$= \sum_{s \in G} \epsilon(s^{-1}) \prod_{r=1}^{n} (A^T)_{rs^{-1}(r)} \quad \text{for } \epsilon(s^{-1}) = \epsilon(s).$$
$$= \sum_{s \in G^{-1}} \epsilon(s) \prod_{r=1}^{n} (A^T)_{rs(r)}$$
$$= \det_{G^{-1}} (A^T).$$

If we treat $\epsilon(s)$ as 1 for all $s \in S_n$ we get the result (b).

COROLLARY 2.8 : If G is closed with respect to taking inverses of elements (i.e. $G^{-1} \equiv G$) then

(a) $\det_G(A) = \det_G(A^T)$

(b) $\mathrm{per}_G(A) = \mathrm{per}_G(A^T)$.

NOTATION : Suppose $G \subset S_n$ and $s \in S_n$. Then we define sG by $sG = \{sg \mid g \in G\}$ and Gs by $Gs = \{gs \mid g \in G\}$.

THEOREM 2.9 : (a) $\det_G(P_s A) = \epsilon(s) \det_{Gs}(A)$

(b) $\mathrm{per}_G(P_s A) = \mathrm{per}_{Gs}(A)$.

PROOF : $P_sA = (a_{s^{-1}(i),j})$. Hence $det_G(P_sA) = \sum_{t \in G} \epsilon(t) \prod_{r=1}^{n} a_{s^{-1}(r),t(r)}$

$$= \sum_{t \in G} \epsilon(t) \prod_{r=1}^{n} a_{r,ts(r)} = \epsilon(s) \sum_{t \in G} \epsilon(ts) \prod_{r=1}^{n} a_{r,ts(r)}$$

$$= \epsilon(s) \sum_{t \in Gs} \epsilon(t) \prod_{r=1}^{n} a_{r,t(r)} = \epsilon(s)\ det_{Gs}(A).$$

In the proof we have used the fact $\epsilon(s)\epsilon(t) = \epsilon(st) = \epsilon(ts)$ and $[\epsilon(s)]^{-1} = \epsilon(s)$.

In a similar manner (b) can be proved.

COROLLARY 2.10 : (a) $det_G(AP_s) = \epsilon(s)\ det_{sG}(A)$

(b) $per_G(AP_s) = per_{sG}(A)$.

PROOF : $det_G(AP_s) = det_{G^{-1}}(P_s^T A^T)$ by Theorem 2.7 and $P_s^T = P_{s^{-1}}$

$$= \epsilon(s^{-1})\ det_{G^{-1}s^{-1}}(A^T)$$

$$= \epsilon(s)\ det_{sG}(A)$$

(b) is similarly proved.

THEOREM 2.11 : If G is any complex in S_n and s is in the normalizer of G, then $det_G(P_sA) = det_G(AP_s)$ and $per_G(P_sA) = per_G(AP_s)$.

PROOF : Clearly $sG = Gs$. Hence $det_G(P_sA) = \epsilon(s)\ det_{Gs}(A) =$ (s) $det_{sG}(A) = det_G(AP_s)$ and $per_G(P_sA) = per_G(AP_s)$ similarly.

COROLLARY 2.12 : If G is a normal subgroup of S_n, then $\det_G(P_sA) = \det_G(AP_s)$ and $\text{per}_G(P_sA) = \text{per}_G(AP_s)$ for every $s \in S_n$. In particular $\det(P_sA) = \det(AP_s)$ and $\text{per}(P_sA) = \text{per}(AP_s)$ by setting $G = S_n$.

COROLLARY 2.13 : If G is a subgroup of S_n and $s \in G$, then
$\det_G(P_sA) = \det_G(AP_s) = \epsilon(s)\det_G(A)$ and
$\text{per}_G(P_sA) = \text{per}_G(AP_s) = \text{per}_G(A)$.

For, in this case, $sG = Gs = G$.
When $G = S_n$ we get the familiar result in determinants and permanents.

THEOREM 2.14 : Let A be an $n \times n$ matrix with rth row and sth row identical. If s in the permutation just transposing r and s, then $s \in G$ implies $\det_G(A) = 0$ when G is a subgroup of S_n.

PROOF : By Corollary 2.13, $\det_G(P_sA) = \epsilon(s)\det_G(A)$. But $P_sA = A$ and $\epsilon(s) = -1$. Hence we get $\det_G(A) = -\det_G(A)$. Thus $\det_G(A) = 0$.

It may be interesting to note that, when s is NOT in G, then $\det_G(A)$ need not vanish. For example consider $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix}$.
Let G be the subgroup $\{I, s\}$ where $s = (1\ 3)$ in S_3.
Then $\det_G(A) = 1.2.5 - 3.2.3 = -8 \neq 0$. Thus, eventhough two rows of A are identical, $\det_G(A) \neq 0$.

COROLLARY 2.15 : If two rows of A are identical then there exists a subgroup G of S_n such that $\det_G(A) = 0$. For we have only to consider the group $G = \{I, s\}$ where s is the transposition of the row indices of the two identical rows of A. In fact <u>any</u> group G containing s will do.

COROLLARY 2.16 : Suppose s interchanges r and u. Let $s \in G$, a subgroup of S_n. Then $\det_G(A)$ is unaltered if we add to uth row any multiple of rth row of A.

This follows from the multilinearity of $\det_G(A)$ as a function of rows of A and the Theorem 2.14.

2.17 GENERAL LAPLACE EXPANSION :

Let G be a subgroup of S_n and $S_n = G_1 \cup G_2 \cup \ldots \cup G_m$ be the right coset decomposition of S_n with respect to G. Then

(a) $\det(A) = \sum_{i=1}^{m} \det_{G_i}(A)$.

(b) $\operatorname{per}(A) = \sum_{i=1}^{m} \operatorname{per}_{G_i}(A)$.

These results are quite obvious. But this innocuous result contains within itself all the known expansions of det(A) and per(A). Hence this result may be aptly called the general Laplace expansion.

Let $(i_1, i_2, \ldots, i_n)$ be a permutation of $(1, 2, \ldots, n)$. Let $1 \leq r < n$. Let $\bar{G}_1$ be the subgroup of S_n fixing all the elements

$(i_1, i_2, \ldots, i_r)$ and let $\overline{G}_2$ be the subgroup of S_n fixing $(i_{r+1}, \ldots, i_n)$. Let $G = \overline{G}_1 \times \overline{G}_2$ be the direct product of $\overline{G}_1$ and $\overline{G}_2$. An element s of G can be uniquely written as $s_1 s_2$ where $s_1 \in \overline{G}_1$ and $s_2 \in \overline{G}_2$. We can also assume that s_1 acts ONLY on $(i_{r+1}, \ldots, i_n)$ and s_2 acts ONLY on $(i_1, i_2, \ldots, i_r)$.

$$\text{Thus } \operatorname{per}_s(A) = \prod_{t=1}^{n} a_{i_t\, s(i_t)} = \prod_{t=1}^{r} a_{i_t\, s(i_t)} \prod_{t=r+1}^{n} a_{i_t s(i_t)}$$

$$= \prod_{t=1}^{r} a_{i_t s_2(i_t)} \prod_{t=r+1}^{n} a_{i_t s_1(i_t)} = \operatorname{per}_{s_2} A[i_1, i_2, \ldots, i_r / i_1, i_2, \ldots, i_r].$$

$\operatorname{per}_{s_1}(i_1, \ldots, i_r / i_1, \ldots, i_r)$. Hence $\det(A) = \sum_{i=1}^{m} \det_{G_i}(A) = \sum_{i=1}^{m} \epsilon(u_i) \det(P_{u_i} A)$ where $G_i = Gu_i$ and $G = G_1 \cup G_2 \cup \ldots \cup G_m$ is the right coset decomposition of S_n with respect to $G = \overline{G}_1 \times \overline{G}_2$.

$$\text{Thus } \det(A) = \sum_{i=1}^{m} \epsilon(u_i) \sum_{s \in G} \epsilon(s) \operatorname{per}_s(P_{u_i} A)$$

$$= \sum_{i=1}^{m} \epsilon(u_i) \sum_{s_1 \in \overline{G}_1} \sum_{s_2 \in \overline{G}_2} \epsilon(s_1 s_2) \operatorname{per}_{s_1 s_2} (P_{u_i} A)$$

$$= \sum_{i=1}^{m} \epsilon(u_i) \sum_{s_2 \in \overline{G}_2} \operatorname{per}_{s_2} (P_{u_i} A)[i_1, \ldots, i_r / i_1, \ldots, i_r] . \epsilon(\sum_{s_1 \in \overline{G}_1} \operatorname{per}_{s_1} (P_{u_i} A)(i_1, \ldots, i_r / i_1, \ldots, i_r) \epsilon(s_1)$$

$$= \sum_{i=1}^{m} \epsilon(u_i) \det(P_{u_i} A)[i_1, \ldots, i_r / i_1, \ldots, i_r]$$

$$\det (P_{u_i} A) (i_1, \ldots, i_r / i_1, \ldots, i_r).$$

This is the familiar Laplace expansion with respect to the columns $(i_1, i_2, \ldots, i_r)$ of A. In a similar manner, by considering the left coset decomposition of S_n with respect to G we get the Laplace expansion with respect to the rows $(i_1, \ldots, i_r)$ of A. A similar result holds for permanents.

It is clear that by expressing G as a direct product of k subgroups and considering the right coset decomposition of S_n with respect to G we get the usual expansion of determinants and permanents for a column partition of A into k classes; for left coset decomposition we get expansions for row partitions.

2.18 *AN EXAMPLE :* *Let H be the alternating subgroup of* S_n. Then $S_n - H$ consists of all odd permutations of S_n and $S_n = H \cup (S_n - H)$ is the right (left) coset decomposition of S_n with respect to H.

$$\det(A) = \det_H(A) + \det_{S_n - H}(A) = \sum_{s \in H} \epsilon(s)\, \mathrm{per}_s(A) + \sum_{s \in S_n - H} \epsilon(s) \mathrm{per}_s(A).$$

$$= \sum_{s \in H} \mathrm{per}_s(A) - \sum_{s \in S_n - H} \mathrm{per}_s(A) = \mathrm{per}_H(A) - \mathrm{per}_{S_n - H}(A).$$

COROLLARY 2.19 : $\det(A) = \mathrm{per}_H(A) - \mathrm{per}_{S_n - H}(A)$ & $\mathrm{per}(A) =$

$$= \mathrm{per}_H(A) + \mathrm{per}_{S_n - H}(A)$$

So $\mathrm{per}_H(A) = \frac{1}{2}(\mathrm{per}(A) + \det(A))$ and $\mathrm{per}_{S_n - H}(A) = \frac{1}{2}[\mathrm{per}(A) - \det(A)]$

consequently the necessary and sufficient condition that $\det(A) = \operatorname{per}(A)$ is that $\operatorname{per}_{S_n - H}(A) = 0$, or equivalently $\operatorname{per}(A) = \operatorname{per}_H(A)$.

THEOREM 2.20 : Let G_1 and G_2 be two conjugate subgroups of S_n. Then there exists an $s \in S_n$ such that

$$\operatorname{per}_{G_n}(A) = \operatorname{per}_{G_1}(P_s^T\, A\, P_s) \quad \text{and}$$

$$\det_{G_2}(A) = \det_{G_1}(P_s^T\, A\, P_s).$$

PROOF : Let $G_2 = s\, G_1\, s^{-1}$. Then $\det_{G_2}(A) = \det_{sG_1s^{-1}}(A)$

$$= \epsilon(s)\, \det_{sG_1}(P_s^T A) = \epsilon(s)\, \epsilon(s)\, \det_{G_1}(P_s^T A P_s)$$

$$= \det_{G_1}(P_s^T\, A\, P_s).$$

Similarly $\operatorname{per}_{G_2}(A) = \operatorname{per}_{G_1}(P_s^T\, A\, P_s)$.

COROLLARY 2.21 : If G is a normal subgroup of S_n, then for any $s \in S_n$, $\det_G(A) = \det_G(P_s^T\, A\, P_s)$ and $\operatorname{per}_G(A) = \operatorname{per}_G(P_s^T\, A\, P_s)$.

2.22 PERMANENTS WITH RESPECT TO GROUP COMPLEXES AND TENSOR PRODUCT OF VECTORS ;

In this section, the very highly useful methods of Marcus and Newman [12] analysing permanents with the help of tensor products are used to extend their results considerably to permanents with respect to group complexes.

2.23 PRELIMINARY IDEAS :

Let V be an n-dimensional unitary space with an inner product represented by (x, y) for $x, y \in V$. Let $M_m(V)$ be the space of all m-multilinear functions on V. In other words $M_m(V)$ consists of all functions $f(x_1, x_2, \ldots, x_m)$ from $V \times V \times \ldots \times V$ (m factors) into the complex field with the property $f(x_1, \ldots, \alpha x_i, \ldots, x_m) = \alpha f(x_1, \ldots, x_m)$ for α, any complex number and $i = 1, 2, \ldots, m$ and $f(x_1, \ldots, x_i + y_i, x_{i+1}, \ldots, x_m) = f(x_1, \ldots, x_i, \ldots, x_m) + f(x_1, \ldots y_i, x_{i+1}, \ldots, x_m)$ for $i = 1, 2, \ldots m$ and $y_i \in V$. Let $V^{(m)}$ be the dual space of $M_m(V)$. In other words $V^{(m)}$ consists of all linear functionals from $M_m(V)$ into the complex field. We define for $u_i \in V$, $i = 1, 2, \ldots, m$, a distinguished element f of $V^{(m)}$ as follows.

f written as $u_1 \otimes u_2 \otimes \ldots \otimes u_m$ has the property that for any $F \in M_m(V)$, $f(F) = F(u_1, u_2, \ldots, u_m)$. In otherwords we have

$$(u_1 \otimes u_2 \otimes \ldots \otimes u_m)(F) \equiv F(u_1, u_2, \ldots, u_m).$$

This element, $u_1 \otimes u_2 \otimes \ldots \otimes u_m$, of $V^{(m)}$ is called the tensor product of $u_1, u_2, \ldots, u_m$ and is said to be decomposable. These tensor products together with their linear combinations over the complex field form a vector space. In fact this vector space is $V^{(m)}$. We introduce an inner product in $V^{(m)}$ by defining inner product for decomposable tensors as follows.

2.24... $(u_1 \otimes u_2 \otimes \cdots \otimes u_m, v_1 \otimes v_2 \otimes \cdots \otimes v_n) = \prod_{i=1}^{m} (u_i, v_i)$

Let G be a complex of S_n. Define a linear operator T_G over (m) by defining its effect on any decomposable tensor as follows.

2.25... $T_G(x_1 \times \ldots \times x_m) = \frac{1}{|G|} \sum_{g \in G} (x_{g^{-1}(1)} \otimes x_{g^{-1}(2)} \otimes \cdots \otimes x_{g^{-1}(m)})$

where $|G|$ is the cardinality of G (assumed to be non-empty).

2.26 *INNER PRODUCT AND PERMANENT* :

Clearly $(T_G(x_1 \otimes \cdots \otimes x_m), y_1 \otimes \cdots \otimes y_m) = \frac{1}{|G|} \sum_{g \in G} \prod_{i=1}^{m} \cdot (x_{g^{-1}(i)}, y_i)$

$$= \frac{1}{|G|} \sum_{g \in G} \prod_{i=1}^{m} (x_i, y_{g(i)})$$

$$= ((x_1 \otimes \cdots \otimes x_m), T_{G^{-1}} (y_1 \otimes \cdots \otimes y_m)).$$

Thus we get an important result.

2.27... $(T_G (x_1 \otimes \cdots \otimes x_m), (y_1 \otimes \cdots \otimes y_m)) = ((x_1 \otimes \cdots \otimes x_m), T_{G^{-1}}(y_1 \otimes \cdots \otimes y_m)) = \frac{1}{|G|} \operatorname{per}_G(AB)$

where $A = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$ is an m x n matrix and $B = (y_1^*, y_2^*, \ldots, y_m^*)$ is an n x m matrix. y^* is the conjugate transpose of y.

LEMMA 2.28 : The linear operator T_G is hermitian iff $G = G^{-1}$.

PROOF : $(T_G(x_1 \otimes \cdots \otimes x_m), (y_1 \otimes \cdots \otimes y_m) = ((x_1 \otimes \cdots \otimes x_m),$

$T_{G^{-1}}(y_1 \otimes \cdots \otimes y_m)) = ((x_1 \otimes \cdots \otimes x_m), T_G(y_1 \otimes \cdots \otimes y_m))$

if $G = G^{-1}$.

Thus T_G is hermitian. The converse follows from the easy fact that $T_G(x_1 \otimes \cdots \otimes x_m) = T_{G^{-1}}(x_1 \otimes \cdots \otimes x_m)$ for all $x_1, x_2, \ldots x_m \in V$ implies $G = G^{-1}$.

2.29 DEFINITION : If G and H are complexes of S_m, let GH denote the multiset of the $|G|\ |H|$ elements of the form g h for $g \in G$ and $h \in H$ (with proper multiplicities). Thus $|GH| = |G|\ |H|$ and in general GH is not a set. This notation differs from the usual definition of the product of two complexes. For example, if G is a subgroup of S_n, GG will be the multiset formed by taking each element of G, $|G|$ times. Hence

2.30... $\operatorname{per}_{GG}(A) = |G| \operatorname{per}_G(A)$

THEOREM 2.31 : Let $A \in M_{m,n}(C)$, $B \in M_{n,m}(C)$. Then for any complexes G and H of S_n,

2.32... $\left|\operatorname{per}_{GH}(AB)\right|^2 \leqslant \operatorname{per}_{GG^{-1}}(AA^*) \operatorname{per}_{H^{-1}H}(B^*B)$.

PROOF : Let $A = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$ and $B = (y_1^*, \ldots, y_m^*)$.

$$(T_G(x_1 \otimes \cdots \otimes x_m), T_{H^{-1}}(y_1 \otimes \cdots \otimes y_m)) = ((x_1 \otimes \cdots \otimes x_m), T_{G^{-1}} T_{H^{-1}} (y_1 \otimes \cdots \otimes y_m))$$

$$= ((x_1 \otimes \cdots \otimes x_m), T_{H^{-1}G^{-1}}(y_1 \otimes \cdots \otimes y_m)) = \frac{1}{|G||H|} \operatorname{per}_{GH}(AB).$$

We have used the fact that $T_{G^{-1}} T_{H^{-1}} = T_{H^{-1}G^{-1}}$ which is quite obvious.

Hence by Cauchy-Schuartz inequality for inner products,

$$|(T_G(x_1 \otimes \cdots \otimes x_m), T_{H^{-1}}(y_1 \otimes \cdots \otimes y_m))|^2$$
$$\leqslant |(T_G(x_1 \otimes \cdots \otimes x_m), T_G(x_1 \otimes \cdots \otimes x_m))| \times |(T_{H^{-1}}(y_1 \otimes \cdots \otimes y_m), T_{H^{-1}}(y_1 \otimes \cdots \otimes y_m))|$$

or $$\left| \frac{\operatorname{per}_{GH}(AB)}{|G||H|} \right|^2 \leqslant \frac{\operatorname{per}_{GG^{-1}}(AA^*)}{|G|^2} \frac{\operatorname{per}_{H^{-1}H}(B^*B)}{|H|^2}$$ and

cancelling $|G|^2 |H|^2$ the theorem follows.

The result 2.32 is a considerable improvement over the result of Marvin Marcus which is as follows.

2.33... If $A \in M_{m,n}(C)$ and $B \in M_{n,m}(C)$, then

$$|\operatorname{per}(AB)|^2 \leqslant \operatorname{per}(AA^*) \operatorname{per}(B^*B)$$

COROLLARY 2.34 : In 2.32 taking $A = T$, $B = T^* J$ where J is an $m \times m$ matrix with all entries unity and setting $T T^* = S$ we get

$$\left|\operatorname{per}_{GH}(TT^{*}J)\right|^{2} \leqslant \operatorname{per}_{GG^{-1}}(TT^{*})\ \operatorname{per}_{H^{-1}H}(J\,T^{*}\,T\,J)$$

or $$\left|\operatorname{per}_{GH}(SJ)\right|^{2} \leqslant \operatorname{per}_{GG^{-1}}(S)\ \operatorname{per}_{H^{-1}H}(JSJ).$$

If row sums of S are $r_1, r_2, \dots, r_m$ with $\sum r_i = r$, then

$$SJ = \begin{bmatrix} r_1 & r_1 & \cdots & r_1 \\ r_2 & r_2 & \cdots & r_2 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ r_m & r_m & \cdots & r_m \end{bmatrix} \quad \text{and } JSJ = \begin{bmatrix} r & r & \cdots & r \\ r & r & \cdots & r \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ r & r & \cdots & r \end{bmatrix}$$

Hence $\operatorname{per}_{GH}(SJ) = |G|\,|H|\ r_1\, r_2 \cdots r_m$ and

$$\operatorname{per}_{H^{-1}H}(JSJ) = |H|^{2}\, r^{m}$$

Thus $$\left|\,|G|\ |H|\ r_1\, r_2 \cdots r_m\right|^{2} \leqslant \operatorname{per}_{GG^{-1}}(S).\,|H|^{2}\, r^{m}$$

If $r \neq 0$ we get $$\operatorname{per}_{GG^{-1}}(S) \geqslant \frac{|G|^{2}\,|r_1\, r_2 \cdots r_m|^{2}}{r^{m}}$$

Let H_m denote the set of all m x m positive semidefinite hermitian matrices. Suppose G is a subgroup of S_m and $S \in H_m$. Clearly GG^{-1} is the elements of G repeated $|G|$ times. Hence we have $\operatorname{per}_{GG^{-1}}(S) = |G|\ \operatorname{per}_G(S)$. Since any element of H_m can be expressed as $T\,T^{*}$ we have proved.

THEOREM 2.35 : If $S \in H_m$ with row sums $r_1, r_2, \dots, r_m$ and $\sum r_i = r \neq 0$,. Then $\operatorname{per}_G(S) \geqslant |G|\ |r_1\, r_2 \cdots r_m|^{2} / r^{m}$.

COROLLARY 2.36 : Suppose $S \in H_m$ with row sums unity. If G is any subgroup of s_m, then $per_G(S) \geq |G| / m^m$. If $G = S_m$, this reduces to $per(S) \geq m!/m^m$. In particular this last result is true for a positive semidefinite symmetric doubly stochastic m x m matrix S.

COROLLARY 2.37 : If G = I, the trivial group of identity permutation only and $S \in H_m$, then $per_I(S) \geq |I| \; |r_1 \; r_2 \cdots r_m|^2 / r^m$ or $s_{11} \; s_{22} \cdots s_{mm} \geq |r_1 \; r_2 \cdots r_m|^2 / r^m$ where $S = (s_{ij})$. In particular for a doubly stochastic matrix in H_m we get

$$s_{11} \; s_{22} \cdots s_{mm} \geq m^{-m}$$

COROLLARY 2.38 : In 2.32 taking H = I and B as a matrix with unit column vectors, clearly $per_{H^{-1}H}(B^*B) = per_i(B^*B) = 1$. Thus we get $|per_G(AB)|^2 \leq per_{GG^{-1}}(AA^*)$. In particular if G is a subgroup of S_m this reduces to $|per_G(AB)|^2 \leq |G| \; per_G(AA^*)$. Taking $G = S_m$ we get $|per\ (AB)|^2 \leq m!\ per\ (AA^*)$. If we put A = I, the unit matrix we get $|per\ (B)|^2 \leq m!$

2.39 SOME RESULTS ON PERMANENT W.R.T A GROUP :

In 2.32 taking G = H = a subgroup of S_m we get

2.40... $|per_G(AB)|^2 \leq per_G(AA^*) \; per_G(B^*B)$

for, in this case, $GH = GG^{-1} = H^{-1}H = G$ repeated $|G|$ times.

COROLLARY 2.41 : Taking $B = I$ we get $\left| \text{per}_G (A) \right|^2 \leqslant \text{per}_G(AA^*)$.

In particular, if $A = T$, a lower triangular matrix we get

$\left| \text{per}_G(T) \right|^2 \leqslant \text{per}_G(TT^*)$. Let $TT^* = S$. Then,

$$\text{per}_G(S) \geqslant \left| \text{per}_G(T) \right|^2 = \left| t_{11}\ t_{22} \cdots t_{mm} \right|^2 = \det(TT^*) = \det S.$$

As we can always write any $S \in H_m$ in the form $S = TT^*$, where T is a lower triangular matrix we have proved.

THEOREM 2.42 : If $S \in H_m$, then for any subgroup G of S_m we have $\text{per}_G(S) \geqslant \det(S)$ (Schur's Theorem).

COROLLARY 2.43 : Taking $G = I$, $\det(S) \leqslant s_{11}\ s_{22} \cdots s_{mm}$. This is Hadamard's determinant theorem.

2.44 SOME RESULTS ON PERMANENTS W.R.T. GROUPS THAT COMMUTE :

Suppose G and H are subgroups of S_m such that $GH = HG$. If K is the <u>set</u> of all elements in GH, then, we know that K is also a subgroup of S_m and the collection GH is K repeated $|G \cap H|$ times. Now 2.32 becomes $|G \cap H|^2 \left| \text{per}_K(AB) \right|^2 \leqslant \text{per}_{GG^{-1}}(AA^*)\ \text{per}_{H^{-1}H}(B^*B)$

Also $|K| = |H|\ |G| \ / \ |G \cap H|$. Thus we get

2.45... $$\left| \frac{1}{|K|}\ \text{per}_K(AB) \right|^2 \quad \frac{1}{|G|}\ \text{per}_G(AA^*) . \ \frac{1}{|H|}\ \text{per}_H(B^*B).$$

Taking $A = T$, $B = T^*$ and $S = TT^*$ we get

2.46... $$\left| \frac{1}{|K|}\ \text{per}_K(AB) \right|^2 \leqslant \frac{1}{|G|}\ \text{per}_G(S) . \ \frac{1}{|H|}\ \text{per}_H(S).$$

In particular if $H \subset G$, then $K = G$ and we have

$$_G(S) |^2 \leq \frac{1}{|G|} \mathrm{per}_G(S) \frac{1}{|H|} \mathrm{per}_H(S) \text{ or}$$

$$\frac{1}{|G|} \mathrm{per}_G(S) \leq \frac{1}{|H|} \mathrm{per}_H(S).$$

have proved the following.

2.48 : If $S \in H_m$ and H and G are subgroups of S_m such $\subset G$, then $\mathrm{per}_G(S)/|G| \leq \mathrm{per}_H(S)/|H|$. In particular, $H = I$, $\mathrm{per}_G(S) \leq |G|\, s_{11}\, s_{22}, \ldots, s_{mm}$.

ARY 2.49 : Suppose $A \in H_m$ where m is r x s and

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{bmatrix}$$

where each A_{ij} is an s x s matrix.

$B = (\mathrm{per}\,(A_{ij}))$ be an r x r matrix. Then clearly $\mathrm{per}(B) =$ $_G(A)$, where G is a subgroup of order $(s!)^r\, r!$ Hence we get $(B)/r!\,(s!)^r \geq \mathrm{per}(A)\,/\,(sr)!$ or

2.50...
$$\mathrm{per}\,(B) \geq \frac{r!\,(s!)^r}{(sr)!}\ \mathrm{per}\,(A).$$

2.51 SOME RESULTS CONNECTING PERMANENTS WITH DETERMINANTS :

Let E represent the alternating subgroup of S_m Then EE is E repeated m!/2 times (assuming $m \geq 2$). Hence

$$\left| per_{EI}(AB) \right|^2 \leqslant per_{EE}(AA^*)\ per_{II}(B^*B) = \frac{m!}{2}\ per_E(AA^*)\ per_I(B^*B).$$

But $per(AB) = per_E(AB) + per_{S_m - E}(AB)$ and

$det(AB) = per_E(AB) - per_{S_m - E}(AB)$. Hence $per_E(AB) = \frac{1}{2}$ $[per(AB) + det(AB)]$ and it follows that,

2.52... $\left| per(AB) + det(AB) \right|^2 \leqslant m!\ [per(AA^*) + det(AA^*)]$ $per_I(B^*B)$. Taking $B = A^*$ this reduces to.

2.53... $per(AA^*) + det\ (AA^*) \leqslant m!\ \ per_I(AA^*)$.

Let $S = AA^* = (s_{ij})$. Then $per\ (S) + det(S) \leqslant m!\ s_{11}\ s_{22} \cdots s_{mm}$, for $S \in H_m$.

If in 2.52, A is nonsingular, setting $B = A^{-1}$ and $AA^* = S$ we get the following.

2.54... $4 \leqslant m!\ [per(S) + det(S)]\ per_I(S^{-1})$.

Equivalently, $m!\ \ per_I\ (S)\ [per\ (S^{-1}) + det\ (S^{-1})] \geqslant 4$.

But $per\ (S^{-1}) \geqslant det(S^{-1})$ by Schur's Theorem. Thus we get

$m!\ per_I(S)\ per(S^{-1}) \geqslant 2$ or $per(S^{-1}) \geqslant \dfrac{2}{m!\ s_{11}\ s_{22} \cdots s_{mm}}$

Compare this with $per\ (S) \leqslant m!\ s_{11}\ s_{22} \cdots s_{mm}$.

Suppose $S^{-1} = (s^{ij})$ then we get

$$m!\, s^{11}\, s^{22} \ldots s^{mm} \geqslant per(S^{-1}) \geqslant \frac{2}{m!\, s_{11}\, s_{22} \ldots s_{mm}} . \quad \text{Thus,}$$

Thus,

2.55... $$s^{11}\, s^{22} \ldots s^{mm}\, s_{11}\, s_{22} \ldots s_{mm} \geqslant 2/(m!)^2$$

DIAGONAL PRODUCTS AND RANK OF POSITIVE MATRICES :

Marvin Marvin [12] conjectured that the rank of a positive square matrix is atmost equal to the number of distinct diagonal products of the matrix. In this section we settle this conjecture affirmatively. Moreover some surprising combinatorial consequences of this conjecture are developed. More than the conjecture itself the method used to settle it-and the consequences will be found extremely interesting.

PRELIMINARY IDEAS :

Let R^+ repr.sent the set of positive real numbers. If $A \in M_{n,n}(R^+)$, then it has $n!$ diagonal products. But these need not be distinct. Let $d(A)$ represent the number of distinct diagonal products of A. Clearly $1 \leqslant d(A) \leqslant n!$. Let $r(A)$ represent the rank of the matrix A.

Marvin Marcus conjectured the following

2.56... For $A \in M_{n,n}(R^+)$, $r(A) \leqslant d(A)$.

This conjecture was verified to be true for $n < 5$ by Westwick. To settle this conjecture affirmatively we prove two theorems quite interesting by themselves.

THEOREM 2.57 : Suppose $\{a_1, a_2, \ldots, a_m\}$ and $\{b_1, b_2, \ldots, b_n\}$ are two <u>sets</u>. Then the number of distinct elements in the mn sums of the form $(a_i + b_j)$ $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$ is atleast $(m+n-1)$. The lower bound is attained if and only if either (1) $m = 1$ or $n = 1$ or (2) $m, n > 1$ and the a_i's arranged in increasing order are in arithmetic progression, the b_j's arranged in the increasing order are in A.P. and the two A. P's have the same common difference.

PROOF : For any multiset S, let D(S) represent the number of distinct elements of S. If a set of real numbers S has the property that the elements of S arranged in increasing order are in A.P. with a common difference K let us write $S \in A_K$.

Now the "if" part of theorem is quite easy to verify. We will prove only the "only if" part. The case $m = 1$ or $n = 1$ is quite trivial. Hence assume that $m, n > 1$. Without loss of generality we can assume that a_i's and b_j's are arranged in increasing order. Let $(i_1, j_1), (i_2, j_2), \ldots, (i_{m+n-1}, j_{m+n-1})$ be called a path from (1, 1) to (m, n) in the matrix $A = (a_i + b_j)$ if $i_1 = 1$, $j_1 = 1$; $i_{m+n-1} = m$, $j_{m+n-1} = n$; and for $r = 1, 2, \ldots, m+n-2$

either $i_{r+1} = i_r$ and $j_{r+1} = j_r+1$ or $i_{r+1} = i_r+1$ and $j_{r+1} = j_r$ and $1 \leq i_r \leq m$, $1 \leq j_r \leq n$.

It is clear that the entries of A on any path are strictly increasing and hence distinct. Each path contains (m+n-1) cells. Thus atleast (m+n-1) entries of A are distinct. If the number of distinct elements in A is (m+n-1) then all the paths must give the <u>same</u> increasing sequence. As $i_r+j_r = r+1$ it follows that $a_i + b_j$ depends only on $i + j$. Thus $a_{i+1}+ b_{j-1} = a_i + b_j$ for $1 \leq i \leq m-1$, $2 \leq j \leq n$. Thus $a_{i+1} - a_i = b_j - b_{j-1}$ and the theorem follows.

COROLLARY 2.58 : Suppose $m, n > 1$. If $D\left\{ a_i + b_j \;\middle|\; \begin{matrix} 1 \leq i \leq m \\ 1 \leq j \leq n \end{matrix} \right\}$ is m+n-1 then $\left\{ a_i+b_j \;\middle|\; \begin{matrix} 1 \leq i \leq m \\ 1 \leq j \leq n \end{matrix} \right\} \in A_K$ for some K. In fact if $a_{i+1} - a_i = b_j - b_{j-1} = k$ for $1 \leq i < m$, $1 < j \leq n$, then

$\{a_i \mid i = 1, 2, \ldots, m\} \in A_K$ and $\{b_j \mid j = 1, 2, \ldots, n\} \in A_K$

and $\left\{ a_i + b_j \;\middle|\; \begin{matrix} 1 \leq i \leq m \\ 1 \leq j \leq n \end{matrix} \right\} \in A_K$.

COROLLARY 2.59 : Suppose $m, n > 1$. If $\left\{ a_i + b_j \;\middle|\; \begin{matrix} 1 \leq i \leq m \\ 1 \leq j \leq n \end{matrix} \right\} \notin A_K$ for any K, then $D\left\{ a_i + b_j \;\middle|\; \begin{matrix} 1 \leq i \leq m \\ 1 \leq j \leq n \end{matrix} \right\} \geq m+n$.

THEOREM 2.60 : Suppose $\{a_1, a_2, \ldots, a_m\}$ and $\{b_1, b_2, \ldots, b_n\}$ are sets of positive numbers. Then $D\left\{ a_i\, b_j \;\middle|\; \begin{matrix} 1 \leq i \leq m \\ 1 \leq j \leq n \end{matrix} \right\} \geq m+n-1$. The lower bound (m+n-1) is attained if and only if.

either (1) $m = 1$ or $n = 1$

(2) $m, n > 1$; the a_i's arranged in increasing order are in geometric progression, the b_j's arranged in increasing order are in geometric progression and the two G.P's have the same common ratio.

COROLLARY 2.61 : Suppose $m, n > 1$. If $D\left\{a_i b_j \;\middle|\; \begin{matrix} 1 \leq i \leq m \\ 1 \leq j \leq n \end{matrix}\right\} = m+n-1$ then the distinct $a_i b_j$'s arranged in increasing order are in G.P.

COROLLARY 2.62 : Suppose $m, n > 1$. If the distinct $a_i b_j$'s arranged in increasing order are not in G.P. then they are atleast $(m+n)$ in number.

PROOF : Theorem 2.60 and the Corollaries 2.61 and 2.62 follow readily from Theorem 2.57 and Corollaries 2.58 and 2.59 applied to the sets $\left\{\log_e a_1, \log_e a_2, \ldots, \log_e a_m\right\}$, and $\left\{\log_e b_1, \log_e b_2, \ldots, \log_e b_n\right\}$ and observing the fact that the log function is monotonic increasing.

2.63 GENERALISATION :

Theorems 2.57 and 2.60 can be easily generalized to K sets of numbers. For example Theorem 2.60 can be generalised as follows.

THEOREM 2.64 : Suppose $\left\{a_{i1}, a_{i2}, \ldots, a_{in_i}\right\}$, $i = 1, 2, \ldots, k$ are k sets of positive numbers. Then $D\left\{\prod_{i=1}^{k} a_{ij_i} \;\middle|\; 1 \leqslant j_i \leqslant n_i\right\} \geqslant \sum_{i=1}^{k} n_i - (k-1)$. The lower bound is attained if and only if

either (1) atleast (k-1) of the n_i's are unity

or (2) atleast two of the n_i's are not unity; for any $n_i \neq 1$ the corresponding $a_{i1}, a_{i2}, \ldots, a_{in_i}$ arranged in increasing order are in G.P.; all these G.P's have the same common ratio.

PROOF : A proof by induction on k is quite straight forward (and is omitted here).

Let us now state a theorem which is apparently a particular case of the conjecture 2.56 but actually equivalent to it. The equivalence will be first established. Then the theorem will be proved.

THEOREM 2.65 : Let $A \in M_{n,n}(R^+)$. If $d(A) \leqslant (n-1)$, then A is singular.

EQUIVALENCE OF 2.56 AND 2.65 :

Obviously 2.56 implies 2.65. Assume that 2.65 is true. When $d(A) = k = (n-1)$ clearly 2.56 and 2.65 are the same. When $d(A) = k \geqslant n$, 2.56 is trivially true. Hence assume that

$d(A) = k \leq (n-2)$. Take any $(k+1) \times (k+1)$ submatrix $A[i_1, i_2, \ldots, i_{k+1} \mid j_1, j_2, \ldots, j_{k+1}]$ and its "complement" $A(i_1, i_2, \ldots, i_{k+1} \mid j_1, j_2, \ldots, j_{k+1})$. Clearly any diagonal product of the former multiplied by any diagonal product of the latter is a diagonal product of A. Hence by Theorem 2.60.

$d(A[i_1, i_2, \ldots, i_{k+1} \mid j_1, j_2, \ldots, j_{k+1}]) +$

$d(A(i_1, i_2, \ldots, i_{k+1} \mid j_1, j_2, \ldots, j_{k+1})) - 1 \leq d(A) = k.$

Hence $d(A[i_1, i_2, \ldots, i_{k+1} \mid j_1, j_2, \ldots, j_{k+1}]) \leq k.$

Applying theorem 2.65 to the matrix $A[i_1, i_2, \ldots, i_{k+1} \mid j_1, j_2, \ldots, j_{k+1}]$ we conclude that it should be singular. Thus any $(k+1) \times (k+1)$ submatrix of A is singular. Hence $r(A) \leq k$ and this is 2.56.

Incidentally, in view of the fact that 2.56 has been verified for all $n < 5$, the above proof shows that the conjecture is verified for all $n \geq 5$ when $k < 5$. We now prove Theorem 2.65 by induction on n.

PROOF OF THEOREM 2.65 :

Let us assume the truth of the theorem for all matrices of order upto $(n-1)$. Let $A \in M_{n,n}(R^+)$. Let us also assume that the distinct diagonal products of A when arranged in increasing order are NOT in G.P. Since $k < 5$ has already been covered there is no loss of generality in assuming $k > 2$. Using Laplace expansion for det A w.r.t the first two rows of A we get

2.66... $\det A = \sum_{j_1 < j_2} \epsilon_{j_1 j_2} \det A[1,2 \mid j_1, j_2] \det A(1,2 \mid j_1, j_2)$

where ϵ_{j_1, j_2} is $(-1)^{j_1+j_2+1}$

$d(A[1, 2 \mid j_1, j_2]) = 1$ clearly implies that $\det A[1,2 \mid j_1, j_2] = 0$. Suppose $d(A[1, 2 \mid j_1, j_2]) = 2$. Then,

$d(A[1, 2 \mid j_1, j_2]) + d(A(1, 2 \mid j_1, j_2)) - 1 \leqslant d(A) \leqslant (n-1)$ by Theorem 2.60. Thus $d(A(1, 2 \mid j_1, j_2)) \leqslant d(A) - 1 \leqslant (n-2)$.

If $d(A(1, 2 \mid j_1, j_2)) = d(A) - 1 = (n-2)$ then by Corollary 2.61 the distinct diagonal products of A arranged in increasing order must be in G.P. contradicting our assumption. Thus $d(A(1,2 \mid j_1, j_2)) \leqslant d(A) - 2 \leqslant n-3$. $A(1, 2 \mid j_1, j_2)$ is of order $(n-2)$ and hence by induction hypothesis it must be singular and hence $\det A(1, 2 \mid j_1, j_2) = 0$. Hence $\det A[1,2 \mid j_1, j_2]$ $\det A(1, 2 \mid j_1, j_2) = 0$ for all $1 \leqslant j_1 < j_2 \leqslant n$. Thus $\det A = 0$ from the Laplace expansion 2.66. Hence A is singular.

Let us now assume that the k distinct diagonal products of A are in G.P. when arranged in increasing order. Consider the set of all n x n positive matrices with exactly $k(\leqslant n-1)$ distinct diagonal products. The condition that these are in G.P. will be that certain continuous functions in the entries of A are equal to zero. Thus the condition that these are NOT in G.P. will be that atleast one of the equalities is not satisfied. The determinant vanishes when atleast one of the equalities is not satisfied. By continuity the determinant will continue to vanish even when all the equalities are satisfied. Thus the theorem must be true even in this case.

This completes the proof of Theorem 2.65 and consequently the conjecture 2.56.

Note that in this proof it is tacitly assumed that the mere condition that there are only $k(< n)$ distinct diagonal products does not <u>force</u> them to be in G.P. In fact we can easily construct a positive matrix of order n with $k\ (< n)$ <u>prescribed</u> positive diagonal products $d_1, d_2, \ldots, d_k$. Towards this end consider an $n \times n$ matrix with the first row consisting of $d_1, d_2, \ldots, d_{k-1}$ and d_k repeated $n-k+1$ times and let all the other rows be filled up by unity. Clearly this satisfies our requirement.

THEOREM 2.66 : If A is an $n \times n$ positive matrix with $d(A) = k \leq (n-1)$, and if the distinct diagonal products are $d_1, d_2, \ldots, d_k$ then each diagonal product d_i occurs an equal number of times with positive and negative signs (this number may be different for different d_i's) in the expansion of det A.

PROOF : Suppose d_i occurs p_i times with positive sign and q_i times with negative sign in det A.

$$\text{Then det } A = \sum_{i=1}^{k} (p_i - q_i) d_i = \sum_{i=1}^{k} n_i \; d_i = 0 \text{ where } n_i = p_i - q_i .$$

Suppose we replace the elements of $A = (a_{ij})$ by their rth powers ($r = 1, 2, 3, \ldots$) and define $A_r = (a_{ij}^r)$. Obviously $d(A_r) = d(A) = k$ ($\leqslant n-1$). Hence det $A_r = \sum n_i \; d_i^r = 0$ for $r = 1, 2, \ldots$ Hence we have

$$\begin{bmatrix} d_1 & d_2 & \cdots & d_k \\ d_1^2 & d_2^2 & \cdots & d_k^2 \\ \vdots & \vdots & \cdots & \vdots \\ d_1^k & d_2^k & \cdots & d_k^k \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The determinant of the coefficient matrix is clearly,

$$d_1 \; d_2 \cdots d_k \begin{vmatrix} 1 & 1 & \cdots & 1 \\ d_1 & d_2 & \cdots & d_k \\ \vdots & \vdots & \cdots & \vdots \\ d_1^{k-1} & d_2^{k-1} & \cdots & d_k^{k-1} \end{vmatrix} \neq 0 \text{ as } d_1 \; d_2 \cdots d_k > 0 \quad \text{and}$$

the determinant is a Vandermonde determinant. Hence it follows that $n_1 = n_2 = \cdots = n_k = 0$. Thus $p_i = q_i$ for $i = 1, 2, \ldots, k$.

ORОLLARY 2.67 : If A is a positive $n \times n$ matrix and $d(A) = k$ ($\leq n-1$) then each diagonal product of A occurs even number of times in the expansion of per A. This follows readily from the fact that

$$\text{per } A = \sum_{i=1}^{k} (p_i + q_i)\, d_i = \sum_{i=1}^{k} 2p_i\, d_i .$$

COROLLARY 2.68 : Suppose A is an $n \times n$ real matrix. Suppose A has exactly k ($\leq n-1$) distinct diagonal sums. Then each diagonal sum d_i occurs the same number of times corresponding to odd and even permutations in S_n (this number may be different for different d_i's).

This is only Theorem 2.66 applied to the matrix $(e^{a_{ij}})$ where $A = (a_{ij})$.

COROLLARY 2.69 : Let A be an $n \times n$ (0, 1) matrix. If per $A = 0$ and A does not have a zero diagonal then each positive diagonal sum occurs an even number of times.

This follows from the fact that the diagonal sums can only be among $\{0, 1, 2, \ldots, n\}$. If per $A = 0$ and A does not have a zero diagonal then the diagonal sums can only be among $\{1, 2, \ldots, n-1\}$. Corollary 2.68 now gives the desired result.

COROLLARY 2.70 : Let A be an $n \times n$ (0, 1) matrix. If the "term rank" of A, i.e., the maximal diagonal sum of A, is atmost $(n-2)$, then each diagonal sum occurs an even number of times.

This is similar to Corollary 2.69.

THEOREM 2.71 : Let A be an m x n positive matrix ($m \leqslant n$). If every m x m submatrix of A has atmost k distinct diagonal products, then $r(A) \leqslant k$.

PROOF : Consider any m x m submatrix of A. By 2.56 its rank is atmost k. Assuming that $k \leqslant m-1$, every (k+1) x (k+1) submatrix of this submatrix is singular. Thus <u>every</u> (k+1) x (k+1) submatrix of A is singular. Thus it follows that $r(A) \leqslant k$.

COROLLARY 2.72 : Suppose A is an m x n real matrix ($m \leqslant n$). Let $S_{m,n}$ denote the set of all one to one functions from (1,2,...,m) into (1, 2, ..., n). For $s \in S_{m,n}$ define the diagonal sum as $\sum_{i=1}^{m} a_{is(i)}$. Suppose A is an m x n (0, 1) matrix. If $\min_s \sum_{i=1}^{m} a_{is(i)} \geqslant 2$, then every diagonal sum of A occurs an even number of times. In particular per (A) is even.

PROOF : The diagonal sums can only be among $\{2, 3, \ldots, m\}$. Hence each diagonal sum occurs an even number of times in each m x m submatrix and hence in the entire matrix A. Per (A) is clearly the number of times the diagonal sum m occurs. Hence per (A) is even.

OROLLARY 2.73 : Suppose A is an m x n (0, 1) matrix ($m \leqslant n$).
f per (A) is odd then $\min_{s} \sum_{i=1}^{m} a_{is(i)} < 2$.

PROOF : This follows readily from Corollary 2.72. This result is uite strange. The mere restriction that per A is odd seems to ave a strangle hold on the minimal diagonal sum of A. In partilar this means that the matrix A must be quite rich in zeros atleast (m-1) zeros).

C H A P T E R - III

MUIR ALGEBRA AND ITS APPLICATIONS

3.0 INTRODUCTION :

T. Muir [13] introduced a symbolic method for expressing permanent of a square matrix as a coefficient of a term in a product. He illustrated the calculation of permanent of a 3 x 3 matrix $A=(a_{ij})$ as follows.

Let x_1, x_2, x_3 be symbols satisfying the properties $x_1 x_2 = x_2 x_1, x_1 x_3 = x_3 x_1, x_2 x_3 = x_3 x_2, x_1^2 = x_2^2 = 0 = x_3^2$. Then per A is the coefficient of $x_1 x_2 x_3$ in the following formal product.

$$(a_{11} x_1 + a_{12} x_2 + a_{13} x_3)(a_{21}x_1 + a_{22}x_2 + a_{23}x_3)(a_{31}x_1 + a_{32}x_2 + a_{33}x_3)$$

He also added that this was just the definition of permanent only.

In this chapter we will develop an algebra in which Muir's ideas are incorporated. It turns out that the algebra so developed gives all the existing formulas for evaluation of permanents and points out a general method of getting new formulas. In addition the algebra seems to be a very powerful tool in combinatorics. We formally develop old and new formulas for evaluation of permanents, formula of inclusion and exclusion and generating function for partition function all though Muir algebra.

3.1 MUIR ALGEBRA :

Let M_n be an associative, commutative algebra over the complex field, of dimension 2^n, with a basis consisting of $1, e_1, e_2, \ldots, e_n, e_1 e_2, e_1 e_3, \ldots, e_{n-1} e_n, e_1 e_2 e_3, \ldots,$ $e_{n-2} e_{n-1} e_n, \ldots, e_1 e_2 \ldots, e_n$ where 1 is the unit element and e_i's have the property $e_i^2 = 0$, the zero element of M_n for $i = 1, 2, \ldots, n$. Commutativity and the property $e_i^2 = 0$ give the complete multiplicative structure of M_n.

M_n can also be thought of in another manner. Let F_n be the free commutative, associative algebra over the complex field generated by n indeterminates $a_1, a_2, \ldots, a_n$. Let G_n be the subalgebra, which is also an ideal, consisting of all the linear combinations of products of a_i's with atleast one a_i occurring with a power greater than one. For example $2 a_1^2 a_2 a_3 + 4 a_1 a_2^3 a_3 + 3 a_1 a_3 a_4^5$ is a typical element of G_n. Consider the algebra F_n/G_n obtained from F_n by "factoring out" G_n. Let e_i be the image of a_i under the natural homomorphism from F_n onto F_n/G_n. Then F_n/G_n with $1, e_1, e_2, \ldots, e_n$ as the generators will be the Muir algebra M_n. M_n, considered as a vector space of dimension 2^n can be considered as a unitary space of dimension 2^n with the usual innerproduct written as (x, y) for $x, y \in M_n$.

3.2 APPLICATION OF MUIR ALGEBRA TO PERMANENTS :

Let $A = (a_{ij}) \in M_{m,n}(C)$ with $m \leqslant n$. Let $S_{m,n}$ be the set of all one to one functions from $\{1,2, \ldots, m\}$ into $\{1,2,\ldots,n\}$. Then we define permanent of A by

3.3... $$\text{per}(A) = \sum_{s \in S_{m,n}} \prod_{i=1}^{m} a_{i,s(i)} \cdot$$

Suppose $\theta = \sum_{i=1}^{n} e_i \in M_n$; then it is easy to see that

3.4... $$\theta^r = r! \sum_{i_1 < i_2 < \cdots < i_r} e_{i_1} e_{i_2}, \ldots, e_{i_r} \quad \text{for } r = 1,2,\ldots,n$$

Given $A = (a_{ij}) \in M_{m,n}(C)$ define $r_i = \sum_{j=1}^{n} a_{ij} e_j \in M_n$ for $i=1,2,\ldots,m$.

Hence $\left(\prod_{i=1}^{m} r_i, \theta^m \right) = m! \text{ per } (A)$ or

3.5... $$\text{per } (A) = \left(\prod_{i=1}^{m} r_i, \frac{\theta^m}{m!} \right)$$

Suppose $\underline{a} = (a_1, a_2, \ldots, a_n)$. Let us denote by $p_r(\underline{a})$ the sum of the products of $a_1, a_2, \ldots, a_n$ taken r at a time. In otherwords $p_r(\underline{a})$ is the rth elementary symmetric function of $a_1, a_2, \ldots, a_n$. With this notation, if $E = (e_1, e_2, \ldots, e_n)$ and $\theta = e_1+e_2+ \ldots e_n$, then $\theta^r = r! \, p_r(E)$ for $r = 1, 2, \ldots, n$. If $R = (r_1, r_2, \ldots, r_m)$ where $r_i = \sum_{j=1}^{n} a_{ij} e_j \in M_n$, then $r_1 r_2 \ldots r_m = p_m(R)$. Thus we get

3.6... $$\text{per } (A) = (p_m(R), p_m(E)).$$

Let us now develop a formula for permanents using columns of A.

Let $A = (a_{ij}) \in M_{m,n}(C)$ with $m \leq n$.

If $C_j = \sum_{i=1}^{m} a_{ij}\, e_i \in M_m$ for $j = 1, 2, \ldots, n$ and $C = (c_1, c_2, \ldots, c_n)$

then we get readily the formula.

3.7... $\quad per(A) = (p_m(C),\ p_m(E))$ if $E = (e_1, e_2, \ldots, e_m)$.

Formulas 3.6 and 3.7 treat rows and columns of A in a symmetric manner (A is not necessarily a square matrix.)

LEMMA 3.8 : Let N be a finite set and x_i, $i \in N$ be indeterminates in a field. If 2^N is the power set of N we have the identity.

$$F(x_1, x_2, \ldots, x_n) = \sum_{S \in 2^N} (-1)^{|N|-|S|} \Big(\sum_{i \in S} x_i\Big)^r = \begin{cases} 0 & \text{if } r < |N| \\ r! \prod_{i \in N} x_i & \text{if } r = |N| \end{cases}$$

PROOF : Let us prove the result for $|N| = 3$, the general case being similar. Take $N = \{1, 2, 3\}$.

$$F = (x_1+x_2+x_3)^r - (x_1+x_2)^r - (x_2+x_3)^r - (x_1+x_3)^r + x_1^r + x_2^r + x_3^r .$$

Putting $x_1 = 0$ clearly $F = 0$, various terms cancelling in pairs. Thus x_1 is a factor of F. Thus $x_1\, x_2\, x_3$ must be a factor of F. Hence if $r < 3$, F must vanish identically. If $r = 3$, then $F = K\, x_1\, x_2\, x_3$ for a constant K. Setting $x_1 = x_2 = x_3 = 1$,

$K = 3^3 - 3.2^3 + 3.1^3 = 3!$ (the general result being $n^n - \binom{n}{1}(n-1)^n + \binom{n}{2}(n-2)^n \ldots = n!$). This proves Lemma 3.8.

Clearly Lemma 3.8 is also true when x_i's are in any commutative ring.

3.9 RYSER'S FORMULA FOR Per(A) :

Let $A = (a_{ij}) \in M_{m,n}(C)$ By 3.7, $\text{per}(A) = (p_m(C), p_m(E))$.

Let N_m denote the set of all m subsets of $N = \{1, 2, \ldots, n\}$.

Then $p_m(C) = \sum_{S \in N_m} \prod_{i \in S} C_i$. Hence, using Lemma 3.8 we get

$$3.10\ldots \quad m!\, p_m(C) = \sum_{S \in N_m} m! \prod_{i \in S} C_i = \sum_{S \in N_m} \sum_{B \in 2^S} (-1)^{m-|B|} \Big(\sum_{i \in B} C_i\Big)^m.$$

where $\sum_{i \in \emptyset} C_i$ is interpreted as 0.

But an s-set of an m-set of an n-set occurs $\binom{n-s}{m-s}$ times as an s-set of the n-set. Hence 3.10 gives

$$m!\, p_m(C) = \sum_{B \in 2^N} (-1)^{m-|B|} \binom{n-|B|}{m-|B|} \Big(\sum_{i \in B} C_i \Big)^m$$

But $\sum_{i \in B} C_i = \sum_{i=1}^{m} R_i(B)\, e_i$, where $R_i(B)$ is the ith row sum of the matrix formed by the columns of A with column indices in B.

Hence $\Big(\sum_{i \in B} C_i\Big)^m = m!\, \Big(\prod_{i=1}^{m} R_i(B)\Big).\ p_m(E)$.

Thus $per(A) = (p_m(C),\ p_m(E)) = (\sum_{B\in 2^N} (-1)^{m-|B|} \binom{n-|B|}{m-|B|} \prod_{i=1}^{m} R_i(B).p_m(E), p_m(E))$

Thus we get the formula.

$$3.11\ldots\quad per\ (A) = \sum_{B\in 2^N} (-1)^{m-|B|} \binom{n-|B|}{m-|B|} \prod_{i=1}^{m} R_i(B)\ .$$

This is Ryser's formula for permanent. We now present a similar but new formula for permanent in terms of column-sums.

3.12 A FORMULA FOR THE PERMANENT OF AN $m \times n$ MATRIX IN TERMS OF COLUMN-SUMS :

Suppose $A = (a_{ij}) \in M_{m,n}(C)$ with $m \leqslant n$.

Let $r_i = \sum_{j=1}^{n} a_{ij}\ e_j \in M_n$ for $i = 1, 2, \ldots, m$. Let $M = (1,2,\ldots,m)$. Then $per\ (A) = (p_m(R),\ p_m(E))$ where $R = (r_1, r_2, \ldots, r_m)$ and $E = (e_1, e_2, \ldots, e_n)$.

$$\text{But } m!\ p_m(R) = m! \prod_{i=1}^{m} r_i = \sum_{B\in 2^M} (-1)^{m-|B|} \left(\sum_{i\in B} r_i\right)^m .$$

But $\sum_{i\in B} r_i = \sum_{j=1}^{n} C_j(B)\ e_j$ where $C_j(B)$ is the jth column-sum of the matrix formed by the rows of A whose indices are in B.

$(\sum_{i\in B} r_i)^m = (\sum_{j=1}^{n} C_j(B)\ e_j)^m$ and hence we get

$$per(A) = (\sum_{B\in 2^M} (-1)^{m-|B|} (\sum_{j=1}^{n} C_j(B)\ e_j)^m,\ p_m(E))$$
$$= \sum_{B\in 2^M} (-1)^{m-|B|}\ p_m(C(B)) \text{ where } C(B) = (C_1(B), C_2(B), \ldots, C_n(B))$$

Thus we get the formula

3.13... $\text{per}(A) = \sum_{B \in 2^M} (-1)^{m-|B|} p_m(C(B))$

COROLLARY 3.14 : If m=n, 3.13 gives

$$\text{Per}(A) = \sum_{B \in 2^M} (-1)^{m-|B|} \prod_{j=1}^{m} C_j(B) \text{ and 3.11 gives}$$

$$\text{Per}(A) = \sum_{B \in 2^N} (-1)^{n-|B|} \prod_{i=1}^{n} R_i(B)$$

We now derive an interesting formula from Lemma 3.8.

Let $r < m \leq n$. Then $\sum_{B \in 2^M} (-1)^{m-|B|} (\sum_{i \in B} r_i)^r = 0$ or

$(\sum_{B \in 2^M} (-1)^{m-|B|} (\sum_{j=1}^{n} C_j(B)\, e_j)^r, p_r(E)) = 0$ clearly this reduces to

$\sum_{B \in 2^M} (-1)^{m-|B|} p_r(C(B)) = 0$ where $C(B) = (C_1(B), C_2(B), C_2(B), \ldots, C_n(B))$.

Thus we have proved the following theorem.

THEOREM 3.15 : Suppose A is an $m \times n$ matrix with $m \leq n$. Then

$$\sum_{B \in 2^M} (-1)^{m-|B|} p_r(C(B)) = \begin{cases} \text{Per}(A) & \text{if } r = m \\ 0 & \text{if } 1 \leq r < m. \end{cases}$$

A similar theorem can be proved for row-sums. We are now ready to prove a very interesting theorem.

THEOREM 3.16 : Suppose A is an $n \times n$ matrix with column sums $C_1(B)$, $C_2(B)$, ..., $C_n(B)$ for a submatrix of A with rows of A whose indices are in $B \in 2^N$. Then for any n complex numbers $u_1, u_2, \ldots, u_n$

3.17... $\text{per}(A) = \sum_{B \in 2^N} (-1)^{n-|B|} \; p_n(C(B)-u)$ where $C(B)-u = (C_1(B)-u_1, \ldots, C_n(B)-u_n)$.

PROOF : Let $1 \quad r \quad n$. Then the coefficient of $u_1 \cdot u_2 \cdot \ldots \cdot u_r$ in $\sum_{B \in 2^N} (-1)^{n-|B|} \; p_n(C(B)-u) = (-1)^r \sum_{B \in 2^N} (-1)^{n-|B|} \; p_{n-r}(0,0,\ldots,0, C_{r+1}(B), C_{r+2}(B), \ldots, C_n(B))$ where there are r zeros in the last vector. But by Theorem 3.15 this is zero. On the otherhand if $r = 0$, $(-1)^r \sum_{B \in 2^N} (-1)^{n-|B|} \; p_n(C(B)) = \text{per}\, A$ and the theorem follows.

COROLLARY 3.18 : If $u_1, u_2, \ldots, u_n$ are such that for every $B \in 2^N$ there exists i such that $C_i(B) = 0$ then $\text{per}(A) = u_1 \cdot u_2 \cdot \ldots \cdot u_n$.

In particular if $A = J_n = (\frac{1}{n})$ then we can take $(u_1, u_2, \ldots, u_n) = (\frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \ldots, \frac{n}{n})$. Then $\text{per}(A) = n! / n^n$.

If A is a permutation matrix we can take $u_i = 1$ for each i. Then $\text{per}(A) = 1$.

We now pass on to an interesting lemma from which a wholly new formula for permanent will be derived.

LEMMA 3.19 : Let $x_1, x_2, \ldots, x_n$ be n indeterminates in a commutative ring. Then we have the identity.

3.20... $F = \sum_{B \in 2^N} (-1)^{n-|B|} \left(\sum_{i \in B} x_i - \sum_{i \in B^c} x_i \right)^n = 2^n \, n! \, x_1 . x_2 . \ldots . x_n .$

PROOF : Proof of this is similar to that of Lemma 3.8. A typical proof of 3.20 can be illustrated by taking $n = 3$. Let us assume that the commutative ring is a field.

$$F = (x_1+x_2+x_3)^3 - (x_1+x_2-x_3)^3 - (x_1-x_2+x_3)^3 - (-x_1+x_2+x_3)^3 + (x_1-x_2-x_3)^3 + (-x_1+x_2-x_3)^3 + (-x_1-x_2+x_3)^3 - (-x_1-x_2-x_3)^3.$$

Clearly $F = 0$ when $x_1 = 0$ as in R.H.S. terms cancel in pairs. Thus $F = k\, x_1\, x_2\, x_3$ for a constant k.

Taking $x_1 = x_2 = x_3 = 1$ we get,

$k = 3^3 - \binom{3}{1} (1)^3 + \binom{3}{2} (-1)^3 - (-3)^3 = 48 = 2^3 . 3!$ In the general case

$k = n^n - \binom{n}{1} (n-2)^n + \binom{n}{2} (n-4)^n - \ldots + (-1)^n (n-2n)^n$. To simplify this consider $f(x) = (x+n)^n - \binom{n}{1} (x+n-2)^n + \binom{n}{2} (x+n-4)^n + \ldots + (-1)^n (x+n-2n)^n$. $f(x) = (x+n)^n - \binom{n}{1} E^{-2} (x+n)^n + \binom{n}{2} E^{-4} (x+n)^n \ldots$ where E is the shift operator defined by $E\, f(x) = f(x+1)$.

Thus $f(x) = (1-E^{-2})^n (x+n)^n = (E^2-1)^n (x-n)^n$. Taking $\Delta = E-1$,

$f(x) = (\Delta^2 + 2\Delta)^n (x-n)^n = (\Delta+2)^n \Delta^n (x-n)^n = (\Delta+2)^n n! = 2^n . n!$

by using well known properties of the difference operator Δ. Putting $x = 0$ we get $k = 2^n . n!$ and this completes the proof of the lemma.

COROLLARY 3.21 : If $n = 2r+1$, $(-1)^{n-|B|} \left(\sum_{i\in B} x_i - \sum_{i\in B^c} x_i\right)^n$

$$= (-1)^{2r+1-|B|} \left(\sum_{i\in B} x_i - \sum_{i\in B^c} x_i\right)^{2r+1} = (-1)^{2r-|B|} \left(\sum_{i\in B^c} x_i - \sum_{i\in B} x_i\right)^{2r+1}$$

$= (-1)^{2r+1-|B^c|} \left(\sum_{i\in B^c} x_i - \sum_{i\in B} x_i\right)^{2r+1}$. Hence taking one half of the terms in Lemma 3.19 we get

$$3.22\ldots \quad \sum_{\substack{B\in 2^N \\ |B|\leq r}} (-1)^{2r+1-|B|} \left(\sum_{i\in B} x_i - \sum_{i\in B^c} x_i\right)^{2r+1} = 2^{2r}(2r+1)!\cdot x_1\cdot x_2\cdot\ldots\cdot x_{2r+1}.$$

If $n = 2r$, in a similar manner we get

$$3.23\ldots \quad \sum_{\substack{B\in 2^N \\ |B|\leq r-1}} (-1)^{2r-|B|} \left(\sum_{i\in B} x_i - \sum_{i\in B^c} x_i\right)^{2r} + \tfrac{1}{2} \sum_{\substack{B\in 2^N \\ |B|=r}} (-1)^r \cdot \left(\sum_{i\in B} x_i - \sum_{i\in B^c} x_i\right)^{2r} = 2^{2r-1}(2r!)x_1\ldots x_{2r}.$$

Note that, in $\sum_{\substack{B\in 2^N \\ |B|=r}} \left(\sum_{i\in B} x_i - \sum_{i\in B^c} x_i\right)^{2r}$ there are an even number of terms which can be arranged in two groups of EQUAL terms.

We are now ready to prove an important modification of Ryser's formula for permanent of square matrices which cuts the calculation by half.

THEOREM 3.24 : Let $A = (a_{ij})$ be an $n \times n$ matrix. Let $C_1, C_2, \ldots, C_n$ be the column sums of A. Let $C_i(B)$ be the ith column-sum of the submatrix of A with rows whose indices are in $B \in 2^N$.

If $r_i = \sum_{j=1}^{n} a_{ij} e_j \in M_n$, then for $B \subseteq 2^N$,

$$\left(\sum_{i \in B} r_i - \sum_{i \in B^c} r_i \right) = \sum_{j=1}^{n} (C_j(B) - C_j(B^c)) e_j = \sum_{j=1}^{n} (2C_j(B) - C_j) e_j$$

for $C_j(B) + C_j(B^c) = C_j$.

CASE-1 : $n = 2r+1$. Then $2^{2r}(2r+1)!\ \mathrm{per}\ (A) = (2^{2r}(2r+1)!\, p_{2r+1}(A),$

$$p_{2r+1}(E)) = \Big(\sum_{\substack{B \in 2^N \\ |B| \leq r}} (-1)^{2r+1-|B|} \Big(\sum_{i \in B} r_i - \sum_{i \in B^c} r_i \Big)^{2r+1},\ p_{2r+1}(E) \Big)$$

$$= \Big(\sum_{\substack{B \in 2^N \\ |B| \leq r}} (-1)^{2r+1-|B|} (2r+1)!\ p_{2r+1}\ (2C(B)-C) \text{ where } C = (C_1, C_2, \ldots, C_n)$$

Thus we get,

$$3.25\ldots \quad 2^{2r}\ \mathrm{per}\ (A) = \sum_{\substack{B \in 2^N \\ |B| \leq r}} (-1)^{2r+1-|B|} \prod_{j=1}^{2r+1} (2\, C_j(B) - C_j)$$

CASE-2 : $n = 2r$. This is similar to Case 1 but for a small difference. We easily get the following.

$$3.26\ldots \quad 2^{2r-1}\ \mathrm{per}\ (A) = \sum_{\substack{B \in 2^N \\ |B| < r}} (-1)^{2r-|B|} \prod_{j=1}^{2r} (2C_j(B) - C_j) + \frac{1}{2} \sum_{\substack{B \in 2^N \\ |B| = r}} (-1)^r \prod_{j=1}^{2r} (2\, C_j(B) - C_j).$$

Formulas 3.25 and 3.26 involve only $B \in P_N$ for which $|B| \leq n/2$ and hence, compared to Ryser's formula, these involve less than half the amount of calculation.

COROLLARY 3.27 : Suppose A is a square matrix with integral entries. If all the column sums are even, then per (A) is even. For $n = 2r+1$, in 3.25, R.H.S has 2^{2r+1} as a factor and L.H.S. is 2^{2r} per (A) and the result follows. For $n = 2r$ we use 3.26 and the fact that $\frac{1}{2} \sum_{\substack{B \in 2^N \\ |B| = r}} (-1)^r \prod_{j=1}^{2r} (2\, C_j(B) - C_j)$ can be written as a similar sum without the factor $\frac{1}{2}$ by taking only one half of the identical pain in the summation. Thus again the result follows.

COROLLARY 3.28 : Suppose A is a square, integral matrix whose rows can be partitioned into (k+1) sets of rows in k of which all the column sums are even. Then per (A) is a multiple of 2^k.

This follows readily from Corollary 3.27 and the general Laplace expansion of per (A) in terms of the sets of rows.

COROLLARY 3.29 : Suppose G_m^r is the collection of all m x m, (0, 1) matrices with each row-sum and column-sum equal to r. If $A \in G_m^{2r}$, then per (A) is even.

This follows readily from Corollary 3.27. In this connection the following conjecture seems to be highly plausible.

CONJECTURE 3.40 : If A is an n x n integral matrix such that any r x n submatrix of A has atleast one column sum odd for r=1,2,...,n then per (A) is odd. In particular per (A) $\neq$ 0.

Since per (A) and det (A) have the same parity this also means that det (A) is odd and in particular det (A) $\neq 0$.

Note that the converse of this conjecture viz.

If per (A) is odd then every $r \times n$ submatrix of A has atleast one column sum odd for $r = 1, 2, \ldots, n$ is certainly true in view of Corollary 2.

T The above result is a bit surprising. It seems that, for large n, it will be rather difficult to construct non-trivial $n \times n$ integral matrices with odd permanent (or determinant). From the results so far derived using Muir algebra the following general principle emerges.

3.41 A GENERAL PRINCIPLE :

From any algebraic expression for $x_1 x_2 \ldots x_n$ in terms of polynomials in $x_1, x_2, \ldots, x_n$ we can extract a formula for permanents.

ILLUSTRATION 3.42 :

$4\, gh\, x_1 x_2 \equiv (g x_1 + h x_2)^2 - (g x_1 - h x_2)^2$ is an identity. From this we get a formula for permanent of $2 \times n$ matrices.

$4\, gh \text{ per } (A) = (4\, gh\, r_1 r_2, p_2(E)) = ((gr_1+hr_2)^2, p_2(E)) - ((gr_1-hr_2)^2, p_2(E))$. Hence 4 gh per (A) = per (A_1)+per A_2 where

A_1 (A_2) is a 2 x n matrix whose identical rows are equal to g times the first row of A plus (minus) h times the second row of A.

ILLUSTRATION 3.43 :

We know that if $x_1 + x_2 + x_3 = 0$, then $x_1^3 + x_2^3 + x_3^3 = 3x_1\, x_2\, x_3$. Taking $x_1 = r_1$, x_2, $x_3 = r_3$, for a 3 x n matrix A, $r_1 + r_2 + r_3 = 0$ if and only if each column sum is zero. Then

$$3 \text{ per } (A) = (3r_1\, r_2\, r_3,\ p_3(E)) = \sum_{i=1}^{3}(r_i^3\ ,\ p_3(E)) = \sum_{i=1}^{3} \text{per } A_i$$

where A_i is the 3 x n matrix each of whose rows is the ith row of A.

Eventhough illustration 3.42 is quite trivial, 3.43 is non-trivial and interesting.

3.44 APPLICATION TO COMBINATORICS - INCLUSION - EXCLUSION FORMULA :

Let $U_1, U_2, \ldots, U_N$ be N objects with weights $W_1, W_2, \ldots, W_N$ (W_i's are elements of a commutative ring). Let $P_1, P_2, \ldots, P_n$ be attributes concerning these objects. Let $A = (a_{ij})$, an N x n matrix be defined by $a_{ij} = 1$ if U_i has the attribute P_j and 0 otherwise. Let $r_i = \sum_{j=1}^{n} a_{ij}\, e_j \in M_n$, i = 1, 2, ..., N. Let $e = \sum_{j=1}^{n} e_j$.

Let N = (1, 2, ..., n). For $B \in 2^N$ let $W(B)$ be the sum of the weights of U_i's possessing all the attributes P_j for $j \in B$. Let $S_r = \sum_{\substack{B \in 2^N \\ |B| = r}} W(B)$. Clearly $W(B) = \sum_{i=1}^{N} W_i((\sum_{j \in B} a_{ij} e_j)^r/r!,\ p_r(E))$

where $r = |B|$

3.45... Hence $S_r = \sum_{i=1}^{N} W_i \left(\frac{r_i^r}{r!}, p_r(E) \right)$.

Clearly U_i has exactly k properties if and only if $r_i^k (\theta - r_i)^{n-k} \neq 0$ in which case $\frac{r_i^k (\theta - r_i)^{n-k}}{k!\,(n-k)!} = p_n(E)$. Thus if W_k represents the sum of weights of U_i's having exactly k properties, then

3.46... $\overline{W}_k = \sum W_i \left(\frac{r_i^k (\theta - r_i)^{n-k}}{k!\,(n-k)!}, p_n(E) \right)$. Hence we get

$$\sum_{k=0}^{n} \overline{W}_k x^k = \sum_{i=1}^{N} W_i \left(\sum_{k=0}^{n} \frac{x^k r_i^k (\theta - r_i)^{n-k}}{k!\,(n-k)!}, p_n(E) \right) = \sum_{i=1}^{N} W_i \left(e^{x r_i} e^{\theta - r_i}, p_n(E) \right)$$

$$= \sum_{i=1}^{N} W_i \left(e^{\theta} \cdot e^{(x-1) r_i}, p_n(E) \right) = \sum_{i=1}^{N} W_i \left(\sum_{r=0}^{n} \frac{\theta^{n-r}}{(n-r)!} \frac{r_i^r}{r!} (x-1)^r, p_n(E \right.$$

$$= \sum_{i=1}^{N} W_i \left(\sum_{r=0}^{n} (x-1)^r \frac{r_i^r}{r!}, p_r(E) \right) = \sum_{r=0}^{n} (x-1)^r \left(\sum_{i=1}^{N} W_i \left(\frac{r_i^r}{r!}, p_r(E) \right) \right)$$

$$= \sum_{r=0}^{n} (x-1)^r S_r \quad \text{using 3.45.}$$

Thus $\sum_{k=0}^{n} \overline{W}_k x^k = \sum_{r=0}^{n} (x-1)^r S_r$

Equating coefficient, of x^k on both sides we get

3.47... $$\overline{W}_k = \sum_{r=k}^{n} (-1)^{r-k} \binom{r}{k} S_r$$

In particular $\overline{W}_0 = \sum_{r=0}^{n} (-1)^r S_r$

Putting $x = y+1$, $\sum_{k=0}^{n} \overline{W}_k (y+1)^k = \sum_{r=0}^{n} y^r S_r$

Equating coefficient of y^r on both sides we get

3.48... $$S_r = \sum_{k=r}^{n} \binom{k}{r} \overline{W}_k .$$

Putting $x = 1/y$ and multiplying both sides by y^n we get

$$\sum_{k=0}^{n} \overline{W}_k y^{n-k} = \sum_{r=0}^{n} (1-y)^r y^{n-r} S_r .$$

If $\overline{W}_{\geq k} = \sum_{r=k}^{n} \overline{W}_r$ clearly $\sum_{k=0}^{n} \overline{W}_k y^{n-k} (1-y)^{-1} = \sum_{k=0}^{n} \overline{W}_{\geq k} y^{n-k}$

Thus $\sum_{k=0}^{n} \overline{W}_{\geq k} y^{n-k} = \sum_{r=0}^{n} (1-y)^{r-1} y^{n-r} S_r$

Equating coefficients of y^{n-k} on both sides,

3.49... $$\overline{W}_{\geq k} = \sum_{r=k}^{n} (-1)^{r-k} \binom{r-1}{r-k} S_r = \sum_{r=k-1}^{n-1} (-1)^{r-k+1} \binom{r}{r-k+1} S_{r+1}$$

or equivalently $\overline{W}_{\geq k} = \sum_{r=0}^{n-k} (-1)^r \binom{r+k-1}{r} S_{r+k}$.

The main result $\sum_{k=0}^{n} \overline{W}_k x^k = \sum_{r=0}^{n} (x-1)^r S_r$ is got elegantly using Muir algebra.

3.50 AN INTERESTING OBSERVATION :

The expressions $(r_i^r , p_r(E))$ and $(r_i^k(\theta - r_i)^{n-k}, p_n(E))$ used in the section are merely expressions for permanents of suitable matrices. Thus it is clearly possible to give a proof of the above results using only permanents. Ryser gave a proof of the formula for permanents using the principle of inclusion - exclusion. Here the principle of inclusion - exclusion can be proved using permanents ! But the connecting link is Muir algebra. Thus Muir algebra seems to be more basic. The interplay of Muir algebra with various combinatorial structures is indeed striking. As an example of this let us derive the generating function for the partition function T_n.

3.51 PARTITION FUNCTION :

Consider an n-set with elements $a_1, a_2, \ldots, a_n$. Let T_n represent the number of ways of partitioning the n-set. Consider the Muir algebra M_n where $\theta = e_1 + e_2 + \ldots + e_n$.

Let $\phi = \theta + \frac{\theta^2}{2!} + \frac{\theta^3}{3!} + \ldots = e^{\theta} - 1$ (this is a finite sum)

Consider the formal expression,

$$\psi = e_1 x^{a_1} + e_2 x^{a_2} + \ldots + e_n x^{a_n} + e_1 e_2 x^{a_1 a_2} + \ldots + e_{n-1} e_n x^{a_{n-1} a_n} + \ldots + e_1 e_2 \ldots e_n x^{a_1 a_2 \ldots a_n} .$$

Then the "coefficient" of $a_1 a_2 \dots a_n$ in $\frac{\Psi^r}{r!}$ gives all the partitions of $a_1, a_2, \dots, a_n$ into r parts in the exponents of x. For example, if n = 4 the partition of $\{a_1, a_2, a_3, a_4\}$ (say) $\{a_1, a_2\}$, $\{a_3, a_4\}$ will appear in $\frac{\Psi^2}{2!}$ as $x^{a_1 a_2 + a_3 a_4}$.

As we are interested only in the number of partitions we can as well take

$$= (a_1 + a_2 + \dots + a_n) + (a_1 a_2 + \dots + a_{n-1} a_n) + \dots + a_1 a_2 \dots a_n$$

$$= \theta + \frac{\theta^2}{2!} + \frac{\theta^3}{3!} + \dots + \frac{\theta^n}{n!} = \phi$$

Thus if $T_n(r)$ is the number of partitions of an n-set into r parts, then $T_n(r) = (\frac{\phi^r}{r!} , p_n (E))$.

Thus $T_n = \sum_{r=1}^{n} T_n(r) = (\sum_{r=1}^{n} \frac{\phi^r}{r!} , p_n (E))$ and hence

3.52... $$T_n = (e^{e^{\theta}-1}, p_n(E)).$$

Suppose $e^{(e^x-1)}-1 = \sum_{r=1}^{\infty} u_r \frac{x^r}{r!}$. Then $T_n = (\sum_{r=1}^{\infty} u_r \frac{\theta^r}{r!} , p_n(E))$

But $(\frac{\theta^r}{r!} , p_n (E)) = \delta_{r,n}$, the Kronecker delta. Thus $T_n = u_n$

which means $\sum_{n=1}^{\infty} T_n \frac{x^n}{n!} = e^{(e^x-1)}-1$, the generating function for T_n.

Also $T_n(r) = (\frac{\emptyset^r}{r!} , p_n(E)) = (\frac{(e^{\theta}-1)^r}{r!} , p_n(E))$

$$= (\frac{1}{r!} \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} e^{\theta k} , p_n(E))$$

$$= \frac{1}{r!} \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} k^n (\frac{\theta^n}{n!} , p_n(E))$$

$$= \frac{1}{r!} \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} k^n .$$

Thus we have proved

3.53... $T_n(r) = \frac{1}{r!} \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} k^n.$

As a corollary, taking r = 1, we get

$$1 = T_n(n) = \frac{1}{n!} \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} k^n \qquad \text{or}$$

3.54... $\sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} k^n = n!$, a well-known result.

C H A P T E R - IV

PERMANENTS AND VANDER WAERDEN CONJECTURE

***

4.0 INTRODUCTION :

Permanents were introduced into mathematics by Cauchy. But permanents are not as well-behaved as determinants. But permanents are very useful in combinatorics where permutation of objects in restricted positions are considered. Due to the difficulty in evaluation of permanents the progress in the study of permanents was very slow. But interest in permanents was revived to a large extent by the famous Van der Waerden Conjecture [1926] on doubly stochastic matrices. This conjecture is still unresolved.

The Chapter arose out of author's unsuccessful attempt at settling the conjecture. But quite a number of interesting results came out of this attempt. This chapter mainly deals with results so obtained. In the beginning a characterisation of doubly stochastic matrices is given. This is applied to latin-squares. Then an important consequence of the result in permanents is derived. Then a number of conjectures more powerful than Vander Waerden's viz., Tverberg's and Djokovic are dealt with. Then permanent is related to multinomial distribution is statistics. Then a new class of doubly stochastic matrices for which vander Waerden conjecture is true is developed.

4.1 CHARACTERIZATION OF DOUBLY STOCHASTIC MATRICES :

Let A be an $n \times n$ real matrix. For $1 \leqslant r, s < n$ let B be an $r \times s$ submatrix of A. Let $\bar{B}$ represent the submatrix of A obtained by omitting the rows and columns in which B lies. For any matrix C let $\sigma(C)$ represent the sum of all the entries in C.

THEOREM 4.2 : Let A be an $n \times n$ non-negative matrix with $\sigma(A) = n$. Then A is doubly stochastic if and only if for some r, s $(1 \leqslant r, s < n)$ all the $r \times s$ submatrices B of A satisfy the condition

$$\sigma(\bar{B}) = \sigma(B) + (n-r-s)$$

PROOF : The condition is necessary. To prove this we may assume without loss of generality that B lies in the top left corner of A as shown in the figure

$$A = \left[\begin{array}{c|c} B & X \\ \hline Y & \bar{B} \end{array}\right]$$

$$\sigma(B) + \sigma(X) = r \qquad \text{for } A \text{ is doubly-stochastic.}$$

$$\sigma(\bar{B}) + \sigma(X) = n-s$$

By subtraction we get $\sigma(\bar{B}) = \sigma(B) + (n-r-s)$.

The condition is sufficient. Let $\sigma(A) = n$, $\sigma(\bar{B}) = \sigma(B) + (n-r-s)$. Let $R_1, R_2, \ldots, R_n$ be the row sums of A.

Consider $\sum \sigma(B)$ where the summation is over all r x s submatrices B of A lying in the first r rows of A.

$$\sum \sigma(B) = \binom{n-1}{s-1} (R_1 + R_2 + \dots + R_r)$$

$$\sum \sigma(\bar{B}) = \binom{n-1}{n-s-1} (R_{r+1} + R_{r+2} + \dots + R_n)$$

$$= \binom{n-1}{s} (n-R_1 - R_2 - \dots - R_r) \text{ for } R_1 + R_2 + \dots + R_n = \sigma(A) = n.$$

But $\sigma(\bar{B}) = \sigma(B) + (n-r-s)$. Thus $\sum \sigma(\bar{B}) = \sum \sigma(B) + \binom{n}{s}(n-r-s)$

Thus $\binom{n-1}{s} (n-R_1-R_2- \dots - R_r) = \binom{n-1}{s-1}(R_1+R_2+ \dots + R_r)+\binom{n}{s}(n-r-s)$

or $\left[\binom{n-1}{s} + \binom{n-1}{s-1}\right] (R_1 + R_2 + \dots + R_r) = n \binom{n-1}{s} - \binom{n}{s}(n-r-s)$

or $\binom{n}{s} (R_1 + R_2 + \dots + R_r) = (n-s) \binom{n}{s} - (n-r-s)\binom{n}{s} = r\binom{n}{s}$

Thus $R_1 + R_2 + \dots + R_r = r.$

But the same argument on <u>any</u> r rows of A shows that the sum of the rowsums of any r rows of A is r. This is possible only if each row sum is 1. In a similar manner considering s columns of A we can show that each column sum is also 1. Since A is a non-negative matrix A must be doubly-stochastic.

Now we prove an important theorem with interesting applications.

THEOREM 4.3 : A set of necessary and sufficient conditions that an $r \times s$ matrix B is a submatrix of an $n \times n$ doubly stochastic matrix with $1 \leqslant r, s < n$ is

(1) all row sums and column sums of B are at most 1

(2) $\sigma(B) \geqslant r + s - n$.

PROOF : Condition (1) is trivially necessary, condition (2) is necessary, for $\sigma(\bar{B}) = \sigma(B) + (n-r-s) \geqslant 0$ in view of Theorem 4.2.

Let us prove the sufficiency. Let us construct a doubly stochastic $n \times n$ matrix A with B on the top left corner. Take $A = \begin{bmatrix} B & X \\ Y & Z \end{bmatrix}$ where X, Y and Z are constructed as follows.

Let the row sums of B be $R_1, R_2, \ldots, R_r$. Let us take all the column vectors of X to be identical, with ith element $\frac{1 - R_i}{n - s} \geqslant 0$. Then $\sum_{i=1}^{r} \frac{1 - R_i}{n - s} = \frac{r}{n-s} - \frac{\sigma(B)}{n - s} \leqslant \frac{r-(r+s-n)}{n-s} \leqslant 1$.

Thus all the column sums of X are atmost equal to 1. Clearly each row sum of $[B \; X]$ is unity by our construction of X.

In a similar manner we can determine Y such all column sums of $\begin{bmatrix} B \\ Y \end{bmatrix}$ are unity and each row sum of Y is at most unity. Take each element of Z to be $\frac{n-s-\sigma(X)}{(n-r)(n-s)} \geqslant 0$. It is easily seen that A so constructed is doubly stochastic.

NOTE : In Theorem 4.3 we have taken $r, s < n$. If either $r = n$ or $s = n$ then the following results are quite obvious.

(a) An $r \times n$ non-negative matrix B $(1 \leq r < n)$ is a submatrix of an $n \times n$ doubly-stochastic matrix if and only if each row sum of B is unity and each column sum is at most unity.

(b) An $n \times s$ non-negative matrix B $(1 \leq s < n)$ is a submatrix of an $n \times n$ doubly stochastic matrix if and only if each column sum of B is unity and each row sum of B is atmost unity.

COROLLARY 4.4 : If B is an $r \times s$ non-negative matrix with each row sum and each column sum atmost equal to unity, then B is the submatrix of a doubly stochastic matrix of suitable size.

PROOF : From Theorem 4.3 the condition B should satisfy is only $\sigma(B) \geq r + s - n$. This is certainly satisfied if n is taken sufficiently large.

EXAMPLE : $B = \begin{bmatrix} .1 & .2 \\ .3 & .2 \end{bmatrix}$ cannot be a submatrix of a 3×3 doubly stochastic matrix, for, $\sigma(B) = .8 < 2+2-3$. But B is certainly a submatrix of a 4×4 doubly stochastic matrix, taking

$$A = \begin{bmatrix} .1 & .2 & .35 & .35 \\ .3 & .2 & .25 & .25 \\ .3 & .3 & .2 & .2 \\ .3 & .3 & .2 & .2 \end{bmatrix}$$

using the construction in Theorem 4.3.

4.5 APPLICATION TO LATIN SQUARES :

Theorem 4.3 can be applied to latin squares in an interesting manner. Suppose we are given an $r \times s$ latin rectangle on n-symbols, say $(1, 2, \ldots, n)$. It is a well known result that this latin rectangle can be extended to an $n \times n$ latin square on $(1, 2, \ldots, n)$ if and only if each symbol in $1, 2, \ldots, n$ occurs atleast $(r+s-n)$ times in the latin rectangle.

We will prove only the necessity of the condition. We know that any permutation matrix is doubly stochastic. Hence if we replace a particular symbol by 1 and all other symbols by 0, a latin square becomes a permutation matrix. Take a particular symbol in $1, 2, \ldots, n$ and replace this symbol by 1 in the given latin rectangle and replace all the other symbols by 0. Then the latin rectangle becomes a part of a permutation matrix. Hence sum of the entries in this matrix must be atleast $(r+s-n)$ by Theorem 4.3 and this means that the chosen symbol must occur atleast $r+s-n$ times in the latin rectangle.

4.6 APPLICATION TO PERMANENTS OF DOUBLY STOCHASTIC MATRICES :

Let $IN_{r,n}$ be the set of all r-subsequences of $\{1, 2, \ldots, n\}$. Let $A = (a_{ij}) \in D_n$, the set of all $n \times n$ doubly stochastic matrices. We define an important function $S_r(A)$ for $A \in D_n$ as follows.

4.7...
$$S_r(A) = \sum_{\alpha, \beta \in IN_{r,n}} \text{per } A[\alpha \mid \beta].$$

In otherwords $S_r(A)$ is the sum of the permanents of all $r \times r$ submatrices of A. Let us also define $F_r(A)$.

$$4.8\ldots \quad F_r(A) = \sum_{\alpha,\beta \in IN_{r,n}} \text{per } A[\alpha|\beta]\, \sigma(A(\alpha|\beta)).$$

Clearly any one of the r! terms in per $A[\alpha|\beta]$ multiplied by any one of the $(n-r)^2$ terms in $\sigma(A(\alpha|\beta))$ gives a term in $S_{r+1}(A)$. But the number of such terms in F is clearly $\binom{n}{r}\binom{n}{r} r!\,(n-r)^2$. But the number of terms in $S_{r+1}(A)$ is $\binom{n}{r+1}\binom{n}{r+1}(r+1)!$ Thus we get

$$F_r(A) = \frac{\binom{n}{r}\binom{n}{r} r!\,(n-r)^2}{\binom{n}{r+1}\binom{n}{r+1}(r+1)!}\; S_{r+1}(A) = (r+1)\, S_{r+1}.$$

Thus we have proved.

$$4.9\ldots \quad (r+1)\, S_{r+1} = \sum_{\alpha,\beta \in IN_{r,n}} \text{per } A[\alpha|\beta]\, \sigma(A(\alpha|\beta)).$$

But, as $A \in \Omega_n$ we also have $\sigma(A(\alpha|\beta)) = \sigma(A[\alpha|\beta]) + n - 2r$ by Theorem 4.2.

$$\text{Hence } (r+1)S_{r+1} = \sum_{\alpha,\beta \in IN_{r,n}} \text{per } A[\alpha|\beta](n-2r+\sigma(A[\alpha|\beta]))$$

$$= (n-2r)S_r + \sum_{\alpha,\beta \in IN_{r,n}} \text{per } A[\alpha|\beta]\, \sigma(A[\alpha|\beta])$$

or $S_{r+1} = \frac{n-2r}{r+1} S_r + \frac{1}{r+1} \sum_{\alpha,\beta \in IN_{r,n}} \text{per } A[\alpha|\beta]\, \sigma(A[\alpha|\beta])$ or

4.10... $S_{r+1} = \frac{n-2r}{r+1} S_r + \frac{1}{r+1} T_r$ (say)

The above result is of fundamental importance.

THEOREM 4.11 : If $A \in D_n$, then $S_{r+1}(A) \leqslant \frac{n-r}{r+1} S_r(A)$ for $1 \leqslant r < n$.

PROOF : $T_r = \sum_{\alpha,\, \beta \in \mathrm{IN}_{r,n}} \operatorname{per} A[\alpha \mid \beta] \, \sigma(A[\alpha \mid \beta])$

$$\leqslant r \sum_{\alpha,\, \beta \in \mathrm{IN}_{r,n}} \operatorname{per} A[\alpha \mid \beta] \text{ for } \sigma(A[\alpha \mid \beta]) \leqslant r.$$

Hence $T_r \leqslant r\, S_r$ and from 4.10 we get

$$S_{r+1} \leqslant \frac{n-2r}{r+1} S_r + \frac{r}{r+1} S_r = \frac{n-r}{r+1} S_r.$$

4.12 THE CASE OF EQUALITY IN 4.11 :

Equality occurs if and only if $\sigma(A[\alpha \mid \beta]) = r$ whenever $\operatorname{per} A[\alpha \mid \beta] > 0$. But every doubly stochastic matrix has atleast one positive diagonal. Consider those $A[\alpha \mid \beta]$'s which have exactly r elements of <u>this</u> particular positive diagonal. Then $\operatorname{per} A[\alpha \mid \beta] > 0$ and hence we must have $\sigma(A[\alpha \mid \beta]) = r$. But $\sigma(A[\alpha \mid \beta]) = r$ implies $\sigma(A[\alpha \mid \beta]) = 0$ and $\sigma(A[\alpha \mid \beta]) = 0$. This is true if and only if all the elements other than those in the particular positive diagonal are zero. This means A is a permutation matrix.

The above argument can be easily understood by taking the positive diagonal to be the main diagonal and this can be always achieved by a suitable row permutation. Then, whenever we take a set of r rows all the elements outside the r x r principal submatrix of A lying on these r rows should be zero. But every non-diagonal element lies outside a suitably chosen r x r principal submatrix. Thus all the non-diagonal elements are zero and hence the matrix must be the unit matrix and the original matrix before row permutation must be a permutation matrix.

COROLLARY 4.13 : $S_r(A) \leqslant \frac{n-r+1}{r} \cdot S_{r-1} \leqslant \frac{(n-r+1)(n-r+2)}{r(r-1)} S_{r-2}$

$$\cdots\cdots \leqslant \frac{(n-r+1)(n-r+2)\ \cdots\ (n-1)}{r(r-1)\ \cdots\ 2} S_1$$

But $S_1 = n$ for $A \in D_n$. Thus $S_r(A) \leqslant \binom{n}{r}$ and the equality holds only for permutation matrices. Hence $\max_{A \in D_n} S_r(A) = \binom{n}{r}$.

In particular, when $r = n$, we get per $A = S_n \leqslant 1$, the equality holding only for permutation matrices.

4.14 APPLICATION OF COVARIANCE :

Suppose $\underline{a} = (a_1, a_2, \ldots, a_n)$ and $\underline{b} = (b_1, b_2, \ldots, b_n)$ are two real vectors. We define covariance between $\underline{a}$ and $\underline{b}$ by

$$\text{Cov}(\underline{a}, \underline{b}) = \frac{1}{n} \sum_{i=1}^{n} a_i b_i - \frac{\sum a_i}{n} \frac{\sum b_i}{n}$$

Hence if we take per $A[\alpha \mid \beta]$ and $\sigma(A[\alpha \mid \beta])$ arranged in lexicographical order with respect to α and β in $IN_{r,n}$ and form vectors P_r and σ_r of dimension $\binom{n}{r}\binom{n}{r}$, then we have

$$\text{Cov}(P_r, \sigma_r) = \frac{1}{\binom{n}{r}^2} \sum_{\alpha, \beta \in IN_{r,n}} \text{per } A[\alpha \mid \beta]\, \sigma(A[\alpha \mid \beta])$$

$$- \frac{S_r}{\binom{n}{r}^2} \frac{\binom{n}{r}^2 \frac{r^2}{n}}{\binom{n}{r}^2}$$

$$= \frac{1}{\binom{n}{r}^2} \left[(r+1)\, S_{r+1} - \left(n-2r + \frac{r^2}{n}\right) S_r\right]$$

$$= \frac{r+1}{\binom{n}{r}^2} \left[\left(S_{r+1} - \frac{n-r}{r+1} S_r\right) + \frac{r}{n} \frac{n-r}{r+1} S_r\right]$$

But $S_{r+1} - \frac{n-r}{r+1} S_r$ and S_r are maximum for permutation matrices only. Thus $\text{Cov}(P_r, \sigma_r)$ is a maximum only for permutation matrices and hence we get $\text{Cov}(P_r, \sigma_r) \leq \frac{(r+1)\, r\, (n-r)}{\binom{n}{r}^2\, n(r+1)} \binom{n}{r} = \frac{r(n-r)}{\binom{n}{r}.n}$, equality holding good only for permutation matrices.

4.15 TVERBERG'S CONJECTURE :

Tverberg [16] conjectured that for $1 < r \leq n$ $S_r(A)$ attains the minimum value for $A \in D_n$ only at $J_n = (\frac{1}{n})$. Vander Waerden conjucture is Tverberg's conjecture for $r = n$.

Let us introduce a new function on D_n. For $A \in D_n$, $0 \leq r \leq n$ define $h_r(A) = \frac{S_r(A)}{S_r(J_n)}$. When $r = 0$ we define $S_r(A) = 1$ for every $A \in D_n$. Thus $h_0(A) = 1$; $h_1(A) = \frac{S_1(A)}{S_1(J_n)} = \frac{n}{n} = 1$.

Djokovic [ ] conjectured that $S_{r+1} \geqslant \frac{(n-r)^2}{n(r+1)} S_r$ for $1 \leqslant r < n$, equality holding only for $A = J_n$. But this is equivalent to the following.

4.16 DJOKOVIC CONJECTURE :

$h_{r+1}(A) \geqslant h_r(A)$ for $A \in D_n$, $1 \leq r < n$ equality holding only for J_n.

Note that 4.16 implies 4.15 for,

$$h_{r+1} \geqslant h_r \Rightarrow h_r \geqslant h_1 \Rightarrow h_r \geqslant 1 \Rightarrow S_r(A) \geqslant S_r(J_n)$$

4.17... 4.16 is equivalent to $\mathrm{Cov}(P_r, \sigma_r) \geqslant 0$ for $1 \leqslant r < n$.

PROOF : $\mathrm{Cov}(P_r, \sigma_r) = \frac{r! \, (n-r)^2}{n} (h_{r+1}(A) - h_r(A))$ and the result follows.

COROLLARY 4.18 : Since we have already proved that $\mathrm{Cov}(P_r, \sigma_r) > 0$ for permutation matrices it follows that Djokovic conjecture is true for permutation matrices. This is an elegant proof.

We now derive some remarkable relations for h_r.

THEOREM 4.19 : Suppose $A \in \Omega_n$ and $p \in [0,1]$. If $A_p = pA+(1-p)J_n$ then $h_n(A_p) = \sum_{r=0}^{n} \binom{n}{r} p^r(1-p)^{n-r} h_r(A)$.

PROOF : Calenisllo's theorem gives, for $n \times n$ matrices A and B,

$$\text{per }(A+B) = \sum_{r=0}^{n} \sum_{\alpha, \beta \in IN_{r,n}} \text{per } A[\alpha \mid \beta] \text{ per } B(\alpha \mid \beta)$$

Hence $\text{per }(A_p) = \text{per }(pA+(1-p)J_n) = \sum_{r=0}^{n} \sum_{\alpha, \beta \in IN_{r,n}} \text{per}(pA)[\alpha \mid \beta]$.

$$\cdot \text{ per } (1-p)J_n) (\alpha \mid \beta)$$

$$= \sum_{r=0}^{n} p^r(1-p)^{n-r} \sum_{\alpha, \beta \in IN_{r,n}} \text{per}A[\alpha \mid \beta] \text{per } J_n(\alpha \mid \beta).$$

But $\text{per }(A_p) = h_n(A_p)\ S_n(J_n)$. Hence we get

$$h_n(A_p) = \sum_{r=0}^{n} p^r(1-p)^{n-r} \frac{S_r(A)}{S_r(J_n)} \binom{n}{r} = \sum_{r=0}^{n} \binom{n}{r} p^r(1-p)^{n-r} h_r(A).$$

Note that in this proof we have not used the fact that $A \in \Omega_n$ or $p \in [0,1]$. In fact the identity in 4.18 is true for any real matrix A and any real p. But we will be using 4.18 for only $A \in \Omega_n$ and $p \in [0, 1]$.

THEOREM 4.20 : Suppose $A \in \Omega_n$, $p \in [0,1]$. If $A_p = pA + (1-p)J_n$, then $h_r(A_p) = \sum_{k=0}^{r} \binom{r}{k} p^k(1-p)^{r-k} h_k(A)$ for $1 \leq r < n$.

PROOF : Though 4.19 seems to be a particular case of 4.20 actually we can derive 4.20 from 4.19.

Suppose p, $q \in [0, 1]$. $A_{pq} = pq\,A + (1-pq)\,J_n$.

But $(A_p)_q = q\,(pA + (1-p)J_n) + (1-q)J_n = pqA + (1-pq)\,J_n = A_{pq}$

Hence $h_n(A_{pq}) = \sum_{r=0}^{n} \binom{n}{r} (pq)^r (1-pq)^{n-r}\, h_r(A)$ and

$$h_n(A_{pq}) = \sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k}\, h_k(A_p).$$

Taking $t = q/(1-q)$ the last results yield

$$\sum_{r=0}^{n} \binom{n}{r} p^r \left(\frac{t}{1+t}\right)^r \left(1 - p\,\frac{t}{1+t}\right)^{n-r} h_r(A) \equiv \sum_{k=0}^{n} \binom{n}{k} \left(\frac{t}{1+t}\right)^k \left(\frac{1}{1+t}\right)^{n-k} h_k(A_p), \text{or}$$

$$\sum_{r=0}^{n} \binom{n}{r} p^r t^r [1+(1-p)t]^{n-r}\, h_r(A) \equiv \sum_{k=0}^{n} \binom{n}{k} t^k\, h_k(A_p).$$

Equating coefficient of t^k on both sides we get

$$\binom{n}{k} h_k(A_p) = \sum_{r=0}^{k} \binom{n}{r} p^r\, h_r(A) \binom{n-r}{k-r} (1-p)^{k-r}. \text{ But } \binom{n}{r}\binom{n-r}{k-r} = \binom{n}{k}\binom{k}{r}.$$

Hence we get $h_k(A_p) = \sum_{r=0}^{k} \binom{k}{r} p^r (1-p)^{k-r}\, h_r(A)$.

COROLLARY 4.21 : For $p \in [0.1]$, $\binom{r}{k} p^k (1-p)^{r-k} \geqslant 0$ and $\sum_{k=0}^{r} \binom{r}{k} p^k (1-p)^{r-k} = (p+1-p)^r = 1$. Thus $h_r(A_p)$ is a weighted average of $h_k(A)$ for $k = 0, 1, 2, \ldots, r$ and hence

$$\min_{k \leqslant r} h_k(A) \leqslant h_r(A_p) \leqslant \max_{k \leqslant r} h_k(A).$$

4.22 BOUNDARY OF D_n :

The set Z_n of all n x n doubly stochastic matrices with positive entries forms the interior of D_n. Thus the boundary $D_n - Z_n$ of D_n consists of doubly stochastic matrices with atleast one entry 0. We will now prove that the Djokovic conjecture is true in D_n if and only if it is true in $D_n - Z_n$, the boundary of D_n.

We need a preliminary result which is quite interesting by itself.

THEOREM 4.23 : Suppose $a_0 \geqslant a_1 \geqslant a_2 \geqslant \dots$ and $b_0 = a_0$ and

$$b_n = p^n a_0 + \binom{n}{1} p^{n-1}(1-p) a_1 + \binom{n}{2} p^{n-2}(1-p)^2 a_2 + \dots + (1-p)^n a_n$$

for n = 1, 2, 3, ... and $p \in [0, 1]$. Then $b_0 \geqslant b_1 \geqslant b_2 \geqslant \dots$

PROOF :
$$b_{r+1} = p^{r+1} a_0 + \binom{r+1}{1} p^r (1-p) a_1 + \binom{r+1}{2} p^{r-1} (1-p)^2 a_2 + \dots + (1-p)^{r+1} a_{r+1}$$

$$= p\left\{p^r a_0 + \binom{r}{1} p^{r-1}(1-p) a_1 + \dots + (1-p)^r a_r\right\} + (1-p) \cdot \left\{p^r a_1 + \binom{r}{1} p^{r-1} (1-p) a_2 + \dots + (1-p)^r a_{r+1}\right\}$$

$$\leqslant p\, b_r + (1-p)\, b_r = b_r$$

for, $p^r a_0 + \binom{r}{1} p^{r-1}(1-p) a_1 + \dots \geqslant p^r a_1 + \binom{r}{1} p^{r-1} (1-p) a_2 + \dots$

Thus $b_{r+1} \geqslant b_r$ for r = 0, 1, 2, ... and the proof is completed.

NOTE : Clearly $\geqslant$ can be replaced by $\leqslant$ throughout in Theorem 4.23.

THEOREM 4.24 : Djokovic conjecture is true for D_n if and only if it is true for $D_n - Z_n$.

PROOF : Necessity of the condition is obvious. Let us assume that Djokovic conjecture is true for $D_n - Z_n$ and prove sufficiency.

If $A \in Z_n$ and $A \notin J_n$ then the 'line segment' from J_n to A when produced meets $D_n - Z_n$ at some 'point' B. Clearly, then, that $A = pB + (1-p) J_n$ for some $p \in (0, 1)$. In fact p can be got as follows. Let $A = (a_{ij})$. Consider $qA + (1-q) J_n = (q\, a_{ij} + \frac{1-q}{n})$. Choose q_o as maximum q for which $q\, a_{ij} + \frac{1-q}{n} \geqslant 0$ for all i, j.

Then $q(a_{ij} - \frac{1}{n}) \geqslant -\frac{1}{n}$ or $q \leqslant \frac{1}{1-na_{ij}}$. Thus $q_o = \frac{1}{1-n(\min\limits_{i,j} a_{ij})}$.

Take $B = q_o A + (1-q_o) J_n$. Clearly $B \in D_n - Z_n$, for, atleast one element of B is 0 by our choice of q_o. Then $p = 1/q_o$. Since $q\, a_{ij} + \frac{1-q}{n} = a_{ij} > 0$ for $q = 1$, q_o must be strictly greater than 1 and thus $p \in (0, 1)$.

$$h_r (A) = \sum_{s=0}^{r} h_s(B) \binom{r}{s} p^s (1-p)^{r-s} \quad \text{for } r = 0, 1, 2, \ldots, n.$$

If we assume that $h_0(B) \leqslant h_1(B) \leqslant \cdots \leqslant h_n(B)$ then by theorem 4.23 and its 'Note' it follows that $h_0(A) \leqslant h_1(A) \leqslant \cdots \leqslant h_n(A)$.

COROLLARY 4.25 : If Djokovic conjecture is true for an $A \in D_n$, then it is true for all matrices in the "segment" A to J_n.

The next theorem is related to the van der Waerden conjecture.

THEOREM 4.26 : If $A \in D_n$ then (i) per $A \geqslant \frac{n!}{n^n}(1-p_o)^n$ where $p_o = 1-n \min_{i,j} a_{ij}$. (ii) per $A_p \geqslant p^n$ per A for $p \in [0,1]$.

PROOF : Let the line segment joining J_n to A meet $D_n - Z_n$ at B when produced. Then $A = p_o B + (1-p_o)J_n$ where $p_o = 1-n \min_{i,j} a_{ij}$.

$$h_n(A) = \sum_{r=0}^{n} \binom{n}{r} p_o^r (1-p_o)^{n-r} \; h_r(B)$$

$$\geqslant (1-p_o)^n, \text{ the first term on R.H.S.}$$

But $h_n(A) = \text{per } A/\text{per } J_n$. Thus Per $A \geqslant \frac{n!}{n^n}(1-p_o)^n$. Also

$$h_n(A_p) = \sum_{r=0}^{n} \binom{n}{r} p^r (1-p)^{n-r} \; h_r(A) \geqslant p^n \; h_n(A).$$

Cancelling per J_n we get per $(A_p) \geqslant p^n$ per A.

Unfortunately these results are not very sharp as they are got by omitting a large number of terms.

COROLLARY 4.27 : For $A \in D_n$ (i) $S_r(A) \geqslant S_r(J_n)\,(1-p_o)^r$ (ii) $S_r(A_p) \geqslant p^r\, S_r(A)$ for $r = 1, 2, \ldots, n$.

These results follow from $h_r(A_p) = \sum_{s=0}^{r} \binom{r}{s} p^s (1-p)^{r-s} h_s(A)$ and arguments similar to those used in Theorem 4.26.

4.28 DEFINITION : Let $h(A) = \min_r h_r(A)$ and $H(A) = \max_r h_r(A)$ for $A \in D_n$.

THEOREM 4.29 : For $A \in D_n$, $h(A_p)$ is a decreasing function of p and $H(A_p)$ is an increasing function of p for $p \in [0, 1]$.

PROOF : $h_r(A_p) = \sum_{s=0}^{r} \binom{r}{s} p^s (1-p)^{r-s} h_s(A)$. Thus

$$\min_{s \leq r} h_s(A) \leq h_r(A_p) \leq \max_{s \leq r} h_s(A)$$

Hence $\min_{r \leq n} \min_{s \leq r} h_s(A) \leq \min_{r \leq n} h_r(A_p) \leq \max_{r \leq n} h_r(A_p) \leq \max_{r \leq n} \max_{s \leq r} h_s(A)$.

4.30... Thus $h(A) \leq h(A_p) \leq H(A_p) \leq H(A)$.

If $0 \leq p_1 < p_2 \leq 1$, then $p_1 = q\, p_2$ for some $q \in [0, 1]$.

$h(A_{p_1}) = h(A_{qp_2}) = h((A_{p_2})_q) \geq h(A_{p_2})$ from 4.30

$H(A_{p_1}) = H(A_{qp_2}) = H((A_{p_2})_q) \leq H(A_{p_2})$ from 4.30.

Thus $h(A_p)$ is decreasing and $H(A_p)$ is increasing with respect to p for $p \in [0, 1]$.

COROLLARY 4.31 : $h(A)$ attains the maximum value 1 at $A = J_n$ and $H(A)$ attains the minimum value 1 <u>uniquely</u> at $A = J_n$.

PROOF : $h(A)$ attains the maximum value at $A = J_n$ by Theorem 4.29 and $h(J_n) = 1$. $H(A)$ attains the minimum value at J_n. We can show the uniqueness as follows. We can easily show that $h_2(A) > 1$ if $A \neq J_n$. Hence $H(A) \geq h_2(A) > 1$ and the uniqueness follows.

That $h_2(A) > 1$ *for* $A \neq J_n$ will now be proved.

THEOREM 4.32 : If $A \in D_n$ and $A \neq J_n$, then $h_2(A) > 1$.

PROOF : $S_{r+1} = \frac{n-2r}{r+1} S_r + \frac{1}{r+1} \sum_{\alpha,\beta \in IN_{r,n}} \text{per } A[\alpha|\beta] \, \sigma(A[\alpha|\beta])$ by 4.10.

Putting $r = 1$, $S_2 = \frac{n-2}{2} S_1 + \frac{1}{2} \sum_i \sum_j a_{ij}^2$. But $S_1 = n$. Thus

$S_2 = \frac{n-2}{2} + \frac{1}{2} \sum_i \sum_j a_{ij}^2$. But $\sum_i \sum_j a_{ij}^2 \geq (\sum_i \sum_j a_{ij})^2/n^2 = 1$

and equality holds in this *cauchy-schwartz inequality* if and only if $a_{ij} = \frac{1}{n}$ for all i,j since $\sum_i \sum_j a_{ij} = n$. Thus, if $A \neq J_n$ we must have $\sum_i \sum_j a_{ij}^2 > 1$. Hence $S_2 \quad \frac{n(n-2)}{2} + \frac{1}{2}$ for $A \neq J_n$ or $S_2 > \frac{(n-1)^2}{2}$. But $S_2(J_n) = \binom{n}{2} \binom{n}{2} \frac{2}{n^2} = \frac{(n-1)^2}{2}$.

Thus $h_2(A) = S_2(A)/S_2(J_n) > 1$ for $A \neq J_n$.

COROLLARY 4.33 : The function $h_2(A)$ attains the minimum value 1 uniquely at $A = J_n$.

THEOREM 4.34 : For a fixed $r = 2, 3, \ldots, n$, $h_r(A)$ has a local minimum at $A = J_n$ in D_n.

PROOF : $h_r(A_p) = \sum_{s=0}^{r} \binom{r}{s} p^s(1-p)^{r-s} h_s(A)$

$$= h_o(A)+p[rh_1(A)-rh_o(A)]+p^2[\binom{r}{2}h_2(A)+\binom{r}{2}-r(r-1)]+\ldots$$

$$= 1 + p^2 \binom{r}{2} [h_2(A) - 1] + \ldots \quad \text{for } h_o(A)=h_1(A)=1.$$

If $A \neq J_n$ then $h_2(A) - 1 > 0$ and hence for sufficiently small $p > 0$ $h_r(A_p) > 1$. But $h_2(A_o) = h_2(J_n) = 1$. This proves the theorem.

In particular per A has a local minimum at $A = J_n$. In fact per $A >$ per J_n if A is in a sufficiently small deleted neighbourhood of J_n in D_n.

4.35 MINIMIZING MATRICES :

In D_n, let A be a matrix such that per $A = \min_{S \in D_n}$ per S. Then A is called a minimizing matrix. If $A = (a_{ij})$ is a minimizing matrix Marcus [12] has shown that

$a_{ij} > 0$ implies per $A(i \mid j)$ = per A.

London [10] has shown that per $A(i \mid j) \geq$ per A for all (i,j). Consider the n^2 pairs (a_{ij}, per $A(i \mid j)$). We note that we can arrange these n^2 pairs with a_{ij}'s in increasing order and per $A(i \mid j)$ in decreasing order, for, first we can arrange a_{ij}'s that are zero in such a way per $A(i \mid j)$ are decreasing. This can be followed by positive a_{ij}'s arranged in increasing order, the corresponding per $A(i \mid j)$'s being constant = per A.

Thus Cov $(a_{ij}$, per $A(i|j)) \leq 0$ with equality if and only if either (i) all a_{ij}'s are positive or (ii) some a_{ij}'s are zero but all per $A(i|j)$ are equal. This result follows from the fact that if we have n pairs $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$ such that $a_1 \leq a_2 \leq \cdots \leq a_n$ and $b_1 \geq b_2 \geq \cdots \geq b_n$ then Cov $(a_i, b_i) \leq 0$ equality holding only if either all a_i's are equal or all b_i's are equal. In our case if all a_{ij}'s are equal then each must be $\frac{1}{n}$ and hence $A = J_n$ and per $A(i|j) =$ per A for all (i, j). If some a_{ij}'s are zero, then all per $A(i|j)$ must be equal.

In any case it is clear that for a minimizing matrix A, Cov. $(a_{ij}$, per $A(i|j)) = 0$ if and only if per $A(i|j) =$ per A for **all** (i, j).

$$\text{But } \sigma(A(i|j)) = n-2 + a_{ij}.$$

Hence Cov $(\sigma(A(i|j))$, per $A(i|j)) =$ Cov $(a_{ij}$, per $A(i|j)) \leq 0$ for any minimizing matrix A. Thus Cov $(P_{n-1}, \sigma_{n-1}) \leq 0$ equality holding only when per $A(i|j) =$ per A for all (i, j). But Djokovic conjecture, for $r = n-1$, is precisely that Cov $(P_{n-1}, \sigma_{n-1}) \geq 0$ equality holding only for $A = J_n$.

At this stage it is interesting to consider the matrix. $M = \frac{1}{2}(I+P)$ where P is an $n \times n$ permutation matrix with 1 in positions $(1, 2), (2, 3), (3, 4), \ldots, (n-1, n), (n, 1)$. Clearly $M \in D_n$. This matrix has the peculiar property that per $M(i \mid j)$ = per $M = \frac{1}{2^{n-1}}$ for all (i,j). Thus $h_n(M) = \frac{1}{2^{n-1}} \cdot \frac{n}{n!}$ and $h_{n-1}(M) = \frac{1}{2^{n-1}} \cdot \frac{n^n}{n!}$. Thus $h_n(M) - h_{n-1}(M) = 0$.

$$\text{But } \frac{d}{d_p} h_n(M_p) = n[h_n(M_p) - h_{n-1}(M_p)] = 0 \text{ at } p = 1.$$

Thus per (M_p) has a stationary value at $p = 1$ i.e. $M_p = M$. If $n = 2$, actually $M = J_n$ and the stationary value is actually a minimum value. But when $n > 2$ the permanent function is neither an absolute maximum nor an absolute minimum as

$$\text{per } I = 1 > \frac{1}{2^{n-1}} > \frac{n!}{n^n} = \text{Per } J_n.$$

when $n = 3$, $S_3(M)$ = per $M = \frac{1}{4}$, per $J_3 = \frac{2}{9}$.

Hence $h_3(M) = \frac{9}{8}$ and $h_2(M) = \frac{9}{8}$, and $h_1(M) = h_0(M) = 1$.

Thus $h_3(M_p) = p^3 \; h_3(M) + 3p^2(1-p) \; h_2(M) + 3 \; P(1-p) \; h_1(M) + (1-p)^3 h_0(M)$

$$= \frac{1}{8} \; (8+3 \; p^2 - 2 \; p^3).$$

Thus $\frac{d}{d_p} \; h_3(M_p) = \frac{3}{4} \; p(1-p)$, $\frac{d^2}{dp^2} h_3(M_p) = \frac{3}{4} \; (1-2p)$.

Thus at $p = 1$, $h_3(M_p)$ is a maximum. In fact $\frac{d}{dp} \; h_3(M_p) = \frac{3}{4}p(1-p) \geqslant 0$ for $p \in [0, 1]$ shows that $h_3(M_p)$ is an increasing function of p in the interval $[0,1]$.

Friedland and Minc [12] proposed the following problem.

PROBLEM 4.36 : Find matrices A on the boundary of Ω_n other than a permutation matrix P or $(J-P)/(n-1)$ so that the permanent be monotone on the segment $(1-\theta) J_n + \theta A$, $0 \leqslant \theta \leqslant 1$.

Of course $M = (J-P)/2$ when $n = 3$ where $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

Thus what we have proved for M when n = 3 is not unexpected. But a similar result for n = 4 will certainly be interesting.

RESULT 4.37 : Per (M_p) is an increasing function of p for $p \in [0, 1]$. when $n = 4$.

PROOF : When $n = 4$, $h_4(M) = h_3(M) = \frac{4}{3}$. By actual calculation $S_2(M) = 5$. Hence $h_2(M) = 10/9$ and $h_1(M) = h_0(M) = 1$. Thus

$$h_4\ (M_p) = \frac{1}{3}\ (3 + 2\ p^2 - p^4)$$

$$\frac{d}{dp}\ h_4\ (M_p) = \frac{4}{3}\ p(1-p^2) \geqslant 0 \text{ for } p \in [0, 1].$$

Thus $h_4(M_p)$ and consequently per (M_p) is an increasing function of p.

Thus we have given an example for Friedland and Minc problem. We conjecture the following.

CONJECTURE 4.38 : Per (M_p) is an increasing function of p for $p \in [0, 1]$ for any n.

4.39 PERMANENTS AND MULTINOMIAL DISTRIBUTION :

Let A be an $n \times n$ non-negative matrix. Let $x^T = (x_1, x_2, \ldots, x_n)$. If $y = A\,x$ and $y^T = (y_1, y_2, \ldots, y_n)$ then

4.40... $$y_1 y_2 \cdots y_n = \sum_{\omega \in ND_{n,n}} \operatorname{per} A[1,2,\ldots,n \mid \omega] \; \frac{x_1^{\omega_1} x_2^{\omega_2} \cdots x_n^{\omega_n}}{\omega_1!\,\omega_2! \cdots \omega_n!}$$

where $ND_{n,n}$ is the set of all non-decreasing n-sequences formed from $\{1, 2, \ldots, n\}$ and ω_i is the number of times i occurs in the sequence ω.

Suppose we take $x_i \geqslant 0$ for $i = 1, 2, \ldots, n$ and $\sum_{i=1}^{n} x_i = 1$. Then we can write,

4.41... $$n!\, y_1 y_2 \cdots y_n = \sum_{\omega \in ND_{n,n}} \frac{n!}{\omega_1!\,\omega_2! \cdots \omega_n!} \, x_1^{\omega_1} x_2^{\omega_2} \cdots x_n^{\omega_n} \operatorname{per} A[1, 2, \ldots, n/\omega]$$

clearly $\frac{n!}{\omega_1! \cdots \omega_n!} \, x_1^{\omega_1} \cdots x_n^{\omega_n}$'s represent multinomial probabilities. Thus $n!\, y_1 y_2 \cdots y_n$ is actually the expectation of $\operatorname{per} A[1,2,\ldots, n/\omega]$ with respect to this multinomial distribution.

Suppose $A = (a_{ij})$ is a column stochastic matrix. Then

$$\sum_{i=1}^{n} y_i = \sum_{i} \sum_{j} a_{ij}\, x_j = \sum_{j} x_j \sum_{i} a_{ij} = 1.$$

Thus $y_i \geqslant 0$ and $\sum y_i = 1$. Hence the maximum value of $y_1 y_2 \cdots y_n$ is $\frac{1}{n^n}$. Hence we get

4.42... $E(\text{per } A\,[1, 2, \ldots, n/w]) \leqslant \frac{n!}{n^n}$

where E is expectation with respect to the multinomial distribution with parameters $(n;\ x_1, x_2, \ldots, x_n)$. Equality in 4.42 holds if and only if each $y_i = 1/n$.

Suppose A is doubly stochastic. Then $y_i = \sum_{j=1}^{n} a_{ij}\, x_j$ is a weighted average of x_j's. Using A.M - G.M inequality, we have

$$y_i = \sum_j a_{ij}\, x_j \geqslant \prod_{j=1}^{n} x_j^{a_{ij}}$$

Thus $\quad y_1 y_2 \cdots y_n \geqslant \prod_{j=1}^{n} x_j^{\sum_i a_{ij}} = \prod_{j=1}^{n} x_j$. Hence we get

4.43... $\quad E(\text{per } A\,[1, 2, \ldots, n \quad w]) \geqslant n!\, x_1\, x_2 \cdots x_n$.

Therefore, if $A \in D_n$, then with respect to the multinomial distribution with parameters $[n, x_1, x_2, \ldots, x_n]$ we have

4.43... $n!\, x_1 x_2 \cdots x_n \leqslant E(\text{per } A\,[1, 2, \ldots, n \quad w]) \leqslant \frac{n!}{n^n}$

In the A.M. - G.M. inequality, equality occurs if and only if all x_j's are equal or $A = J_n$. Thus equality occurs in 4.43 only if each $x_i = \frac{1}{n}$ or $A = J_n$. Thus

$$E(\text{per } A\,[1, 2, \ldots, n \quad w]) = \frac{n!}{n^n}$$

only for the multinomial distribution with parameters $[n; \frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}]$ or when $A = J_n$.

4.45 INTERPRETATION OF 4.44 :

Suppose $x_1 = x_2 = \ldots = x_n = \frac{1}{n}$. Then $E(\text{per } A[1, 2, \ldots, n \mid \omega]) = \frac{n!}{n^n}$. Expectation being an average, this means.

4.46... $\min_{\omega} \text{ per } A[1, 2, \ldots, n \mid \omega] \leqslant \frac{n!}{n^n} \leqslant \max_{\omega} \text{ per } A[1, 2, \ldots, n \mid \omega]$.

Thus, if per $[1, 2, \ldots, n \mid \omega]$ is independent of ω, then it must be equal to $n!/n^n$. But in this case we can also show that $A = J_n$. For, taking $\omega = [1, 1, \ldots, 1]$, we get per $A[1, 2, \ldots, n \mid \omega] = n!\, a_{11}\, a_{21} \ldots a_{n1} = \frac{n!}{n^n}$. Thus $a_{11}\, a_{21}\, a_{31} \ldots a_{n1} = \frac{1}{n^n}$. Since $\sum_{i=1}^{n} a_{i1} = 1$ this implies $a_{11} = a_{21} = \ldots = a_{n1} = \frac{1}{n}$. In a similar manner we can show that $a_{1j} = a_{2j} = \ldots = a_{nj} = \frac{1}{n}$ for $j = 1, 2, \ldots, n$. Thus $A = J_n$.

COROLLARY 4.47 : If $A \neq J_n$, then we have

$$\min_{\omega} \text{ per } A[1, 2, \ldots, n \mid \omega] < \frac{n!}{n^n} < \max_{\omega} \text{ per } A[1, 2, \ldots, n \mid \omega]$$

Van der Waerden conjecture states that per $A > \frac{n!}{n^n}$ if $A \in D_n$ and $A \neq J_n$. In view of the Corollary 4.47 it may be thought that if we can show that $\max_{\omega} \text{ per } A[1, 2, \ldots, n \mid \omega] = \text{per } A[1, 2, \ldots, n \mid 1, 2, \ldots, n] = \text{per } A$ then van der Waerden conjecture will be atonce affirmatively resolved.

Unfortunately it is not true that $\max_{\omega} \text{per } A[1,2,\ldots,n \mid \omega] = \text{per } A$ for every $A \in \Omega_n$ as the following counter example shows.

$$A = \frac{1}{24}\begin{bmatrix} 11 & 5 & 8 \\ 13 & 11 & 0 \\ 0 & 8 & 16 \end{bmatrix} \in \Omega_3. \quad \text{Then per } A = \frac{3804}{24^3}.$$

$$\text{But per } A[1,2,3 \mid 1,1,3] = \text{per } \frac{1}{24}\begin{bmatrix} 11 & 11 & 8 \\ 13 & 13 & 0 \\ 0 & 0 & 16 \end{bmatrix} = \frac{4576}{24^3}$$

and hence $\max_{\omega} \text{per } A[1,2,3 \mid \omega] \neq \text{per } A$.

4.48 RYSER'S CONJECTURE AND JURKAT'S COUNTER EXAMPLE.:

Ryser conjectured that per $(AB) \leq \min(\text{per } A, \text{per } B)$ for $A, B \in \Omega_n$. If this conjecture is true then vander Waerden conjecture is immediately resolved. For, then, per $A \geq \text{Per}(AJ_n) = \text{per } J_n$ for every $A \in \Omega_n$. Unfortunately Rysor's conjecture is not true and Jurkat gave the following counter example.

$$A = \frac{1}{24}\begin{bmatrix} 11 & 5 & 8 \\ 13 & 11 & 0 \\ 0 & 8 & 16 \end{bmatrix}, \quad B = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Clearly $A, B \in \Omega_3$. Per $AB = \frac{3840}{24^3}$, per $A = \frac{3804}{24^3}$. Thus per $AB >$ per A

Let us analyse why this happens.

4.49 DEFINITION : A non-negative $n \times n$ matrix A is said to be column dominating or simply cod if per $A = \max_{\omega} \text{per } A[1,2,\ldots,n \mid \omega]$

for $w \in ND_{n,n}$. Let $\bar{D}_n$ be the subset of D_n, consisting of all cod matrices.

THEOREM 4.50 : If A, $B^T \in \bar{D}_n$ then per (AB) $\leq$ min (per A, per B).

PROOF : $$\text{per}(AB) = \sum_{w \in ND_{n,n}} \frac{\text{per } A[1,2,\ldots,n \mid w]\,\text{per } B[w \mid 1,2,\ldots,n]}{w_1!\, w_2! \ldots w_n!}$$

$$\leq \text{per } A \sum \frac{\text{per } B[w \quad 1,2,\ldots,n]}{w_1!\, w_2! \ldots w_n!}$$

But $$\text{per } J_n = \text{per}(J_n B) = \sum_{B} \frac{\text{per } J_n[1,2,\ldots,n \mid w]\,\text{per } B[w \mid 1,2,\ldots,n]}{w_1!\, w_2! \ldots w_n!}$$

$$= \text{per } J_n \sum_{w} \frac{\text{per } B[w \mid 1,2,\ldots,n]}{w_1! \ldots w_n!}$$

and hence $$\sum_{w} \frac{\text{per } B[w \quad 1,2,\ldots,n]}{w_1! \ldots w_n!} = 1$$

and consequently Per (AB) $\leq$ per A.

Also per (AB) = per ($B^T A^T$) per B^T = per B if $B^T \in D_n$.

COROLLARY 4.51 : If per (AB) $>$ min (per A, per B) then either $A \notin \bar{D}_n$ or $B^T \notin \bar{D}_n$.

In Jurkat's example $A \notin \bar{D}$

THEOREM 4.52 : Van der Waerden conjecture is true for $A \in \bar{D}_n$.

PROOF : Clearly $J_n \in \bar{D}_n$. If $A \in \bar{D}_n$, then per J_n = per $(AJ_n) \leq$ min (per A, per J_n). Thus per $A \geq$ per J_n.

Ryser also conjectured that per $(AA^T) \leq$ per A for $A \in D_n$. But Morris Newman gave a counter example. Nevertheless for $A \in \bar{D}_n$ the conjecture is certainly true for per $(AA^T) \leq$ min (per A, per A) if $A \in \bar{D}_n$.

Let us see the structure of $\bar{D}_n$ now. For a permutation matrix P clearly per $P[1, 2, \ldots, n \mid w] = 0$ if $w \neq [1,2,\ldots,n]$. But per P = 1. Hence $P \in \bar{D}_n$. Also $J_n \in \bar{D}_n$.

THEOREM 4.53 : $\bar{D}_2 = D_2$.

PROOF : Any element of D_2 is of the form $A = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$ for $p \in [0, 1]$. Per $A = p^2 + (1-p)^2$. But per $A[12 \mid 11] = 2p(1-p)$ and $p^2 + (1-p)^2 - 2p(1-p) = (2p-1)^2 \geq 0$. Thus $A \in \bar{D}_2$. Thus $D_2 \subset \bar{D}_2$. But $\bar{D}_2 \subset D_2$. Hence $D_2 = \bar{D}_2$.

THEOREM 4.54 : For each $n \geq 3$, $D_n - \bar{D}_n$ is nonempty.

PROOF : Let us prove the theorem by induction on n. For n = 3,

$$A = \frac{1}{24} \begin{bmatrix} 11 & 5 & 8 \\ 13 & 11 & 0 \\ 0 & 8 & 16 \end{bmatrix} \in D_3 - \bar{D}_3$$ and hence $D_3 - \bar{D}_3$ is nonempty.

Suppose $D_n - \bar{D}_n$ is nonempty for $n = m$. Let $A \in D_m - \bar{D}_m$. Take $B = A \oplus I_1$, the direct sum of A and the unit matrix of order 1. $\text{Per } B[1,2, \ldots, m, m+1 \mid \bar{w}, m+1] = \text{per } A[1, 2, \ldots, m \mid \bar{w}]$ $\text{per } A = \text{Per } B$ where $\bar{w} \in ND_{m,m}$ is such that $\text{per } A[1,2, \ldots, n \mid \bar{w}] > \text{per } A$. Hence $B \notin D^c_{m+1}$. Thus it $D_m - \bar{D}_m$ is nonempty then $D_{m+1} - \bar{D}_{m+1}$ is also nonempty and hence by induction the theorem follows.

THEOREM 4.55 : For $n \geq 3$, $\bar{D}_n$ is <u>not</u> a convex set.

PROOF : We know that $P \in \bar{D}_n$, if P is a permutation matrix of order n. Also D_n is the convex hull of all the permutation matrices. If $\bar{D}_n$ is convex then $\bar{D}_n \supset D_n$ which is impossible for $D_n - \bar{D}_n$ is nonempty for $n \geq 3$. Thus $\bar{D}_n$ cannot be a convex set.

THEOREM 4.57 : If $A \in \bar{D}_n$ then AP, $PA \in \bar{D}_n$ for any permutation matrix P of order n.

PROOF : The property of column dominance is symmetric with respect to columns and symmetric with respect to rows and hence the theorem follows.

THEOREM 4.58 : If $A_1 \in \bar{D}_{n_1}$ and $A_2 \in \bar{D}_{n_2}$, then $A_1 \oplus A_2 \in \bar{D}_{n_1+n_2}$

PROOF : There exist $w_1 \in ND_{n_1,n_1}$ and $w_2 \in ND_{n_2,n_2}$ such that

$\text{per } A_1[1,2,\ldots,n_1 \mid w_1] > \text{per } A_1$ and $\text{per } A_2[1,2,\ldots,n_2 \mid w_2] > \text{per } A_2$.

If $\bar{w}_2 = w_2 + (n_1, n_1, \ldots, n_1)$, then, clearly, per $(A_1 \oplus A_2)[1,2,\dots,n_1+n_2 \mid w_1, \bar{w}_2]$ per $(A_1 \oplus A_2)$ = per A_1 per A_2 for per $(A_1 \oplus A_2)[1,2,\dots,n_1+n_2 \mid w_1,\bar{w}_2]$ = per $A_1[1,2,\dots,n_1 \mid w_1]$. per $A_2[1, 2, \dots, n_2 \mid w_2]$.

DEFINITION 4.59 : If $A^T \in \bar{D}_n$ let us call A row-dominating matrix or simply rod. If A is both row-dominating and column dominating let us call it doubly dominating or simply dod.

J_n and permutation matrices are clearly dod.

4.60 CLASSES OF MATRICES SATISFYING VAN DER WAERDEN CONJECTURE :

Various subsets of D_n that satisfy vander waerden conjecture were given by a number of authors. The most-important subset is the set of all positive semi-definite symmetric doubly stochastic matrices treated extensively in the pioneering works of Marvin Marcus. Sasser and Slater extended this class to the set of all normal doubly stochastic matrices with numerical range in $[-\frac{\pi}{2n}, \frac{\pi}{2n}]$. Friedland extended this still further to doubly stochastic matrices with numerical range in $[-\frac{\pi}{2n}, \frac{\pi}{2n}]$.

The collection of cod and rod matrices seems to be a distinct class of matrices satisfying vander waerden conjecture. The structure of this class has not been throughly determined. In particular the case of equality in per $(AB) \leq$ (per A, per B) for $A, B^T \in \bar{D}_n$, has not at all been touched.

C H A P T E R - V

APPLICATION OF PERMANENTS TO GRAPH THEORY

5.0 INTRODUCTION :

In graph theory a complete set of invariants of graphs in likely to be of immense help and hence attempts have been made to get such a set. While some trivial sets are certainly there a useful set has not so far been found. Harary proposed characteristic polynomial or equivalently the spectrum of the adjacency matrix of a graph as a complete set of invariants. But quickly non-isomorphic graphs with the same spectrum have been found. Indeed quite a lot of work Has been done on cospectral graphs. In this chapter we analyse the reasons for failure of characteristic polynomial to be a complete invariant for a graph and give a modified polynomial which is likely to characterize graphs. Incidentlly some interesting results are got. Through brute verification by hand the polynomial was found by the author to characterize graphs with atmost six vertices. Through the use computer even graphs with seven vertices were found to be characterized by the proposed polynomial. This chapter deals with the attempt to find the polynomial and the prosperties of the polynomial.

5.1 PRELIMINARY IDEAS :

A graph G is a pair (V,E) where V is a finite non-empty set (elements of V are called vertices) and E is a subset of the set of all unordered pairs of distinct elements of V (the elements of E are called edges). E is allowed to be empty. Let $|V| = p$ and $|E| = q$.

A graph G can be completely characterized by its adjacency matrix $A = (a_{ij})$, a $p \times p$, (0,1) matrix with leading diagonal elements zero and other elements described as follows. Suppose

$V = \{v_1, v_2, \ldots, v_p\}$. Then, for $i \neq j$, $a_{ij} = 1$ if $(v_i, v_j) \in E$ and $a_{ij} = 0$ if $(v_i, v_j) \notin E$.

Eventhough this preliminary idea of graph is given we are going to take for granted quite a number of standard results from graph theory as and when necessary. Algebraic graph theory essentially starts with the adjacency matrix of a graph. An important problem in graph theory is the construction of a complete set of invariants for a graph. A complete set of invariants for a graph may be defined a collection of numbers (ordered or unordered) such that a graph uniquely determines the collection and the collection uniquely determines the graph. In otherwords non-isomorphic graphs must have different collections and isomorphic graphs must have the same collection. Such a complete set of invariants will also solve the graph isomorphism problem.

To this end, Harary proposed the characteristic polynomial of the adjacency matrix viz. $\det (x I - A) = \sum_{r=0}^{p} a_r x^r$, as the complete invariant polynomial for a graph, in the sense that $(a_0, a_1, \ldots, a_p)$ constitutes a complete set of invariants for a graph. But, unfortunately, the characteristic polynomial was easily shown, by counterexamples, to be an 'incomplete' invariant polynomial. In fact the term "cospectral graphs" denoting non-isomorphic graphs with the same characteristic polynomial or equivalently the same "spectrum" has become a standard term in graph theory literature.

Let us analyse various candidates for the complete invariants for a graph. If A is the adjacency matrix of a graph then det A and Per A are unsuitable candidates as they will be zero for "sparse" adjacency matrices and thus cannot distinguish graphs with very small number of edges. Nevertheless these are certainly invariants. What about the characteristic polynomial? Consider the matrix xI-A. The presense of a large number of zeros in this matrix will adversely affect the determinant of the matrix due to considerable loss of information. Another factor is the cancellation of a considerable number of terms due to the rules for signs of each term in a determinant. To obviate this difficulty we can consider permanent of xI-A. But still the problem of zeros remains. Thus we are led to a modified matrix. Suppose $A=(a_{ij})$ is the adjacency matrix of a graph G. Let x be a real indeterminate (assumed to take only positive values). Define $A(x) = (x^{a_{ij}})$. The elements of this matrix are 1 and x. Thus all the zeros are conveniently removed. In fact A(x) is got from A by replacing all the zeros by 1 and 1's by x. Will per A(x) be a complete invariant polynomial? From the complete list of all graphs with atmost six vertices given in Harary's book "Graph Theory" per A(x) was calculated. Interestingly enough per A(x) is CERTAINLY A COMPLETE INVARIANT POLYNOMIAL for graphs with atmost 4 vertices. This was actually verified by brute calculations. For graphs with more than 4 vertices calculation of permanent, a badly behaving function, is quite tedious and hence search for a counter example by brute calculation was given up and theoretical attack was started.

What are the symmetries of the permanent function and a graph? The adjacency matrices A and B of two isomorphic graphs are related by an equality, $B = PAP^T$ for a suitable permutation matrix P. In fact such an equality is a necessary and sufficiency condition for isomorphism of graphs represented by A and B. It is easy to see that

5.2... $B = PAP^T \Longleftrightarrow B(x) = PA(x)\, P^T$ for every $x \in (0, \infty)$.

Also Per B(x) = Per $(P\, A(x)\, P^T)$ = Per A(x) and this shows that for isomorphic graphs the polynomials per A(x) indeed coincide. But unfortunately per (PA(x)Q) = per A(x) for any two permutation matrices P and Q of suitable order even if $Q \neq P^T$. Thus if A and PAQ, where $Q \neq P^T$, represent adjacency matrices of two non-isomorphic graphs G_1 and G_2 respectively then Per A(x) cannot distinguish G_1 and G_2.

Thus we are led to the problem: can A and PAQ represent adjacency matrices of two non-isomorphic graphs G_1 and G_2?. Clearly $QPAQQ^T = QPA$ is also an adjacency matrix of G_2 if PAQ is one. As product of two permutation matrices is again a permutation matrix we are led to the problem:

5.3... Can A and PA represent adjacency matrices of non-isomorphic graphs?

THEOREM 5.4 : Suppose A and PA are adjacency matrices of graphs G_1 and G_2 respectively. If P, considered as representing a permutation, is of odd order then G_1 and G_2 must be isomorphic.

PROOF : If PA is an adjacency matrix, then it must be symmetric and hence $PA = A^T P^T = AP^T$ for A is also symmetric. Thus $A = PAP$ from which we get $A = PAP = P^2AP^2 = \ldots = P^rAP^r = \ldots$ for all positive integral r.

Let $P^{2m-1} = I$ for some positive integer m. Such an m must exist as P is of odd order. Then $A = P^m AP^m$ and $P^m(P^mAP^m)(P^m)^T = P^{2m} A = PA$. Thus A and PA are adjacency matrices of two isomorphic graphs.

INTERPRETATION :

By permuting rows of an adjacency matrix of a graph we cannot get another adjacency matrix of a non-isomorphic graph if the permutation is of odd order.

THEOREM 5.5 : Suppose A and PA are adjacency matrices of graphs on n vertices. *If P represents a cycle of length n, then $A \equiv 0$.*

PROOF : As in the proof of Theorem 5.4 we have $A = P^r AP^r$ for all positive integral r. Let $P = (\delta_{is(j)})$ where s is a complete cycle of length n.

$A = PAP$ gives $a_{ij} = \sum_k \sum_l \delta_{is(k)}\, a_{kl}\, \delta_{ls(j)} = a_{s^{-1}(i)s(j)}$ for every (i,j). In a similar manner, from $A = P^r AP^r$ we get,

5.6... $$a_{ij} = a_{s^{-r}(i)s^r(j)} \quad \text{for every } (i,j).$$

Also $(PA)_{ij} = \sum_k \delta_{is(k)}\, a_{kj} = a_{s^{-1}(i)j}$. Hence,

$a_{s^{-1}(i)j} = a_{s^{-1}(j)i} = a_{is^{-1}(j)}$ for every (i,j) by symmetry of A and PA. In particular we must have

5.7... $$0 = (PA)_{ii} = a_{s^{-1}(i)i} = a_{is(i)} \text{ for each } i.$$

Hence using 5.6 we get,

5.8... $$0 = a_{ii} = a_{s^{-r}(i)s^r(i)}\ .\quad \text{Thus } a_{is^{2r}(i)} = 0 \text{ for each } i.$$

Also $a_{s^{-r}(i)\,s^r(j)} = 0$ if $s(s^{-r}(i)) = s^r(j)$ using 5.8. Thus we have,

5.9... $$a_{is^{-2r+1}(i)} = 0 \quad \text{or} \quad a_{is^{2r-1}(i)} = 0 \text{ for each } i.$$

Combining 5.8 and 5.9 we get $a_{i\,s^r(i)} = 0$ for each i and r. If s is a complete cycle of length n then for every j there exists an r such that $s^r(i) = j$. Thus $a_{ij} = 0$ for every (i,j). Thus $A \equiv 0$.

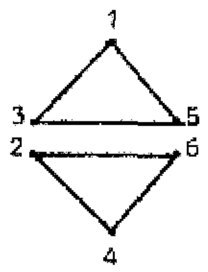
NOTE : Even if s is not a complete cycle we have $a_{is^r(i)} = 0$ for every i and r. Suppose s, as a permutation of $(1, 2, \ldots, n)$ is expressible as a product of k disjoint cycles. Let $(i_1, i_2, \ldots, i_r)$ be a typical cycle. Then $a_{i_a i_b} = 0$ for $a \neq b$ and $a,b = 1, 2, \ldots, r$ in view of the fact $a_{is^r(i)} = 0$ for every i and r. Thus A has an $r \times r$ principal zero submatrix. Considering all the k cycles it is clear that A must be the adjacency matrix of a k-partite graph. Thus we have proved.

THEOREM 5.10 : *If A and PA are adjacency matrices of graphs G_1 and G_2 respectively and P, considered as a permutation, is the product of k disjoint cycles then G_1 and G_2 are k-partite graphs.*

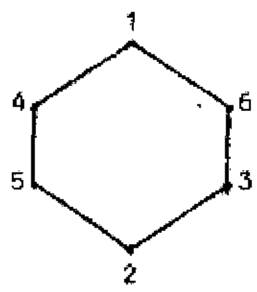
COUNTEREXAMPLE 5.11 :

Theorem 5.10 is quite interesting. But still there is no guarantee that G_1 and G_2 are isomorphic. In fact they need not be isomorphic as the following example shows.

G_1



G_2 =



Both are 3-partite. Their adjacency matrices are

$$A = \left[\begin{array}{cc|cc|cc} 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{array}\right] \qquad B = \left[\begin{array}{cc|cc|cc} 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$\text{Take } P = \left[\begin{array}{cccccc} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

Then $B = PA$.

Here P corresponds to the permutation

$(1\ 2)\ (3\ 4)\ (5\ 6)$.

This counter example shows that Per $A(x)$ is NOT a complete invariant polynomial as G_1 and G_2 have the same polynomials though they are non-isomorphic.

The failure is mainly due to the fact that the zeros in the main diagonal and zeros in non-diagonal positions have distinct meaning, as far as graphs are concerned, but are treated alike in per $A(x)$. Thus there is a clear need for distinguishing diagonal zeros and non-diagonal zeros. To this end we define a new matrix from the adjacency matrix.

DEFINITION 5.12 :

If A is an adjacency matrix of a graph, let $A(x,y) = (\bar{a}_{ij})$ where $\bar{a}_{ii} = x$, $\bar{a}_{ij} = y$ if $a_{ij} = 1$ and $\bar{a}_{ij} = 1$ if $a_{ij} = 0$ and $i \neq j$. In other words

5.13... $$A(x,y) = xI + yA + (J - A - I)$$

where I is the unit matrix and J is a matrix with each entry unity.

Alternatively $A(x,y) = (x-1)\, I + (y-1)\, A + J$. We now propose a conjecture.

CONJECTURE 5.14 :

For a graph with adjacency matrix A, per $A(x,y)$ is a complete invariant polynomial (in two variables x and y).

Note that we do not have the possibility that $P\, A(x,y) = B(x,y)$ for a permutation matrix $P(\neq I)$. For, if $B(x,y) = PA(x,y)$ then

$$(x-1)\, I + (y-1)\, B + J = P \left\{ (x-1)\, I + (y-1)\, A + J \right\}$$
$$\text{or } (x-1)\, I + (y-1)\, B = (x-1)\, P + (y-1)\, PA.$$

Equating coefficients of x and y we must have $I = P$ and $B = PA$ and thus $A = B$ and $P = I$.

DEFINITION 5.51 : Let $F(G;\, x,y) = \text{per } A(x,y) = \sum_{r=0}^{n} \sum_{s=0}^{r} C_{rs}\, x^{n-r}\, y^{s}$

Then the conjecture means that $\left\{ C_{rs} \mid 0 \leqslant s \leqslant r \leqslant n \right\}$ constitute a complete collection of invariants for a graph G with adjacency matrix A.

We can make the polynomial homogeneous by introducing another variable Z by defining

5.16... $A(x,y,z) = xI + yA + Z(J-A-I)$

$$= (x-z)\, I + (y-z)\, A + zJ.$$

Then $F(G;\ x,\ y,\ z) = \sum_{r=0}^{n} \sum_{s=0}^{r} C_{rs}\, x^{n-r}\, y^{s}\, z^{r-s}.$

THEOREM 5.16 : If $\overline{G}$ is the complement of the graph G, then

$$F(\overline{G};\ x,\ y,\ z) = F(G;\ x,\ z,\ y)$$

PROOF : If A is the adjacency matrix of G, then J-A-I is the adjacency matrix of $\overline{G}$ and thus,

$F(\overline{G};\ x,\ y,\ z) = \mathrm{Per}\left\{xI + y(J-A-I) + zA\right\} = F(G;\ x,\ z,\ y).$

COROLLARY 5.17 : $C_{r,s}(\overline{G}) = C_{r,r-s}(G).$

Thus if we know the invariant polynomial for G then that of $\overline{G}$ is also immediately known. As a consequence, for a self complementary graph G, we have $F(G;\ x,\ y,\ z) = F(G;\ x,\ z,\ y)$. In other words the polynomial of G must be symmetric in y and z. No such property of characteristic polynomial is known for a general graph.

5.18 SOME ARGUMENTS IN FAVOUR OF THE CONJECTURE :

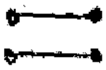
It is quite easy to verify the following results.

(a) $C_{2,2}$ is the number of edges in the graph;

(b) $C_{3,3}$ is twice the number of triangles in the graph.

(a') $C_{2,0}$ is the number of edges in the complementary graph;

(b') $C_{3,0}$ is twice the number of triangles in the complementary graph;

(c) $C_{4,4}$ is number of subgraphs (not necessarily induced) of the form  plus twice the number of squares in the graph.

(d) $C_{5,5}$ = 2 $\{$ number of subgraphs of the form  plus number of subgraphs of the form  $\}$

(c') and (d') are $C_{4,0}$ and $C_{5,0}$ that are similar to (c) and (d) but relate to the complementary graph.

Using the invariants (a), (b), (c), (d), (a'), (b'), (c'), (d') fixed by the permanent polynomial we are able to verify the permanent polynomial COMPLETELY CHARACTERIZES graphs with atmost 6 vertices. This was actually carried out on the basis of the complete list of graphs with atmost 6 vertices given by Harary in his book "Graph Theory".

The above verification together with verification of the conjecture for graphs with 7 vertices using computer in I.I.T., Madras shows that the permanent polynomial is much better than the characteristic polynomial in distinguishing graphs. But permanent polynomial is rather more difficult to calculate than the characteristic polynomial. There seems to be no general method of finding the characteristic polynomial of $\overline{G}$ from that of G if G is not regular. But permanent polynomial is superior in this respect.

5.19 RELATIONS AMONG C_{ij}'s :

The C_{ij}'s are not all completely independent. We shall establish a few relations among them.

$$F(G;\ x,y,z) = \sum_i \sum_j C_{ij}\, x^{n-i}\, y^j\, z^{i-j} = \text{Per}\ (xI + yA + z(J-I-A))$$

$$F(G;\ x,\ y,\ y) = \sum_i \sum_j C_{ij}\, x^{n-i}\, y^i = \text{per}\ (xI + y(J-I))$$

$$= \text{per}\ (yJ + (x-y)\ I)$$

$$= \sum_{r=0}^{n} \binom{n}{r}\ r!\ y^r (x-y)^{n-r}.$$

$$\text{Thus } \sum_j C_{ij} = \sum_{r=0}^{i} \binom{n}{r}\ r!\ \binom{n-r}{n-i}\ (-1)^{i-r} = \frac{n!}{(n-i)!} \sum_{r=0}^{i} \frac{(-1)^{i-r}}{(i-r)!}$$

This gives (n+1) linear equalities satisfied by C_{ij}'s.

5.20 CONSEQUENCE OF SYMMETRY OF A :

Consider the coefficient of $x^{n-2}\ y\ z$ namely C_{21}. x^{n-2} can be got only from (n-2) diagonal terms. Thus yz should be got from a 2 x 2 principal submatrix. Due to symmetry we can get only y^2 or z and hence $C_{21} = 0$. It is easily seen that $C_{21} = 0$ is also sufficient for a (0, 1) matrix A with zero diagonal to be symmetric. $C_{21} = 0$ is another equality apart from the (n+1) already got.

5.21 AN INTERESTING RESULT :

Suppose A is the adjacency matrix of a graph G with n vertices. Let $\widetilde{A} = xI + yA + z(J-I-A)$. Then let us define $S_r(\widetilde{A})$

as the sum of the permanent of all r x r submatrices of $\bar{A}$ for $r = 1, 2, \ldots, n$. Let $\nabla = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}$.

THEOREM 5.22 : $S_{n-1}(\bar{A}) = \nabla \mathrm{per}(\bar{A}) = \nabla F(G; x, y, z)$.

PROOF : $\frac{\partial F}{\partial x}$ gives the sum of the permanental cofactors of the diagonal terms ; $\frac{\partial F}{\partial y}$ gives the sum of the permanental cofactors of all the positions occupied by y; $\frac{\partial F}{\partial z}$ gives the sum of the permanental cofactors of all the positions occupied by z. The fact that every position in $\bar{A}$ is occupied by exactly one element from $\{x, y, z\}$ completes the proof.

COROLLARY 5.23 : $S_{n-r}(\bar{A}) = \frac{1}{r!} \nabla^{r} F$ for $r = 0, 1, 2, \ldots, (n-1)$.

PROOF : The only thing to be proved is the factor r!. If we take r elements in independent positions in $\bar{A}$ then we can take rth derivative w.r.t. x, y, z in r! ways and hence $\nabla^{r} F = r!\, S_{n-r}(\bar{A})$.

If we define $S_0(\bar{A}) = \frac{1}{n!} \nabla^{n} F$ then 5.23 holds even for $r = n$. But it is clear that $S_0(\bar{A}) = 1$.

5.24 ADDING AN ISOLATED POINT TO A GRAPH :

Suppose G is a graph on n points with adjacency matrix A. Let us add an isolated point to G and get G_1. If B in the adjacency matrix of G_1 it can be written in the form

$$B = \left[\begin{array}{c|c} 0 & 0_{1\times n} \\ \hline 0_{n\times 1} & A \end{array}\right] \qquad \text{Thus } \bar{B} = \left[\begin{array}{c|c} x & zJ_{1\times n} \\ \hline zJ_{n\times 1} & \bar{A} \end{array}\right]$$

where $J_{r \times s}$ is an $r \times s$ matrix with each entry unity.

Clearly per $(\bar{B}) = x$ per $(\bar{A}) + z^2 S_{n-1}(\bar{A})$ (expanding by first row).

If per $(\bar{A}) = F$ and per $(\bar{B}) = F_1$ then we get

$$5.25\ldots \qquad F_1 = xF + z^2 \nabla F = (x + z^2 \nabla) F$$

COROLLARY 5.26 : If we add a point to G and join this point to all the vertices of G and get G_2 then the permanent polynomial F_2 of G_2 is given by $F_2 = (x + y^2 \nabla) F$.

Suppose G_1 and G_2 are two graphs. Let $G_1 \bigcup G_2$ represent the disjoint union of these two graphs. Let K_n represent the complete graph on n points. Let $G_1 + G_2$ represent the graph obtained by taking disjoint union of G_1 and G_2 and joining every vertex of G_1 with every vertex of G_2. If G^c represents the complement of G. Then clearly $G_1 + G_2 = (G_1^c \bigcup G_2^c)^c$. With these notations we have proved.

$$5.27\ldots \qquad F_{G \bigcup K_1} = (x + z^2 \nabla) F_G$$

$$5.28\ldots \qquad F_{G+K_1} = (x + y^2 \nabla) F_G$$

In general we have $F_{GU(K_1UK_1U \ldots r \text{ terms})} = (x + z^2 \nabla)^r F_G$

$$F_{G+K_n} = (x + y^2 \nabla)^n F_G$$

$$F_{K_n} = (x + y^2 \nabla)^{n-1} x$$

$$F_{K_1UK_1U \ldots n \text{ terms}} = (x + z^2 \nabla)^{n-1} x.$$

5.29 A MODIFIED POLYNOMIAL :

Let A be any $n \times n$ adjacency matrix. Let B be any $n \times n$ real matrix. Let $S_r(B)$ represent the sum of all the permanents of all $r \times r$ submatrices of B for $1 \leqslant r \leqslant n$. Let C_r be the collection of all $r \times r$ principal submatrices of A. Define

5.30... $$S_{r,s}(A) = \sum_{B \in C_r} S_s(B) \quad \text{for } 1 \leqslant s \leqslant r \leqslant n.$$

For any function $f(x, y, z)$ of x, y, z write $\left[f(x, y, z)\right] = f(1,1,1)$. Then it is quite easy to see that $S_{r,s}(A)$ is related to the function $\text{per}(\bar{A}) = F(G, x, y, z)$ in the following manner

5.31... $$S_{r,s}(A) = \frac{1}{(n-r)!\,s!\,(r-s)!}\left[\frac{\partial^{n-r+s} F}{\partial x^{n-r}\,\partial y^{s}}\right]$$

By Taylor's Theorem we also have

5.32... $$F(x,y,1) = \sum_r \sum_s \frac{1}{r!}\,\frac{1}{s!}\left[\frac{\partial^{r+s} F}{\partial x^{r}\,\partial y^{s}}\right](x-1)^r (y-1)^s$$

Thus, from 5.31 and 5.32 we get

5.33... $$F(x,y,1) = \sum_{r=0}^{n} \sum_{s=0}^{n-r} \frac{1}{r!}\,\frac{1}{s!}\,S_{n-r,s}(A)\,r!\,s!\,(n-r-s)!\,(x-1)^r (y-1)$$

$$= \sum_{r=0}^{n} \sum_{s=0}^{n-r} (n-r-s)!\; S_{n-r,s}(A)\,(x-1)^r (y-1)^s.$$

$F(x,y,z)$ can be got from $F(x,y,1)$ by multiplying each term in $F(x,y,1)$ by a suitable power of z to make its degree n. Thus F is completely determined by $S_{r,s}(A)$ for $1 \leq s \leq r \leq n$. Note that in the expression of $F(x,y,1)$ in powers of $(x-1)$ and $(y-1)$ to be valid we have to <u>define</u> $S_{r,0} = \binom{n}{r}$ for $r = 0,1,2, \ldots, n$.

It can also be seen that $S_{r,s}(A) \frac{1}{(r-s)!} \left[\nabla^{r-s} \left(\frac{\partial^{n-r} f}{\partial x^{n-r}} \right) \right]$ the conjecture on complete set of invariants takes the following form.

5.34... If A is the adjacency matrix of a graph G then.

$\left\{ S_{r,s}(A) \mid 1 \leq s \leq r \leq n \right\}$ completely characterizes the graph G.

Unfortunately the polynomial $F(x, y, 1) = \sum_{r} \sum_{s} (n-r-s)! \cdot S_{n-r,s}(A)\,(x-1)^r\,(y-1)^s$ is rather inconvenient to use. Clearly
$F(x+1, y+1, 1) = \sum_{r} \sum_{s} (n-r-s)!\, S_{n-r,s}(A)\, x^r y^s$,

is better for algebraic manipulation. If this is homogenised by introducing z the polynomial would be better and we may also drop the factor $(n-r-s)!$

Essentially what we are suggesting is the following polynomial

$$\sum_{r=0}^{n} \sum_{s=0}^{n-r} S_{n-r,s}(A)\, x^r y^s z^{n-r-s} \quad \text{or equivalently,}$$

5.35... $$f(x, y, z) = \sum_{r=0}^{n} \sum_{s=0}^{r} S_{r,s}(A)\, x^{n-r} y^s z^{r-s}.$$

Let us call $f(x, y, z)$ as homogeneous permanent polynomial (H.P.P.). How is this related to $\bar{A}$? Suppose w is the linear operator on the vector space of all polynomials in Z, with coefficients polynomials in (x,y), defined by the property $w(z^r) = \frac{1}{r!} z^r$ for r = 0, 1, 2, ... clearly this is an invertible linear operator with inverse w^{-1} defined by $w^{-1}(z^r) = r!\, z^r$.

Let $\overline{\overline{A}} = (x+z)I + (y+z)A + Z(J-A-I) = xI + yA + zJ$.

Then $f(x,y,z) = w(\text{per } \overline{\overline{A}})$. But per $(\overline{\overline{A}}) = F(x+z, y+z, z)$ if $F(x,y,z) = \text{per } \overline{\overline{A}}$. Hence we have

5.36... $f(x,y,z) = w\left\{F(x+z, y+z, z)\right\} \equiv w\,\overline{F}$ (say).

In what way is f better than F? This is answered by the remarkable multiplicative property of f.

THEOREM 5.37 : Let G_1 and G_2 be two graphs with H.P.P. f_{G_1} and f_{G_2} respectively. Then $f_{G_1 \cup G_2} = f_{G_1} \cdot f_{G_2}$.

PROOF : Suppose A_1 and A_2 are adjacency matrices of G_1 and G_2 respectively. Then the adacency matrix of $G_1 \cup G_2$ is $A_1 \oplus A_2$, the direct sum of A_1 and A_2. It is easy to see that the following convolution formula holds.

5.38... $S_{r,s}(A_1 \oplus A_2) = \sum_{k=0}^{r} \sum_{l=0}^{s} S_{k,l}(A_1)\, S_{r-k,s-l}(A_2)$.

In the above convolution any meaningless term is interpreted as zero. If A is an n x n adjacency matrix $S_{r,s}(A)$ is meaningful only for $0 \leqslant s \leqslant r \leqslant n$.

$$f_{G_1} = \sum_r \sum_s S_{r,s}(A_1)\, x^{n_1 - r}\, y^s\, z^{r-s}$$

$$f_{G_2} = \sum_r \sum_s S_{r,s}(A_2)\, x^{n_2 - r}\, y^s\, z^{r-s}$$

where n_1 and n_2 are the number of vertices in the graphs G_1 and G_2 respectively.

$$f_{G_1 \cup G_2} = \sum_r \sum_s S_{r,s}(A_1 \oplus A_2)\, x^{n_1+n_2}\, y^s\, z^{r-s}$$

$$= f_{G_1} \cdot f_{G_2}$$

in view of the convolution formula for $S_{r,s}(A_1 \oplus A_2)$

COROLLARY 5.39 : (a) $f_{\bigcup_{i=1}^{k} G_i} = \prod_{i=1}^{k} f_{G_i}$

clearly $f_{K_1} = w\left\{\text{per}\,(xI + zJ)\right\} = w(x+z) = x+z$. Thus if $\bigcup_{i=1}^{k} K_1 = nK_1$ we get $f_{nK_1} = (x+z)^n$;

(b) $f_{nK_1 \cup G} = (x+z)^n f_G$;

(c)
$$\begin{aligned} \text{Per}\,(J-I) &= \text{Per}\,(xI + zJ) \quad \text{at } x=-1,\ z=1 \\ &= \left\{w^{-1}(f_{nK_1})\right\} \quad \text{at } x=-1,\ z=1 \\ &= \left\{w^{-1}(x+z)^n\right\} \quad \text{at } x=-1,\ z=1 \\ &= w^{-1} \sum \binom{n}{r} x^r z^{n-r} \text{ at } x=-1,\ z=1 \\ &= \sum_{r=0}^{n} \binom{n}{r} (n-r)!\,(-1)^r = n! \sum_{r=0}^{n} \frac{(-1)^r}{r!}. \end{aligned}$$

This is the well known formula for the number of derangements of n objects.

Another property of f which justifies its superiority over F is the following theorem. If g and h are homogeneous polynomials of degree n in x,y,z let us write $g \leqslant h$ if all the corresponding coefficients satisfy the same inequality.

THEOREM 5.40 : If H is a spanning subgraph of G, then $f_H \leqslant f_G$ and equality holds if and only if H is isomorphic to G.

PROOF : Suppose H is a spanning subgraph of G, then for suitable labelling of H and G their adjacency matrices A_H and A_G satisfy the inequality $A_H \leqslant A_G$ (i.e., $A_H[i|j] \quad A_G[i|j]$ for every (i,j)). Hence, for any r,s $(s \leqslant r)$ $S_{r,s}(A_H) \leqslant S_{r,s}(A_G)$ and consequently

$$\sum_r \sum_s S_{r,s}(A_H)\, x^{n-r}\, y^s\, z^{r-s} \leqslant \sum_r \sum_s S_{r,s}\,(A_G)\, x^{n-r}\, y^s\, z^{r-s}.$$

Obviously equality holds if and only if $S_{r,s}(A_H) = S_{r,s}(A_G)$ for all r,s. In view of the fact that $A_H \leqslant A_G$ this can happen if and only if $A_H = A_G$. This proves the theorem.

NOTE : Suppose H and G are graphs with the same number of vertices and $f_H \leqslant f_G$ Theorem 5.40 does not claim or prove that H is a spanning subgraph of G.

Consider G_n, the set of all graphs on n vertices. We can partially order G_n by writing $H \leqslant_1 G$ if and only if H is isomorphic to a spanning subgraph of G. Equality in G_n is isomorphism of graphs. Let us also consider P_n, the f-polynomials of graphs in G_n. We can partially order P_n by requiring $f_H \leq_2 f_G$ if and only if $f_H \leq f_G$. In view of Theorem 5.40 we can conjecture that there is a homomorphism of G_n into P_n as partially ordered sets (i.e., a mapping of G_n into P_n which preserve partial order). In fact this homomorphism will turn out to be isomorphism if our conjecture is true.

5.41 ANOTHER FORM OF THE CONJECTURE :

Let A represent the adjacency matrix of a graph on n vertices. An s-chain in A is a collection of s 1's in independent positions (two positions are said to be independent in a matrix if they are in different rows and different columns). Let $C_s(A)$ be the collection of all s-chains in A. For any element $a \in C_s(A)$, let R(a), the rank of a, be the number of rows in the smallest principal submatrix of A containing all the positions occupied by a. Clearly $R(a) \geq s$ if $a \in C_s(A)$. If $a \in C_s(A)$, then a contributes one to $S_{R(a),s}$ and in general contributes $\binom{n-R(a)}{r-R(a)}$ to $S_{r,s}$. This is quite obvious from the definition of $S_{r,s}$. Thus $S_{r,s}(A) = \sum_{a \in C_s(A)} \binom{n-R(a)}{r-R(a)}$. In particular $S_{s,s}(A) = \sum_{a \in C_s(A)} \binom{n-R(a)}{r-R(a)}$. But $R(a) \geq s$ for $a \in C_s(A)$. Hence $S_{s,s}(A)$ is the number of s-chains with rank exactly equal to s.

Suppose $t_{r,s}$ is the number of elements of $C_s(A)$ with rank exactly equal to r. Then obviously we have, 5.42... $$S_{r,s}(A) = \sum_{u=s}^{r} \binom{n-u}{r-u} t_{u,s} .$$

Thus $t_{r,s}$'s completely determine $S_{r,s}$'s. The converse is also true. For 5.42 is equivalent to

5.43... $$\begin{bmatrix} S_{s,s} \\ S_{s+1,s} \\ \cdot \\ \cdot \\ \cdot \\ S_{n,s} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \binom{n-s}{1} & \binom{n-s-1}{0} & 0 & \cdots & 0 \\ \binom{n-s}{2} & \binom{n-s-1}{1} & \binom{n-s-2}{0} & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \binom{n-s}{n-s} & \binom{n-s-1}{n-s-1} & \binom{n-s-2}{n-s-2} & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} t_{s,s} \\ t_{s+1,s} \\ \cdot \\ \cdot \\ \cdot \\ t_{n,s} \end{bmatrix}$$

The $(n-s+1) \times (n-s+1)$ matrix is clearly non-singular and $S_{r,s}$'s completely determine $t_{r,s}$'s. Thus our conjecture takes the following form.

CONJECTURE 5.44 : A graph is completely characterized by $t_{r,s}$'s where $t_{r,s}$ is the number of s-chains with rank r in the adjacency matrix of the graph.

LEMMA 5.45 : If $a \in C_s(A)$ then $s \leqslant R(a) \leqslant 2s$.

PROOF : $R(a) \geqslant s$ is quite obvious. Suppose I is the set of row suffixes of elements in the s-chain a and J, the set of column suffixes.

Then $|I| = |J| = s$. Also $|I \cup J| = R(a)$. Hence
$$R(a) = |I \cup J| \leqslant |I| + |J| = 2s.$$

Let us now consider the maximum number of non-zero $t_{r,s}$'s clearly $s \leq r \leq 2s$ and $1 \leq s \leq n$. Thus the combination of r, s for which $t_{r,s}$ may be non-zero is $N = \left\{(r,s) \mid 1 \leq s \leq n,\ s \leq r \leq \min(2s,n)\right\}$

For $s \leq \frac{n}{2}$ the combinations are (s,s), $(s+1, s)$, ..., $(2s, s)$

For $s > \frac{n}{2}$ the combinations are (s,s), $(s+1, s)$, ..., (n, s).

$$\text{Thus } |N| = \sum_{s \leq [\frac{n}{2}]} (s+1) + \sum_{s > [\frac{n}{2}]} (n-s+1) = [2+3+ \dots + (t+1)] + [1 + 2 + \dots + n-t]$$

where $t = [\frac{n}{2}]$.

$$\text{Thus } |N| = \frac{(t+1)(t+2)}{2} - 1 + \frac{(n-t)(n-t+1)}{2} = \frac{n^2 - n(2t-1) + 2t^2 + 2t}{2}$$

CASE 1 : n is even. Then $t = \frac{n}{2}$ and $|N| = \frac{n(n+4)}{4}$

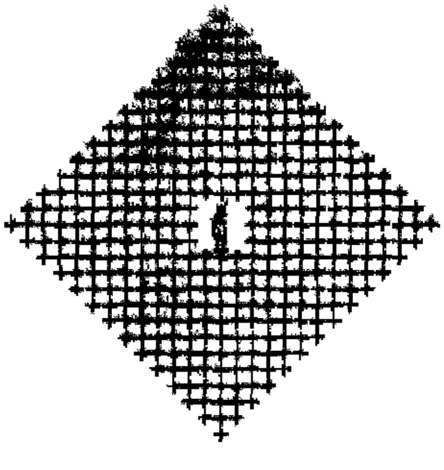
CASE 2 : n is odd. Then $t = \frac{n-1}{2}$ and $|N| = \frac{n(n+4)-1}{4}$

Thus $|N| = \left[\frac{n(n+4)}{4}\right]$.

But the adjacency matrix is determined by $\frac{n(n-1)}{2}$ elements, and $|N| \div \frac{n(n-1)}{2} \doteq \frac{1}{2}$ for large n. Thus, if the conjecture is true, then a graph is characterized by about half the number of elements as compared to the number of elements required to construct the adjacency matrix.

5.46 A PARTICULAR CASE :

Suppose $n = 2m$. What is $t_{2m,m}$? Clearly $a \in C_m(A)$ contributing to $t_{2m,m}$ must have row suffixes and column suffixes disjoint. Thus a typical element will be of the form (i_1, j_1), (i_2, j_2), ..., (i_m, j_m) where all i's and j's are different. But such a term after introducing the 'transposes' (j_1, i_1), (j_2, i_2), ..., (j_m, i_m) will contribute a one-factor of the graph. Thus $t_{2m,m}$ is 2^m times the number of one-factors of a graph. Thus, indirectly, the f-polynomial of a graph determines the number of one-factors.



R E F E R E N C E S

[1] BALASUBRAMANIAN, K. Diagonal sums of doubly stochastic matrices, Sankhya, 1977, Vol.39, Series B.

[2] BALASUBRAMANIAN, K. On equality of some elements in Matrices, Linear Algebra and Appl. 18, 1978.

[3] BALASUBRAMANIAN, K. Constancy of functions restricted to a subset, Linear and Multilinear Algebra, Vol.7, 1979.

[4] BALASUBRAMANIAN, K. Maximal diagonal sums, Linear and Multilinear Algebra, Vol.7, 1979.

[5] BALASUBRAMANIAN, K. Determinants and Permanents with respect to Group Complexes, Vol.V, No.1-2, March 1977, Mathematika Sciences.

[6] CLAUDE BERGE, Topological Spaces, Oliver and Boyd (1963).

[7] DJOKOVIC, D.Z., On a conjecture by Van der Waerden, Mat. Vesnik (19) 4 (1967).

[8] SHMUEL FRIEDLAND, Matrices Satisfying the Van der Waerden Conjecture, Linear Algebra and Appl. 8 (1974).

[9] JURKAT, W.B. and RYSER, H.J., Matrix Factorization of determinants and permanents, J. Algebra, 1966.

[10] DAVID LONDON, Some Notes on the Van der Waerden conjecture, Linear Algebra and Appl. 4 (1971).

[11] MARWIN MARCUS and HENRYK MINC, A Survey of Matrix Theory and Matrix Inequalities - Allyn. and Bacon (1964).

[12] HENRYK MINC, Permanents, Encyclopedia of Mathematics and its Applications, Vol.6 - Addison Wesley, 1978.

[13] MUIR, T., On a class of permanent symmetric function, Proc. Roy. Soc. Edinburgh 11 (1882).

[14] RYSER, H.J., Combinatorial Mathematics, Carus Monograph No.14, Wiley, New York. (1963).

[15] SASSER, D.W., and SLATER, M.L., On the inequality $\sum x_i y_i \geqslant \frac{1}{n} \sum x_i \sum y_i$ and the Van der Waerden Conjecture, J. Combinatorial Theory 3 (1967).

[16] HELGE TVERBERG, On the permanent of a bistochastic matrix Math - Scand 12 (1963).

[17] EDWARD TZU-HSIA WANG, Maximum and Minimum diagonal sums of doubly stochastic matrices, *Linear Algebra and Appl.* 8(6) (1974).

[18] MANN, H.B., Design of Experiments.

*****