On stochastic matrices of a given Frobenius type

By Nobuko IWAHORI

§ 1. Introduction.

A real square matrix $P=(p_{ij})$ of degree n is called stochastic if all p_{ij} are $\geqq 0$ and $\sum_{j=1}^{n} p_{ij}=1$ for $i=1, 2, \cdots, n$. Let us denote by Ω the set $\{1, 2, \cdots, n\}$. With a stochastic matrix $P=(p_{ij})$ we associate a subset Supp (P) of $\Omega\times\Omega$, called the support of P, defined by

$$\mathrm{Supp}\,(P)=\{(i, j)\in\Omega\times\Omega\,;\ p_{ij}\neq 0\}\,.$$

Let us denote by $S(n)$ the set of all stochastic matrices of degree n. The question we shall consider in this paper is the following.

Suppose a subset G of $\Omega\times\Omega$ is given. We denote by $S(n)_G$ the subset of $S(n)$ consisting of all P in $S(n)$ such that Supp $(P)=G$. We assume that $S(n)_G$ is not empty. For each P in $S(n)$, denote by $L^*(P)$ the set of all row vectors $u=(u_1, u_2, \cdots, u_n)$ fixed by P (i.e. $uP=u$) and such that $u_1\geqq 0, \cdots, u_n\geqq 0, \sum_i u_i=1$. Then how can one describe the set $\Delta_G=\bigcup_{P\in S(n)_G} L^*(P)$ in terms of G?

We can reduce this question (see § 7) to the case where G is of Frobenius type, i.e. the case where for any P in $S(n)_G$, Ω consists of a single ergodic class and Ω has no transient class (in the sense of Doob [3, Chap. 5]), or equivalently, every P in $S(n)_G$ is permutation-irreducible in the sense of Frobenius (see Gantmacher [4, Chap. 2]; also see § 2 below.)

Let now G be a subset of Ω^2 of Frobenius type. Our main result is as follows:

(i) Δ_G is a convex, bounded subset of $\boldsymbol{R}^n$. Hence the closure $\bar{\Delta}_G$ of Δ_G in $\boldsymbol{R}^n$ is a compact, convex subset of $\boldsymbol{R}^n$.

(ii) Let E_G be the set of all extreme points of $\bar{\Delta}_G$. Then E_G is a finite set. Let $E_G=\{u_1, \cdots, u_m\}$. Then Δ_G is given by

$$\Delta_G=\{\alpha_1u_1+\cdots+\alpha_mu_m\,;\ \alpha_1>0, \cdots, \alpha_m>0, \sum_i\alpha_i=1\}\,.$$

(iii) There exists a finite set $E_G^*=\{v_1, \cdots, v_k\}$ such that

$$E_G\subset E_G^*\subset\bar{\Delta}_G\,,$$
$$E_G^*=\{\alpha_1v_1+\cdots+\alpha_kv_k\,;\ \alpha_1>0, \cdots, \alpha_k>0, \sum_i\alpha_i=1\}\,.$$

In order to define the set E_G^*, let us introduce a notion of a G-cycle. A cycle over Ω of length m is a sequence $\langle i_1, \cdots, i_m\rangle$ of mutually distinct elements in Ω identified up to cyclic permutations:

$$\langle i_1, i_2, \cdots, i_m\rangle=\langle i_2, i_3, \cdots, i_m, i_1\rangle= \cdots =\langle i_m, i_1, \cdots, i_{m-1}\rangle .$$

The subset $\{i_1, \cdots, i_m\}$ of Ω is called the support of the cycle $C=\langle i_1, \cdots, i_m\rangle$ and is denoted by Supp (C). A cycle $\langle i_1, \cdots, i_m\rangle$ over Ω is a G-cycle if $m=1, (i_1, i_1)\in G$, or $m>1, (i_1, i_2)\in G, (i_2, i_3)\in G, \cdots, (i_m, i_1)\in G$. We denote by $\langle G\rangle$ the set of all G-cycles. For each C in $\langle G\rangle$, we associate a row vector $v_C=(\alpha_1, \cdots, \alpha_n)\in \boldsymbol{R}^n$ by

$$\alpha_i=\begin{cases}\dfrac{1}{m}, & \text{if } i\in \operatorname{Supp}(C)\\ 0, & \text{otherwise}\end{cases}$$

where m is the length of C.

Now we will prove (see Theorem 5 below) that we may take as our set E_G^* the finite set $\{v_C ; C\in\langle G\rangle\}$.

The quantity $\lambda^*(P, C)$ introduced in §9 may be regarded as the "probability of running on the cycle C" in the homogeneous Markov chain represented by P. In fact, for a stochastic matrix P of support G of Frobenius type, the unique fixed row vector $L^*(P)$ is given by the following formula:

$$L^*(P)=\sum_{C\in\langle G\rangle}\lambda^*(P, C)\cdot v_C$$

Our method is as follows: when G is a Frobenius type, $L^*(P)$ consists of a single point for each $P\in S(n)_G$ (see e.g. §7 below). Thus one has a surjective map $L^*: S(n)_G\to\Delta_G$. This map is "linearized" as follows: we construct some convex space of matrices, to be denoted by $T(n)_G$ in §8 and a bijection $\sigma: T(n)_G\to S(n)_G$ such that $\tau=L^*\circ\sigma: T(n)_G\to\Delta_G$ is affine-linear. Then immediately Δ_G is known to be convex.[1)]

We then consider the free $\boldsymbol{R}$-module $\langle G\rangle_{\boldsymbol{R}}$ generated by $\langle G\rangle$ and define a convex subset Θ_G of $\langle G\rangle_{\boldsymbol{R}}$ by

$$\Theta_G=\left\{\sum_{C\in\langle G\rangle}\mu_C\cdot C;\ \mu_C>0 \text{ (for every } C),\ \sum_C\mu_C=1\right\} .$$

Define now a linear map $\psi: \langle G\rangle_{\boldsymbol{R}}\to\boldsymbol{R}^n$ by $\psi(C)=v_C$ (for every $C\in\langle G\rangle$). Our main result is equivalent to $\psi(\Theta_G)=\Delta_G$ and this is proved by constructing a surjective linear map $\rho: \Theta_G\to T(n)_G$ such that $\psi=\tau\circ\rho$ (see §9).

1) We owe this construction of $T(n)_G$ and σ to Professor T. Kato; he pointed out this construction which was vaguely visible in our first version of this work. This construction simplified very much our former version.

We have proved several facts about matrices and determinants as lemmas; one of them is a new (we hope) proof of a theorem of Bott-Mayberry [1] which gives an expansion formula of a determinant in terms of rooted trees in the sense of graph theory.

As an application of the main theorem, we shall settle the following question: let G be a subset of $\Omega\times\Omega$ of Frobenius type. Define n functions $v_1, \cdots, v_n$ on $S(n)_G$ by

$$L^*(P)=(v_1(P), \cdots, v_n(P)) .$$

Then given i and j, where $v_i(P)\geqq v_j(P)$ for all P in $S(n)_G$? In other words, the question is to find "the most important position in a communication network" for all distribution of information-probability. (cf. Kemeny-Snell [5, Chap. 8]). By our main result, we see that $v_i(P)\geqq v_j(P)$ for all P in $S(n)_G$, if and only if every G-cycle through j passes through i. (See Theorem 7 below).

Finally we will give a criterion for the injectivity of the map $L^*\colon S(n)_G\to\Delta_G$. It turns out that L^* is injective if and only if $|G|=n+d(G)$, where $|G|$ is the cardinality of G and $d(G)$ is the dimension of the linear space spanned by the v_C $(C\in\langle G\rangle)$.

The author would like to express her hearty thanks to Professor T. Kato for his crucial remarks which simplified a great deal of the first version of this paper.

§2. Paths, cycles, trees, and pseudo-trees.

Let Ω be the set consisting of positive integers $1, 2, \cdots, n$. We denote by Ω^m the cartesian product $\Omega\times\Omega\times\cdots\times\Omega$ (m factors). An element $(i_1, i_2, \cdots, i_m)$ of Ω^m is called a *path* over Ω of length $m-1$ from i_1 to i_m if $i_1, i_2, \cdots, i_m$ are mutually distinct. In particular, an element i of Ω is a path of length 0.

Now let us denote by π the permutation of Ω^m defined by

$$\pi(i_1, i_2, \cdots, i_m)=(i_2, i_3, \cdots, i_m, i_1) .$$

Then the set Ω_0^m consisting of all paths over Ω is stable under $\pi\colon \pi(\Omega_0^m)=\Omega_0^m$. Thus the cyclic group $Z_m=\{1, \pi, \cdots, \pi^{m-1}\}$ acts on Ω_0^m and we get a partition of Ω_0^m into Z_m-orbits. A Z_m-orbit thus obtained is called a *cycle* over Ω of length m. For a path $(i_1, i_2, \cdots, i_m)$, we denote by $\langle i_1, i_2, \cdots, i_m\rangle$ the cycle containing $(i_1, i_2, \cdots, i_m)$. Thus we have

$$\langle i_1, i_2, \cdots, i_m\rangle=\langle i_2, i_3, \cdots, i_m, i_1\rangle=\cdots=\langle i_m, i_1, \cdots, i_{m-1}\rangle .$$

Let $(j_1, \cdots, j_k)$ be a path over Ω of length $k-1$ with $k\leqq m$. We denote by

$C^{(m)}(j_1, \cdots, j_k)$ the set of all cycles $\langle i_1, i_2, \cdots, i_m\rangle$ over Ω of length m such that $j_1=i_1, \cdots, j_k=i_k$. We denote by $C(j_1, \cdots, j_k)$ the union of all $C^{(m)}(j_1, \cdots, j_k)$, $(m=k, k+1, \cdots, n)$.

LEMMA 1. *$C(i)$ is a disjoint union of $C(i, 1), \cdots, C(i, i-1), C(i, i+1), \cdots, C(i, n)$ and $\langle i\rangle$. Also $C(i)$ is a disjoint union of $C(1, i), \cdots, C(i-1, i), C(i+1, i), \cdots, C(n, i)$ and $\langle i\rangle$.*

PROOF. Obvious.

Let G be a subset of Ω^2. A path $(i_1, \cdots, i_m)$ is called a G-path if $(i_k, i_{k+1})\in G$ for $k=1, \cdots, m-1$. A cycle C of length >1 is called a G-cycle if every path in C is a G-path. Thus $\langle i_1, \cdots, i_m\rangle$ $(m>1)$ is a G-cycle if and only if

$$(i_1, i_2)\in G, (i_2, i_3)\in G, \cdots, (i_m, i_1)\in G .$$

A cycle $\langle i\rangle$ of length 1 is called a G-cycle if (i, i) is in G. Let $(j_1, \cdots, j_k)\in \Omega_0^k$. Then we denote by $C_G(j_1, \cdots, j_k)$ the set of all G-cycles contained in $C(j_1, \cdots, j_k)$.

Let T be a subset of Ω^2 and let i be an element of Ω. T is called an *oriented rooted tree over i* if

(ORT_1) for any element j in $\Omega-\{i\}$, there exists a unique T-path over Ω from j to i, and

(ORT_2) there is no T-cycle.

We denote by $\mathfrak{T}(i)$ the set of all oriented rooted trees over i, given a subset G of Ω^2, we denote by $\mathfrak{T}(i)_G$ the set of all oriented rooted trees T over i such that $T\subset G$.

Let $C=\langle i_1, \cdots, i_m\rangle$ be a cycle and let T be a subset of Ω^2. T is called an *oriented pseudo-tree over C* if

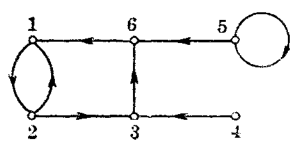
(OPT_1) C is the unique T-cycle over Ω, and

(OPT_2) $T-\{(i_1, i_2)\}, T-\{(i_2, i_3)\}, \cdots, T-\{(i_m, i_1)\}$ are all oriented rooted trees over $i_1, i_2, \cdots, i_m$ respectively.

Given a cycle C, we denote by $T(C)$ the set of all oriented pseudo-trees over C. For a given $G\subset\Omega^2$, we denote by $T(C)_G$ the set of all oriented pseudo-trees T over C such that $T\subset G$.

We note finally that a subset G of Ω^2 is characterized by its directed gragh Γ_G (See Ore [6] for the definition of a directed graph), defined as follows: the vertices of Γ_G are elements of Ω, the directed edges of Γ_G are bijective with elements of G. Thus, for example, if $n=6$ and if the graph of G is given by

(2.1)



then G consists of $(1, 2)$, $(2, 1)$, $(2, 3)$, $(3, 6)$, $(4, 3)$, $(5, 5)$, $(5, 6)$ and $(6, 1)$.

It is convenient to denote a sebset G of Ω^2 by a matrix M_G of degree n defined as follows: if $(i, j) \in G$, then the (i, j)-th entry of this matrix is the symbol $*$, if $(i, j) \notin G$, then the (i, j)-th entry is 0. For example, for G given in (2. 1), we have

$$M_G = \begin{pmatrix} 0 & * & 0 & 0 & 0 & 0 \\ * & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & * \\ * & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

§ 3. Subsets of Ω^2 of Frobenius type.

Let $\Omega = \{1, 2, \cdots, n\}$ and let G be a subset of Ω^2. Let us recall here a binary relation on Ω associated with G which is well-known in the theory of homogeneous Markov chain. Let i and j be elements of Ω. We write iGj if

(i) either $i = j$, or

(ii) $i \neq j$ and there exists a G-path from i to j.

Then we have

(3.1) iGi for every i in Ω,

(3.2) if iGj and jGk, then iGk $(i, j, k \in \Omega)$,

i.e. this binary relation defines a structure of a quasi-partially ordered set on Ω. Let us consider the associated partially ordered set: let $i \in \Omega$ and $j \in \Omega$. We write $i \equiv j$ (mod. G) if we have both iGj and jGi. Then $\equiv$ (mod. G) is an equivalence relation on Ω.

We denote by Ω/G the set of all equivalence classes thus defined. We denote by $[i]_G$ the equivalence class containing i. Now it is easy to see that there exists a unique structure of a partially ordered set on Ω/G such that

$$[i]_G \geqq [j]_G \Leftrightarrow iGj\,.$$

We note that the minimal (resp. non-minimal) elements of Ω/G are called *G-ergodic* (resp. *G-transient*) classes in Ω in the theory of homogeneous Markov chains (see Doob [3]). We denote by $g(G)$ the number of minimal elements of Ω/G.

DEFINITION 3.1. A subset G of Ω^2 is called of Frobenius type if Ω/G consists of a single element.

DEFINITION 3.2. A subset G of Ω^2 is called of quasi-Frobenius type if $g(G)=1$. We note that for any equivalence class $\Sigma=[i]_G$ in Ω/G, $G_1=G\cap\Sigma^2$ is a subset of Σ^2 which is of Frobenius type.

Let $A=(a_{ij})$ be a complex square matrix of degree n. We associate with A a subset Supp(A) of Ω^2, called the support of A, as follows:

$$\mathrm{Supp}(A)=\{(i, j)\in\Omega^2;\ a_{ij}\neq 0\}\ .$$

It is easy to see that for every permutation τ of Ω, one has

$$\mathrm{Supp}(P_\tau A P_\tau^{-1})=\tau(\mathrm{Supp}(A))$$

where P_τ is the permutation matrix associated with τ. $A=(a_{ij})$ is called non-negative if every entry a_{ij} is non-negative. Then Def. 3.1 is justified by the following:

LEMMA 2. *Let A be a non-negative matrix of degree n and let G be the support of A. Then the following conditions are equivalent.*

(i) *G is of Frobenius type.*

(ii) *A is permutation-irreducible,* i.e. *there exists no permutation matrix J such that JAJ^{-1} takes the following form:*

$$\begin{pmatrix} A_1 & 0 \\ A_2 & A_3 \end{pmatrix}$$

PROOF. (i)$\Rightarrow$(ii). Suppose that G is of Frobenius type. Suppose also that A is not permutation-irreducible. Then we may assume A is already of the form mentioned above. Let k be the degree of A_1. Then kGp implies $p\leq k$, as is easily seen by induction on the length of the G-path from k to p. Thus we have $k\not\equiv n$ (mod. G) which is impossible since G is of Frobenius type.

(ii)$\Rightarrow$(i). Suppose that A is permutation-irreducible. Let G be not of Frobenius type, and let Ω_1 be an ergodic class of Ω. Then Ω_1 is a proper subset of Ω. Let $i\in\Omega_1$ and $j\in\Omega$. Then iGj must imply $j\in\Omega_1$ since $[i]_G$ is a minimal element of Ω/G. Hence the (i, j)-th entry of A is 0 if $i\in\Omega_1$ and $j\notin\Omega_1$. But this means the existence of a permutation matrix J which brings JAJ^{-1} into a form mentioned above. q.e.d.

We note that a subset G of Ω^2 of Frobenius type is completely determined by G-cycles: namely, let $C_1, \cdots, C_k$ be the set of all G-cycles and consider the subset $\tilde{G}$ of Ω^2 consisting of $(i, i)\in\Omega^2$ such that $C_p=\langle i\rangle$ for some p and of $(i, j)\in\Omega_0^2$ such that $C_p\in C(i, j)$ for some p. Then $G=\tilde{G}$. In fact, one has obviously $\tilde{G}\subset G$. Now let $(i, j)\in G$. If $i=j$, then $\langle i\rangle$ is a G-cycle and we have $(i, i)\in\tilde{G}$. If $i\neq j$, then there is a G-path $(j, k_1, \cdots, k_r, i)$ from j to i. We have then G-cycle $\langle i, j, k_1, \cdots, k_r\rangle$ i.e. $(i, j)\in\tilde{G}$.

§ 4. A lemma on matrices.

Let x_{ij} $(1\leqq i, j\leqq n; i\neq j)$ be n^2-n indeterminates over the complex number field $\boldsymbol{C}$. We consider the square matrix $X=(x_{ij})$ of degree n where

$$x_{ii}=-(x_{i1}+\cdots+x_{i,i-1}+x_{i,i+1}+\cdots+x_{in})$$

for $i=1, \cdots, n$. Thus $\sum_{j=1}^{n} x_{ij}=0$ for $i=1, \cdots, n$, i.e. denoting by a a column vector of degree n with entries 1 everywhere, one has $Xa=0$. Hence 0 is a root of the characteristic polynomial $\varphi(t)=\det(tI-X)$. We claim that 0 is a simple root of $\varphi(t)$. Suppose $\varphi(t)$ is divisible by t^2. Then, for every $n\times n$ complex matrix $A=(a_{ij})$ satisfying $\sum_{j=1}^{n} a_{ij}=0$ (for $i=1, \cdots, n$), 0 must be a multiple eigen value. But this is impossible as one can see from the following example.

$$A=\begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ -1 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{4.1}$$

In fact, one gets by a simple computation that

$$|tI-A|=t\prod_{\nu=1}^{n-1}(t+\omega^{\nu}-1) \qquad \left(\omega=\exp\frac{2\pi i}{n}\right).$$

Thus the row vectors $b\neq 0$ such that $bX=0$ are uniquely determined up to scalar factors. In particular, denoting by $\varDelta_{ij}$ the (i, j)-th cofactor of X and putting

$$b_i=(\varDelta_{1i}, \varDelta_{2i}, \cdots, \varDelta_{ni}) \text{ for } i=1, \cdots, n\,,$$

we have $b_iX=0$ $(i=1, \cdots, n)$. $\varDelta_{pq}$ is a polynomial in x_{ij} $(i\neq j)$ and $\varDelta_{pq}\neq 0$ as one sees by an example given in (4.1). Hence $b_i\neq 0$. Therefore there exists non-zero rational functions $k_2, \cdots, k_n$ in x_{ij} $(i\neq j)$ such that

$$b_i=k_ib_1 \qquad (i=2, \cdots, n)\,.$$

We claim now $k_2=\cdots=k_n=1$. In fact, for example, since

$$k_2=\frac{\varDelta_{12}}{\varDelta_{11}}=\frac{\varDelta_{22}}{\varDelta_{21}}=\cdots=\frac{\varDelta_{n2}}{\varDelta_{n1}}\,,$$

we see that $k_2=\dfrac{\varDelta_{12}}{\varDelta_{11}}$ does not contain $x_{12}, x_{13}, \cdots, x_{1n}$:

$$\frac{\partial k_2}{\partial x_{12}}=\frac{\partial k_2}{\partial x_{13}}=\cdots=\frac{\partial k_2}{\partial x_{1n}}=0\,.$$

Similarly, $k_2=\frac{\Delta_{i2}}{\Delta_{i1}}$ does not contain $x_{i1},\cdots,x_{i,i-1},x_{i,i+1},\cdots,x_{in}$ for every $i=1,\cdots,n$. Thus k_2 is a constant not depending on $x_{ij}\,(i\neq j)$. So we may compute the value of k_2 by taking a matrix $A=(a_{ij})$ such that $\sum_{j=1}^{n}a_{ij}=0$ for $i=1,\cdots,n$. Taking the matrix (4.1), we get $k_2=1$ easily. We get $k_2=\cdots=k_n=1$ similarly. Thus we have shown that

(4.2) $$\Delta_{i1}=\Delta_{i2}=\cdots=\Delta_{in}\qquad(i=1,\cdots,n)\,.$$

In other words, we have proved the following:

LEMMA 3. *Let $A=(a_{ij})$ be a complex matrix of degree n such that $\sum_{j=1}^{n}a_{ij}=0$ for $i=1,\cdots,n$. Denote by Δ_{ij} the cofactor of a_{ij} in A. Then*

$$\Delta_{i1}=\Delta_{i2}=\cdots=\Delta_{in}\qquad(\mathrm{i}=1,\cdots,n)\,.$$

§ 5. A proof of the theorem of Bott-Mayberry.

Let $\Omega=\{1,\cdots,n\}$ and let $A=(a_{ij})$ be an $n\times n$ complex matrix. Given a non-empty subset T of Ω^2, we put after [1]

(5.1) $$v(A,T)=\prod_{(i,j)\in T}a_{ij}\,.$$

Now, we define $u_i(A)$ $(i\in\Omega)$ by

(5.2) $$u_i(A)=\sum_{T\in\mathfrak{T}(i)}v(A,T)\,.$$

Given a cycle C over Ω, we define $\lambda(A,C)$ by

(5.3) $$\lambda(A,C)=\sum_{T\in T(C)}v(A,T)\,.$$

We also define $\lambda_{ij}(A)$ for $(i,j)\in\Omega^2$ as follows:

(5.4) $$\lambda_{ij}(A)=\sum_{C\in\mathfrak{C}(i,j)}\lambda(A,C)\qquad(\text{for }i\neq j)$$
$$\lambda_{ii}(A)=\lambda(A,\langle i\rangle)\,.$$

LEMMA 4. $\sum_{j=1}^{n}\lambda_{ij}(A)=\sum_{j=1}^{n}\lambda_{ji}(A)$ *for* $i=1,\cdots,n$.

PROOF. We have by Lemma 1,

$$\sum_{j=1}^{n}\lambda_{ij}(A)=\sum_{C\in\mathfrak{C}(i)}\lambda(A,C)$$

and

$$\sum_{j=1}^{n} \lambda_{ji}(A)= \sum_{C\in C(i)} \lambda(A, C) ,$$

which completes the proof.

LEMMA 5. *$\lambda_{ij}(A)=u_i(A)\cdot a_{ij}$ for every (i, j) in Ω^2.*

PROOF. Define a map φ: $T(\langle i\rangle)\rightarrow\mathfrak{T}(i)$ by $\varphi(T)=T-\{(i, i)\}$. Then obviously φ is bijective. Now, since we have $v(A, T)=a_{ii}v(A, \varphi(T))$ for every T in $T(\langle i\rangle)$, we get $\lambda_{ii}(A)=a_{ii}\cdot u_i(A)$.

Now let $(i, j)\in\Omega_0^2$. Define a map ψ: $\bigcup_{C\in C(i,j)} T(C)\rightarrow\mathfrak{T}(i)$ by $\psi(T)=T-\{(i, j)\}$. Then from the definitions of $T(C)$, $C(i, j)$ and $\mathfrak{T}(i)$ given in §2, it is easy to see that ψ is bijective. Moreover, we have $v(A, T)=a_{ij}v(A, \psi(T))$. Hence we get $\lambda_{ij}(A)=u_i(A)\cdot a_{ij}$, q.e.d.

Now putting $a_i=\sum_{j=1}^{n} a_{ij}$ $(i=1, \cdots, n)$, one obtains easily from Lemmas 4 and 5 the following:

LEMMA 6. $u_i(A)\cdot a_i=\sum_{j=1}^{n} u_j(A)\cdot a_{ji}$ $(i=1, \cdots, n)$.

We note that $u_i(A)$ depends only on off-diagonal entries of A. I.e. if $A-B$ is a diagonal matrix, then we have $u_i(A)=u_i(B)$ for $i=1, \cdots, n$.

Now let x_{ij} $(1\leqq i, j\geqq n: i\neq j)$ be n^2-n indeterminates over C. We consider again the square matrix $X=(x_{ij})$ of degree n defined in §4. Then, from $\sum_{j=1}^{n} x_{ij}=0$ $(i=1, \cdots, n)$, one has by Lemma 6

$$0= \sum_{j=1}^{n} u_j(X)\cdot x_{ji} \qquad (i=1, \cdots, n),$$

i.e. the row vector $u=(u_1(X), \cdots, u_n(X))$ satisfies $uX=0$. Hence u is a scalar multiple of the row vector b_1 $(=b_2= \cdots =b_n)$ considered in §4: $u=kb_1$, where k is a rational function in the x_{ij} $(i\neq j)$. We have thus,

$$\frac{u_1(X)}{\Delta_{11}} = \cdots = \frac{u_n(X)}{\Delta_{n1}} =k ,$$

where as in §4, Δ_{ij} is the (i, j)-th cofactor of X. Now from $k=\frac{u_1(X)}{\Delta_{11}}$ and $u_1(X)=\sum_{T\in\mathfrak{T}(1)} v(X, T)$, we see that k does not depend on $x_{12}, x_{13}, \cdots, x_{1n}$. Similarly $k=\frac{u_i(X)}{\Delta_{i1}}$ does not depend on $x_{i1}, \cdots, x_{i,i-1}, x_{i,i+1}, \cdots, x_{in}$. Thus k is a constant. In order to compute the value of k, take the matrix A given in (4.1). Then we get $k=(-1)^{n-1}$. Thus we get (noting $\Delta_{i1}=\Delta_{ii}$ by Lemma 3)

$$u_i(X)= \sum_{T\in\mathfrak{T}(i)} v(X, T)=(-1)^{n-1}\Delta_{i1}=(-1)^{n-1}\Delta_{ii} .$$

In other words, we have proved the following theorem which is equivalent to a theorem of Bott-Mayberry [1].

THEOREM 1. *(Bott-Mayberry). Let* $A=(a_{ij})$ *be an* $n\times n$ *complex matrix such that* $\sum_{j=1}^{n} a_{ij}=0$ *for* $i=1, \cdots, n$. *Then, for each* $i=1, \cdots, n$, $\sum_{T\in\mathfrak{T}(i)} v(A, T)$ *is equal to the* (i, j)*-th cofactor of the matrix* $-A$.

§ 6. Weakly symmetric matrices.

An $n\times n$ complex matrix $A=(a_{ij})$ is called *weakly symmetric* if $\sum_{j=1}^{n} a_{ij}=\sum_{j=1}^{n} a_{ji}$ for $i=1, \cdots, n$.

Now let $\Omega=\{1, 2, \cdots, n\}$ and let C be a cycle over Ω. We denote by $l(C)$ the length of C. We associate with C a weakly symmetric matrix $E_C=(z_{ij})$ as follows:

$$z_{ij}=\begin{cases} 0, & \text{if } i\neq j \text{ and } C\in C(i, j), \\ 1/l(C), & \text{if } i\neq j \text{ and } C\in C(i, j), \\ 0, & \text{if } i=j \text{ and } C\neq\langle i\rangle, \\ 1, & \text{if } i=j \text{ and } C=\langle i\rangle. \end{cases} \tag{6.1}$$

Obviously we have $\sum_{i,j} z_{ij}=1$.

LEMMA 7. *Let* $A=(a_{ij})$ *be a weakly symmetric matrix of degree* n. *Then* $u_1(A)=u_2(A)=\cdots=u_n(A)$.

PROOF. Put $a_i=\sum_{j=1}^{n} a_{ij}=\sum_{j=1}^{n} a_{ji}$ $(i=1, \cdots, n)$ and denote by $D=(d_{ij})$ the diagonal matrix defined by $d_{ij}=0$ (for $i\neq j$), $d_{ii}=a_i$ $(i=1, \cdots, n)$. Now let Δ_{ij} be the (i, j)-th cofactor of $A-D$. Then, applying Lemma 3 on $A-D$ and on ${}^t(A-D)$ (the transpose of $A-D$), one has

$$\Delta_{i1}=\Delta_{i2}=\cdots=\Delta_{in},\ \Delta_{1i}=\Delta_{2i}=\cdots=\Delta_{ni} \qquad (i=1, \cdots, n)\,.$$

Hence all the Δ_{ij} coincide. On the other hand, by Theorem 1 we have

$$u_i(A)=u_i(A-D)=(-1)^{n-1}\Delta_{ii}\,.$$

Therefore we get $u_1(A)=\cdots=u_n(A)$, q.e.d.

THEOREM 2. *Let* A *be a weakly symmetric complex matrix of degree* n. *Then*

$$\tilde{\lambda}(A)\cdot A=\sum_{C} l(C)\cdot\lambda(A, C)\cdot E_C\,, \tag{6.2}$$

where the summation is taken over all cycles C *over* $\Omega=\{1, \cdots, n\}$ *and* $\tilde{\lambda}(A)$ *is defined by*

$$\tilde{\lambda}(A)=u_1(A)=\cdots=u_n(A)\,. \tag{6.3}$$

Furthermore,

$$\left(\sum_{i,j} a_{ij}\right)\cdot\tilde{\lambda}(A)=\sum_{C} l(C)\cdot\lambda(A, C)\,. \tag{6.4}$$

Proof. Denote by $Y=(y_{ij})$ the right hand side of the equation (6.2) in the theorem. Then from the definition of E_C, one gets easily

$$y_{ij}=\lambda_{ij}(A) \qquad \text{for every } (i, j) \text{ in } \Omega^2\,.$$

Putting $u_1(A)=\cdots=u_n(A)=\tilde{\lambda}(A)$ (Lemma 7), we get, by Lemma 5, $Y=\tilde{\lambda}(A)\cdot A$. Hence comparing the sum of all entries of these matrix we have $\sum_{i,j} y_{ij}=\tilde{\lambda}(A)\cdot\sum_{i,j} a_{ij}$. On the other hand, since the sum of all entries of E_C is 1, one has

$$\sum_{i,j} y_{ij}=\sum_{C} l(C)\cdot\lambda(A, C)\,.$$

Thus $\sum_{C} l(C)\lambda(A, C)=\tilde{\lambda}(A)\cdot\sum_{i,j} a_{ij}$ which completes the proof.

§ 7. Stochastic matrices of a given Frobenius type.

A real matrix $P=(p_{ij})$ of degree n is called stochastic if

(i) $p_{ij}\geqq 0$ (for all i and j), and

(ii) $\sum_{j=1}^{n} p_{ij}=1$ (for all i).

We denote by $S(n)$ the set of all stochastic matrices of degree n. $S(n)$ is compact, convex subset of the set of all non-negative matrices of degree n. Furthermore, $S(n)$ is closed with respect to multiplication. Now for any P in $S(n)$, the following limit exists always (cf. Doob [3]):

$$L(P)=\lim_{m\to\infty}\frac{1}{m}(P+P^2+\cdots+P^m)\,.$$

$L(P)$ is also in $S(n)$.

Lemma 8. *Let P be a stochastic matrix of degree n and let G be the support of P. Then $g(G)$ is equal to the multiplicity of the eigen value 1 of P. $g(G)$ is also equal to the rank of the matrix $L(P)$.*

Proof. It is well-known (cf. Gantmacher [4]) that

(i) every eigen value of P has the absolute value not exceeding 1,

(ii) if a is an eigen value of P with $|a|=1$, then a is a root of unity and a is a simple root of the minimal equation of P.

Thus there is a complex non-singular matrix Q such that

$$QPQ^{-1}=\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}$$

where X is a diagonal matrix whose diagonal entries are roots of unity and Y

is a square matrix whose all eigen values are smaller than 1 in the absolute value. Thus $\lim_{m\to\infty} Y^m=0$, hence $L(Y)=0$. We have therefore

$$Q\cdot L(P)\cdot Q^{-1}=\begin{pmatrix} L(X) & 0 \\ 0 & 0 \end{pmatrix}.$$

From this fact, it is immediate to see that the rank of $L(P)$ is equal to the rank of $L(X)$, which is nothing but the multiplicity of the eigen value 1 of the matrix P. Now it is known (cf. Gantmacher, loc. cit.) that this multiplicity is also equal to $g(G)$, q.e.d.

Thus, if P is a stochastic matrix with a quasi-Frobenius support G, $L(P)$ is of rank 1. On the other hand, the column vector

$$a=\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

satisfies the equation $L(P)a=a$. Hence 1 is a simple eigen value of $L(P)$. Now since $L(P)^2=L(P)$ (cf. Doob, loc. cit.), $L(P)x$ is a scalar multiple of a for every column vector x. Thus all the row vectors of $L(P)$ are the same:

$$L(P)=\begin{pmatrix} u_1 & u_2 & \cdots & u_n \\ u_1 & u_2 & \cdots & u_n \\ \cdot & \cdot & \cdots & \cdot \\ u_1 & u_2 & \cdots & u_n \end{pmatrix}$$

Since $L(P)P=L(P)$ (cf. Doob, loc. cit.), the row vector

$$u=(u_1, u_2, \cdots, u_n)$$

is an eigen vector of P: $uP=u$. Furthermore $\sum_{i=1}^{n} u_i=1$. We note that u is characterized by these two properties since 1 is a simple eigen value of P. We denote this vector u by $u=L^*(P)$.

Now let us formulate the problem we shall consider below. Let G be a given subset of Ω^2 of quasi-Frobenius type. We denote by $S(n)_G$ the set of all stochastic matrices P in $S(n)$ such that G is the support of P. We want to determine the set consisting of the $L^*(P)$ when P ranges over the set $S(n)_G$. We denote this set by Δ_G.

Let us generalize this problem slightly. Let G be any subset of Ω^2 such that $S(n)_G$ is not empty. For every P in $S(n)_G$ we denote again by $L^*(P)$ the set of all row vectors $(u_1, \cdots, u_n)=u$ such that

(i) $u_1 \geqq 0, \cdots, u_n \geqq 0$,

(ii) $u_1 + \cdots + u_n = 1$,

(iii) $uP = u$.

It is well known that $L^*(P) \neq \phi$. We denote again by Δ_G the union of all $L^*(P)$ for $P \in S(n)_G$. Clearly these are natural extensions of the definition above for the case where G is of quasi-Frobenius type.

The question of determination of the set Δ_G is now reduced to the case where G is of Frobenius type as follows.

Let G be a subset such that $S(n)_G \neq \phi$. Let

$$\Omega = \Omega_1 \cup \cdots \cup \Omega_g \cup \cdots \cup \Omega_r$$

be the partition of Ω into the equivalence classes under the binary relation $\equiv$ (mod. G) (see § 2). Thus, $\Omega_1, \cdots, \Omega_r$ are the elements of Ω/G. Let $\Omega_1, \cdots, \Omega_g$ ($g = g(G)$) be the totality of the minimal elements of Ω/G. Applying a suitable permutation on Ω if necessary, we may assume that

$$\text{if } i \in \Omega_p,\ j \in \Omega_q,\ p < q, \text{ then } i < j. \tag{7.1}$$

Then every $P=(p_{ij})$ in $S(n)_G$ has the following so-called normal form:

$$P = \begin{pmatrix} P_1 & 0 & 0 & \cdots 0 & 0 & 0 & \cdots 0 \\ 0 & P_2 & 0 & \cdots 0 & 0 & 0 & \cdots 0 \\ \cdot & \cdot & \cdot & \cdots \cdot & \cdot & \cdot & \cdots \cdot \\ 0 & 0 & 0 & \cdots P_g & 0 & 0 & \cdots 0 \\ P_{g+1,1} & P_{g+1,2} & P_{g+1,3} & \cdots P_{g+1,g} & P_{g+1} & 0 & \cdots 0 \\ \cdot & \cdot & \cdot & \cdots \cdot & \cdot & \cdot & \cdots \cdot \\ P_{r1} & P_{r2} & P_{r3} & \cdots P_{r,g} & P_{r,g+1} & P_{r,g+2} & \cdots P_r \end{pmatrix}$$

Now it is known (see Gantmacher [4, Chap. 2]) that none of the matrices $P_{g+1}, P_{g+2}, \cdots, P_r$ has any eigen value λ such that $|\lambda| \geqq 1$. Let u be in $L^*(P)$ and let $u=(u_1, \cdots, u_g, \cdots, u_r)$ be the block-decomposition of entries of u according to the partition $\Omega = \Omega_1 \cup \cdots \cup \Omega_g \cup \cdots \cup \Omega_r$. Then $uP=u$ implies $u_r P_r = u_r$. Since 1 is not an eigen value of P_r, we have $u_r = 0$. Then from $u_{r-1}P_{r-1} + u_r P_{r,r-1} = u_{r-1}$ we get $u_{r-1}P_{r-1} = u_{r-1}$, and so $u_{r-1} = 0$. Thus $u_r = u_{r-1} = \cdots = u_{g+1} = 0$. Hence $uP=u$ implies $u_1 P_1 = u_1, \cdots, u_g P_g = u_g$.

Obviously, $P_1, \cdots, P_g$ are stochastic matrices whose supports are all of Frobenius type. Hence $L^*(P_i)$ ($i=1, \cdots, g$) consists of a single vector e_i and each u_i ($i=1, \cdots, g$) is a scalar multiple of e_i:

$$u_i = k_i e_i \qquad (i=1, \cdots, g).$$

Obviously $k_1 \geqq 0, \cdots, k_g \geqq 0$ and $k_1 + \cdots + k_g = 1$. Now put $\Omega_i = \{j+1, j+2, \cdots, j+n_i\}$ and $e_i = (\xi_{i1}, \cdots, \xi_{in_i})$. We denote by $\bar{e}_i$ the row vector of degree n defined by $\bar{e}_i = (\eta_1, \cdots, \eta_n)$ where $\eta_k = 0$ for $k \notin \Omega_i$, $\eta_{j+1} = \xi_{i1}, \eta_{j+2} = \xi_{i2}, \cdots, \eta_{j+n_i} = \xi_{in_i}$. Then we have $u = k_1\bar{e}_1 + \cdots + k_g\bar{e}_g$. Thus u is in the convex hull $\bar{\Delta}$ of the set $\bar{\Delta}_1 \cup \cdots \cup \bar{\Delta}_g$, where we put $\bar{\Delta}_i = 0 \times \cdots \times 0 \times \Delta_{G_i} \times 0 \times \cdots \times 0$ and $G_i = G \cap \Omega_i^2$ $(i=1, \cdots, g)$. Conversely if u is in $\bar{\Delta}$, then u is of the form $u = k_1 z_1 + \cdots + k_g z_g$ with $z_i \in \Delta_{G_i}$, $k_i \geqq 0$ $(i=1, \cdots, g)$ and $\sum_{i=1}^{g} k_i = 1$. Hence there exists $P_i \in S(n)_{G_i}$ $(i=1, \cdots, g)$ such that $z_i = L^*(P_i)$. Then it is easily seen that there exists P in $S(n)_G$ such that $u \in L^*(P)$. Thus we get

(7.2) Δ_G coincides with the convex hull $\bar{\Delta}$ of the set $\bar{\Delta}_1 \cup \cdots \cup \bar{\Delta}_g$

where $\bar{\Delta}_i = 0 \times \cdots \times 0 \times \Delta_{G_i} \times 0 \times \cdots \times 0$ and $G_i = G \cap \Omega_i^2$ $(i=1, \cdots, g)$.

Thus we have reduced our problem to the case where G is a subset of Ω^2 of Frobenius type, which we shall assume from now on.

LEMMA 9. *Let G be a subset of Ω^2 of Frobenius type, and let $P \in S(n)_G$, $u = (u_1, \cdots, u_n) = L^*(P)$. Then $u_1 > 0, \cdots, u_n > 0$.*

PROOF. Since G is of Frobenius type, it is easy to see that $\mathfrak{T}(i)_G \neq \phi$. Hence $u_i(P) = \sum_{T \in \mathfrak{T}(i)} v(P, T) = \sum_{T \in \mathfrak{T}(i)_G} v(P, T) > 0$. By Lemma 6, the vector $(u_1(P), \cdots, u_n(P)) = u(P)$ satisfies $u(P)P = u(P)$. Therefore $u(P)$ is a scalar multiple of u : $u(P) = ku$. Then we get $k = u_1(P) + \cdots + u_n(P) > 0$ and one has $u_i > 0$ $(i=1, \cdots, n)$, q.e.d.

§ 8. Space $T(n)_G$.

This section is due to T. Kato.

Denote by $T(n)$ the set of all non-negative, weakly symmetric matrices $T = (t_{ij})$ of degree n such that $\sum_{i,j} t_{ij} = 1$. Given a subset G of Ω^2 of Frobenius type we denote by $T(n)_G$ the subset of $T(n)$ consisting of all T in $T(n)$ such that $\mathrm{Supp}(T) = G$. Then for every $T = (t_{ij})$ in $T(n)_G$, we have

$$t_i = \sum_{j=1}^{n} t_{ij} = \sum_{j=1}^{n} t_{ji} > 0 \text{ for } i=1, \cdots, n.$$

Now we define two mappings $\sigma : T(n)_G \to S(n)_G$ and $\tau : T(n)_G \to \Delta_G$ as follows: let $T = (t_{ij})$ be in $T(n)_G$ and $t_i = \sum_{j=1}^{n} t_{ij} = \sum_{j=1}^{n} t_{ji}$, then

(8.1) $$\tau(T) = (t_1, \cdots, t_n),$$

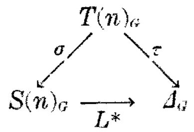
(8.2) $$\sigma(T) = (p_{ij}), \ p_{ij} = \frac{t_{ij}}{t_i} \qquad (\text{for } (i, j) \in \Omega^2).$$

Clearly $P=(p_{ij})=\sigma(T)$ is then in $S(n)_G$. Moreover $\sum_{i=1}^{n} t_i p_{ij}=\sum_{i=1}^{n} t_{ij}=\sum_{i=1}^{n} t_{ji}=t_j$. Since $t_1>0, \cdots, t_n>0$ and $\sum_{i=1}^{n} t_i=1$, one has $\tau(T)=L^*(P)\in \Delta_G$. Thus σ (resp. τ) maps $T(n)_G$ into $S(n)_G$ (resp. into Δ_G).

THEOREM 3. (*T. Kato*) *Let G be a subset of Ω^2 of Frobenius type. Then*

(i) *$\sigma: T(n)_G \to S(n)_G$ and $\tau: T(n)_G \to \Delta_G$ are both surjective.*

(ii) *the following diagram is commutative:*



i.e. $L^*\circ\sigma=\tau$.

(iii) *Δ_G is convex.*

PROOF. (ii) was shown above already. Let $P=(p_{ij})$ be in $S(n)_G$. Put $u=(u_1, \cdots, u_n)=L^*(P)$. Then $u_1>0, \cdots, u_n>0$ by Lemma 9. Now set $t_{ij}=u_i p_{ij}$ for $(i, j)\in \Omega^2$. Then, since $\sum_{i=1}^{n} p_{ji}=1$, one has

$$\sum_{i=1}^{n} t_{ij}=\sum_{i=1}^{n} u_i p_{ij}=u_j=\sum_{i=1}^{n} u_j p_{ji}=\sum_{i=1}^{n} t_{ji}\,,$$

i.e. $(t_{ij})=T$ is weakly symmetric. Moreover $\mathrm{Supp}(T)=G$, and $\sum_{i,j} t_{ij}=\sum_i u_i=1$. Hence $T\in T(n)_G$. Clearly we get $\sigma(T)=P$. Thus σ is surjective. Then by (ii), τ is also surjective. (iii) is then an immediate consequence of the surjectivity and the fact that τ is a linear mapping, and that $T(n)_G$ is convex. q.e.d.

COROLLARY. *$\sigma: T(n)_G \to S(n)_G$ is bijective.*

§9. Extreme points of Δ_G.

Let G be a subset of Ω^2 of Frobenius type. We denote by $\langle G\rangle$ the set of all G-cycles. We also denote by $\langle G\rangle_{\boldsymbol{R}}$ the real vector space spanned by $\langle G\rangle$ freely, i.e. $\langle G\rangle_{\boldsymbol{R}}$ is the real vector space consisting of all formal linear combinations $\sum_{C\in\langle G\rangle} \mu_C\cdot C$ with coefficients μ_C in the real number field $\boldsymbol{R}$, and $\langle G\rangle$ is linearly independent over $\boldsymbol{R}$.

Now we define Θ_G to be the subset of $\langle G\rangle_{\boldsymbol{R}}$ as follows:

$$\Theta_G=\left\{\sum_{C\in\langle G\rangle} \mu_C\cdot C\,;\, \mu_C>0 \text{ (for all } C\in\langle G\rangle) \text{ and } \sum_C \mu_C=1\right\}.$$

Thus, if $\langle G\rangle$ consists of k elements, Θ_G is an $(k-1)$-dimensional open simplex.

Next we define a linear mapping $\tilde{\rho}$ from $\langle G\rangle_{\boldsymbol{R}}$ into the space $M(n, \boldsymbol{R})$ of all $n\times n$ real matrices as follows:

(9.1) $$\tilde{\rho}\Big(\sum_{C} \mu_C \cdot C\Big) = \sum_{C} \mu_C \cdot E_C$$

(see §6 for the definition of the matrix E_C). We define moreover a linear mapping $\tilde{\tau}$: $M(n, \boldsymbol{R}) \to \boldsymbol{R}^n$ as follows:

(9.2) $$\tilde{\tau}(X) = (x_1, \cdots, x_n)$$

where $X = (x_{ij})$, $x_i = \sum_{j=1}^{n} x_{ji}$ $(i = 1, \cdots, n)$.

We denote $\tilde{\tau}(E_C)$ by v_C. The restriction of $\tilde{\tau}$ on $T(n)_G$ obviously coincides with the map τ: $T(n)_G \to \Delta_G$ given in §8. We note that $v_C = (\xi_1^C, \cdots, \xi_n^C)$ is given as follows:

(9.3) $$\xi_i^C = \begin{cases} 1/l(C), & \text{if } C \in C(i), \\ 0, & \text{otherwise.} \end{cases}$$

We claim that, if $\theta = \sum_{C} \mu_C \cdot C$ is in Θ_G, then $\tilde{\rho}(\theta) \in T(n)_G$. In fact, put $\tilde{\rho}(\theta) = \sum \mu_C \cdot E_C = T = (t_{ij})$; then T is weakly symmetric and $\sum_{i,j} t_{ij} = 1$ because this is the case for E_C and $\sum \mu_C = 1$. Now the support of T is G, since one can check easily the following equalities:

$$t_{ii} = \begin{cases} \mu_{\langle i \rangle}, & \text{if } (i, i) \in G \\ 0, & \text{otherwise,} \end{cases}$$

$$t_{ij} = \sum_{C \in C(i,j)_G} \frac{1}{l(C)} \mu_C, \qquad (i \neq j).$$

Thus, restricting $\tilde{\rho}$ on Θ_G, we define a map ρ: $\Theta_G \to T(n)_G$.

Next we define a map ρ^*: $T(n)_G \to \Theta_G$ as follows: let $T \in T(n)_G$ and $C \in \langle G \rangle$. Define $\lambda^*(T, C)$ by

(9.4) $$\lambda^*(T, C) = \frac{l(C) \cdot \lambda(T, C)}{\lambda(T)}, \quad \text{where } \lambda(T) = \sum_{C \in \langle G \rangle} l(C) \cdot \lambda(T, C).$$

Then clearly

(9.5) $$\lambda^*(T, C) > 0 \text{ for } T \in T(n)_G, C \in \langle G \rangle$$

(9.6) $$\sum_{C \in \langle G \rangle} \lambda^*(T, C) = 1.$$

ρ^*: $T(n)_G \to \Theta_G$ is defined by

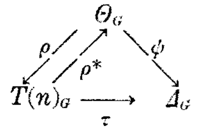
(9.7) $$\rho^*(T) = \sum_{C \in \langle G \rangle} \lambda^*(T, C) \cdot C.$$

We define ψ: $\Theta_G \to \Delta_G$ by

(9.8) $$\psi = \tau \circ \rho.$$

Thus we get the following diagram:

(9.9)



Lemma 10. *Let G be a subset of Ω^2 of Frobenius type and $T \in T(n)_G$. Then*

(i) $T = \sum_{C \in \langle G \rangle} \lambda^*(T, C) \cdot E_C$,

(ii) $\lambda(T) = u_1(T) = u_2(T) = \cdots = u_n(T)$.

Proof. Immediate from Theorem 2.

Lemma 11. (i) $\rho \circ \rho^* = \mathrm{id.}$,

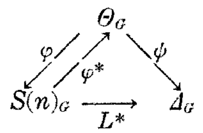
(ii) ρ *is surjective.*

(iii) $\tau = \psi \circ \rho^*$.

Proof. (i) is equivalent to (i) of Lemma 10. (ii) is consequence of (i). Now as to (iii), $\psi \circ \rho^* = \tau \circ \rho \circ \rho^* = \tau$, q.e.d.

Let us now translate these properties of the diagram (9.9) into those of the following diagram

(9.10)



making use of the bijection $\sigma: T(n)_G \to S(n)_G$. Here $\varphi: \Theta_G \to S(n)_G$ is defined by

(9.11)
$$\varphi = \sigma \circ \rho$$

and $\varphi^*: S(n)_G \to \Theta_G$ is defined by

(9.12)
$$\varphi^* = \rho^* \circ \sigma^{-1} .$$

Then Lemma 11 can be translated as follows:

Lemma 12. (i) $\varphi \circ \varphi^* = \mathrm{id.}$,

(ii) φ *is surjective,*

(iii) $L^* = \psi \circ \varphi^*$

Remark. In general, we have $\rho^* \circ \rho \neq \mathrm{id.}$, $\varphi^* \circ \varphi \neq \mathrm{id.}$

Let us interpret (iii) of Lemma 12 as follows:

Theorem 4. *Let G be a subset of Ω^2 of Frobenius type. Let P be a stochastic matrix with support G. Then the fixed row vector $L^*(P)$ of P is given by the following formula:*

$$L^*(P) = \sum_{C \in \langle G \rangle} \lambda^*(P, C) \cdot v_C$$

where C ranges over the set $\langle G\rangle$ of all G-cycles; v_C is a row vector defined by (9.3); $\lambda^*(P, C)$ *is defined as follows:*

$$\lambda^*(P, C)=\frac{l(C)\cdot\lambda(P, C)}{\lambda(P)}, \text{ where } \lambda(P)=\sum_{C\in\langle G\rangle} l(C)\cdot\lambda(P, C),$$

PROOF. $L^*(P)=(\psi\circ\varphi^*)(P)=(\psi\circ\rho^*\circ\sigma^{-1})(P)$. Put $T=\sigma^{-1}(P)$. Then

$$L^*(P)=\psi(\rho^*(T))=\psi\Big(\sum_{C\in\langle G\rangle}\lambda^*(T, C)C\Big)=\sum_C \lambda^*(T, C)\psi(C).$$

and $\psi(C)=\tau(\rho(C))=\tau(E_C)=v_C$. Hence we get

$$L^*(P)=\sum_C \lambda^*(T, C)\cdot v_C.$$

Now let us show $\lambda^*(T, C)=\lambda^*(P, C)$ for all $C\in\langle G\rangle$. In fact, putting $P=(p_{ij})$, $\sigma^{-1}(P)=T=(t_{ij})$ and $t_i=\sum_j t_{ij}=\sum_j t_{ji}$, we have $t_i p_{ij}=t_{ij}$. Then it is easy to see that $t_1\cdots t_n\cdot\lambda(P,C)=\lambda(T,C)$. Hence $t_1\cdots t_n\cdot\lambda(P)=\lambda(T)$ and we get $\lambda^*(P,C)=\lambda^*(T,C)$, q.e.d.

THEOREM 5. *Let G be a subset of Ω^2 of Frobenius type. Then the set Δ_G of all row vectors fixed by some stochastic matrix in $S(n)_G$ consists of all row vectors of the form $\sum\mu_C\cdot v_C$, $\mu_C>0$ (for every $C\in\langle G\rangle$), $\sum_{C\in\langle G\rangle}\mu_C=1$, where v_C is the row vector associated with a G-cycle C by* (9.3).

COROLLARY. *Let L_G be the convex hull of the set $E_G^*=\{v_C; C\in\langle G\rangle\}$. Then*
(i) *L_G coincides with the closure $\bar{\Delta}_G$ of Δ_G in R^n.*
(ii) *$E_G^*=\{v_C; C\in\langle G\rangle\}$ contains the set E_G of all extreme points of $\bar{\Delta}_G$.*
(iii) *Let $A_G=\{\sum_{C\in\langle G\rangle}\xi_C\cdot v_C; \xi_C\in R$ (for every $C\in\langle G\rangle$), $\sum_C\xi_C=1\}$ be the smallest affine subspace of R^n containing $E_G^*=\{v_C; C\in\langle G\rangle\}$. Then Δ_G is the interior of the convex hull L_G of $E_G^*=\{v_C; C\in\langle G\rangle\}$ in the space A_G.*

PROOF. The theorem is a re-statement of the fact that $\psi:\Theta_G\to\Delta_G$ $\psi(\sum\mu_C\cdot C)=\sum\mu_C\cdot v_C$ is surjective. However this is obvious since $\psi=\tau\circ\rho$ and $\rho:\Theta_G\to T(n)_G$ and $\tau: T(n)_G\to\Delta_G$ are both surjective. The corollary is an immediate consequence of the theorem and the properties of extreme points (cf. Bourbaki [2]), q.e.d.

Thus, for a subset G of Ω^2 of Frobenius type, Δ_G is an open cell of the affine space A_G spanned by Δ_G. Now by the reduction of the general case in §7, we have the following:

THEOREM 6. *Let G be a subset of Ω^2 such that $S(n)_G\neq\phi$. Let $\Omega_1,\cdots,\Omega_g$ ($g=g(G)$) be the minimal elements of Ω/G. Put $G_i=G\cap\Omega_i^2$ ($i=1,\cdots,g$). Then Δ_G consists of all row vectors of the from*

$$\sum_{i=1}^{g}\sum_{C\in\langle G_i\rangle}\xi_C^i v_C,$$

where ξ_C^i *are real numbers such that*

$$\xi_C^i \geqq 0 \qquad (\textit{for } i=1, \cdots, g, C\in\langle G_i\rangle),$$

$$\sum_{i=1}^{g} \sum_{C\in\langle G_i\rangle} \xi_C^i = 1,$$

and that, for a fixed i, ξ_C^i $(C\in\langle G_i\rangle)$ *are all zero or all positive.* v_C *is a row vector defined by* (9.3). *In particular,* Δ_G *is convex and* $\bar{\Delta}_G$ *has only finitely many extreme points.*

REMARK 1. For a subset G of Ω^2 of Frobenius type, a statement for $T(n)_G$ analogous to Theorem 5 is valid by Lemma 10 and by the surjectivity of ρ. For example, $\tilde{E}_G^* = \{E_C; C\in\langle G\rangle\})$ contains the set of all extreme points $\tilde{E}_G$ of the closure $\overline{T(n)_G}$ of $T(n)_G$ in $M(n, \boldsymbol{R})$, and $\tilde{E}_G \subset \tilde{E}_G^* \subset \overline{T(n)_G}$.

REMARK 2. In the corollary to Theorem 5, $E_G^* \neq E_G$ in general. For example, if $G=\Omega^2$, and $C=\langle 1, 2, \cdots, n\rangle$, $v_C\in E_G^* - E_G$.

§10. Some applications.

LEMMA 13. *Let* G, G' *be subsets of* Ω^2 *such that* $G\subset G'$. *Then* (i) *If* G *is of quasi-Frobenius type, so is* G'. (ii) *If* G *is of Frobenius type, so is* G'.

PROOF. Let $i, j\in\Omega$. Then iGj implies $iG'j$. Hence $i\equiv j$ (mod. G) implies $i\equiv j$ (mod. G'). Thus there is induced a map π: $\Omega/G \to \Omega/G'$. π is obviously order-preserving and surjective. Thus our assertion (ii) is obvious. Assume now G is of quasi-Frobenius type. Let p' and q' be minimal elements of Ω/G'. Then there are elements p, q in Ω/G such that $\pi(p)=p', \pi(q)=q'$. Since Ω/G has the unique minimal element r, we have $p\geqq r, q\geqq r$. Then $p'\geqq\pi(r), q'\geqq\pi(r)$. Hence by the minimality of p' and q', we have $p'=\pi(r)=q'$. Thus $g(G')=1$, i.e. G' is of quasi-Frobenius type, q.e.d.

LEMMA 14. *Let* $P, P_1, P_2, \cdots$ *be elements of* $S(n)$ *such that* $\lim_{m\to\infty} P_m = P$. *If the support of* P *is of quasi-Frobenius type, then the supports of* P_m *is also of quasi-Frobenius type for sufficiently large* m *and we have*

$$\lim_{m\to\infty} L^*(P_m) = L^*(P)$$

PROOF. Let $G, G_1, G_2, \cdots$ be the support of $P, P_1, P_2, \cdots$ respectively. Then we have obviously $G_m \supset G$ for sufficiently large m. Hence G_m is a quasi-Frobenius type by Lemma 13. Now put $L^*(P_m)=u_m$, $L^*(P)=u$ $(m=1, 2, \cdots)$. Since u_m is contained in a compact set, in order to show $\lim_{m\to\infty} u_m = u$, it is enough to show $v=u$ for every convergent subsequence $u_{m_1}, u_{m_2}, \cdots \to v$. Thus we may assume that $\lim_{m\to\infty} u_m = v$ exists. Then $u_m P_m = u_m$ $(m=1, 2, \cdots)$ implies $vP=v$. Hence v is a scalar multiple of u. But since the sum of the components of v is clearly 1, we have $v=u$, q.e.d.

REMARK. The map $P \to L(P) = \lim_{m\to\infty} \frac{1}{m}(P+ \cdots +P^m)$ from $S(n)$ into itself is not continuous if $n \geqq 2$. For example let

$$P(t)=\begin{pmatrix} 1-t & t \\ 2t & 1-2t \end{pmatrix}, \qquad 0 \leqq t \leqq \frac{1}{2},$$

$$Q=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then $\lim_{t\to 0} P(t) = Q$. But

$$L(P(t))=\begin{pmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \end{pmatrix} \text{ for } 0<t<\frac{1}{2}, \quad L(Q)=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

LEMMA 15. *Let T be a pseudo-tree over a Ω^2-cycle C. Then $S(n)_T$ consists of a single matrix Q. Moreover, $L^*(Q)=v_C$.*

PROOF. Since any rooted tree has cardinality $n-1$, any pseudo-tree has cardinality n. Thus every row of T (in the matrix form) contains exactly one $*$. I.e. for any $i \in \Omega$, there is exactly one $j \in \Omega$ such that $(i, j) \in T$. Hence every matrix Q in $S(n)_T$ is of the following form: every row of Q has 1 as its entry at exactly one place and all other entries are 0. Thus $Q \in S(n)_T$ is uniquely determined.

Now by a suitable permutation of Ω, we may assume that $C=\langle 1, 2, \cdots, r\rangle$. Then Q has the following form:

$$Q=\begin{pmatrix} Q_1 & 0 \\ Q_2 & Q_3 \end{pmatrix}$$

where Q_1 is a stochastic matrix of degree $r=l(C)$, and all the eigen values of Q_3 are smaller than 1 in the absolute value. Moreover, Q_1 is a cyclic permutation matrix of r letters. Therefore we get

$$L(Q)=\begin{pmatrix} L(Q) & 0 \\ * & 0 \end{pmatrix}$$

i.e. $L^*(Q)=(L^*(Q_1), 0, \cdots, 0)=(1/r, 1/r, \cdots, 1/r, 0, \cdots, 0)=v_C$, q.e.d.

THEOREM 7. *Let G be a subset of Frobenius type in Ω^2. Define n functions $\tilde{u}_i$ on $S(n)_G$ by $L^*(P)=\tilde{u}=(\tilde{u}_1(p), \cdots, \tilde{u}_n(p))$, $P \in S(n)_G$. Then*

(i) *$\tilde{u}_i(P) \geqq \tilde{u}_j(P)$ for all $P \in S(n)_G$ if and only if $C(i)_G \supset C(j)_G$*

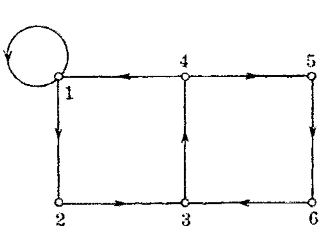
(ii) *if $C(i)_G \supsetneqq C(j)_G$, then $\tilde{u}_i(P) > \tilde{u}_j(P)$ for every $P \in S(n)_G$*

(iii) *$\tilde{u}_i(P) = \tilde{u}_j(P)$ for all $P \in S(n)_G$ if and only if $C(i)_G = C(j)_G$.*

PROOF. Suppose $\tilde{u}_i(P)\geqq\tilde{u}_j(P)$ for all $P\in S(n)_G$. Let C be any cycle in $C(j)_G$. Since G is of Frobenius type, there exists a pseudo-tree T in $T(C)_G$. Clearly T is of quasi-Frobenius type. Since $T\subset G$, $S(n)_T$ is contained in the closure of $\overline{S(n)_G}$. Hence, by Lemma 14, we have $\tilde{u}_i(Q)\geqq\tilde{u}_j(Q)$ for every $Q\in S(n)_T$. Now, by Lemma 15, $C\in C(j)_G$ implies $\tilde{u}_j(Q)>0$. Therefore we have $\tilde{u}_i(Q)>0$. But this gives us $C\in C(i)_G$ by Lemma 15. Thus $C(j)_G\subset C(i)_G$.

Now if $C(j)_G=C(i)_G$, then we have $\tilde{u}_i(P)=\tilde{u}_j(P)$ on $S(n)_G$ by Theorem 4. Also if $C(i)_G\supsetneqq C(j)_G$, then we have $\tilde{u}_i(P)>\tilde{u}_j(P)$ by Theorem 4. (Note that $\lambda^*(P,C)>0$ for every G-cycle C). These considerations complete the proof of Theorem 7.

EXAMPLE. Let $G\subset\Omega^2\,(n=6)$ be given as frllows:



i.e. $$G=\begin{pmatrix} * & * & 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & 0 \end{pmatrix}$$

Then $\langle G\rangle$ consists of 3 G-cycles:

$$C_1=\langle 1\rangle,\quad C_2=\langle 1,2,3,4\rangle,\quad C_3=\langle 3,4,5,6\rangle.$$

Then

$$C(1)_G=\{C_1, C_2\},\quad C(2)_G=\{C_2\},\quad C(3)_G=\{C_2, C_3\},$$
$$C(4)_G=\{C_2, C_3\},\quad C(5)_G=\{C_3\},\quad C(6)_G=\{C_3\}.$$

Thus we have by Theorem 7, that

$$\tilde{u}_3(P)=\tilde{u}_4(P),\quad \tilde{u}_5(P)=\tilde{u}_6(P)\quad \text{for all } P\in S(6)_G.$$

Also we have

$$\tilde{u}_1(P)>\tilde{u}_2(P),\quad \tilde{u}_3(P)>\tilde{u}_2(P)\quad \text{for all } P\in S(6)_G.$$

§ 11. A criterion for the injectivity of the map L*.

Let G be a subset of Ω^2 of Frobenius type. We denote by $|G|$ the number of elements in the set G. We denote by $c(G)$ the number of G-cycles and by $d(G)$ the dimension of the affine subspace A_G of R^n spanned by Δ_G. Our purpose now is to derive several relations between these numbers, $|G|, c(G), d(G)$.

To begin with, put $\langle G\rangle=\{C_1,\cdots,C_k\}$ where $k=c(G)$. Put

(11.1) $$v_p=v_{C_p}=(\xi_{p1},\cdots,\xi_{pn})\,(p=1,\cdots,k).$$

Then by definition we have

$$d(G)+1=\text{rank of}\begin{pmatrix} \xi_{11} & \cdots & \xi_{1n} & 1 \\ \xi_{21} & \cdots & \xi_{2n} & 1 \\ \cdot & \cdots & \cdot & \cdot \\ \xi_{k1} & \cdots & \xi_{kn} & 1 \end{pmatrix}. \tag{11.2}$$

Since $\sum_i \xi_{pi}=1$, the last column of the matrix in (11.2) is a linear conbinations of the other columns. Therefore

$$d(G)+1=\text{rank of}\begin{pmatrix} \xi_{11} & \cdots & \xi_{1n} \\ \xi^{21} & \cdots & \xi_{2n} \\ \cdot & \cdots & \cdot \\ \xi_{k1} & \cdots & \xi_{kn} \end{pmatrix}. \tag{11.3}$$

Thus we got

LEMMA 16. *$d(G)+1$ is the dimension of the linear subspace spanned by $l(C)\cdot v_C$ $(C\in\langle G\rangle)$.*

We note that $l(C)\cdot v_C=(\varepsilon_1, \cdots, \varepsilon_n)$ is given as follows:

$$\varepsilon_i=\begin{cases}1, & \text{if } C\in C(i), \\ 0, & \text{otherwise.}\end{cases} \tag{11.4}$$

THEOREM 8. *Let G be a subset of Ω^2 of Frobenius type. Then*

(i) $c(G)-1\geqq|G|-n\geqq d(G)$,

(ii) *L^*: $S(n)_G\to\Delta_G$ is injective if and only if $|G|=n+d(G)$.*

PROOF. (i) Obviously we have $\dim\Theta_G=c(G)-1$ and $\dim S(n)_G=|G|-n$. Also one has $\dim\Delta_G=d(G)$ by definition. Since σ: $T(n)_G\to S(n)_G$ is a homeomorphism, we have $\dim T(n)_G=\dim S(n)_G=|G|-n$.

Since ρ: $\Theta_G\to T(n)_G$, τ: $T(n)_G\to\Delta_G$, are surjective affine-linear maps, we get

$$\dim\Theta_G\geqq\dim T(n)_G\geqq\dim\Delta_G,$$

i.e. (i) is proved. (ii) L^*: $S(n)_G\to\Delta_G$ is injective if and only if τ: $T(n)_G\to\Delta_G$ is injective, because of the bijection σ: $T(n)_G\to S(n)_G$ and $L^*\circ\sigma=\tau$. However, since τ is an affine-linear map, τ is injective if and only if $\dim T(n)_G=\dim\Delta_G$, i.e. (ii) is proved.

COROLLARY. *Let G be a subset of Ω^2 of Frobenius type such that*

$$L^*: S(n)_G\to\Delta_G$$

is injective. Then $|G|\leqq 2n-1$.

PROOF. $|G|=n+d(G)\leqq n+(n-1)$ since $d(G)\leqq n-1$, q.e.d.

Finally, as to the actual computation for $d(G)$, we note the following

LEMMA 17. *Let G be a subset of Ω^2 of Frobenius type and let $f(G)$ be the dimension of the real vector space F_G consisting of all $n\times n$ real weakly symmetric matrices $Q=(q_{ij})$ such that*

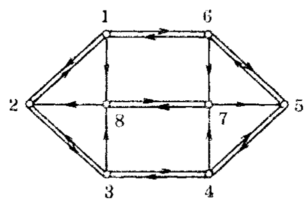
$$\sum_j q_{ij}=\sum_j q_{ji}=0 \qquad (i=1, \cdots, n)$$
$$q_{ij}=0 \quad (\text{for } (i, j)\in G).$$

Then $d(G)=|G|-n-f(G)$.

PROOF. Let $v=(v_1, \cdots, v_n)$ be any point in Δ_G and let Y_v be the full inverse image of v under the map $\tau: T(n)_G\to\Delta_G$. Since τ is a surjective, affine-linear map, Y_v is a convex subset of dimension $\dim T(n)_G-\dim\Delta_G=|G|-n-d(G)$. Fix an element $T_0=(t_{ij}^0)$ in Y_v. Then for any $T=(t_{ij})\in Y_v$, one has $T-T_0\in F_G$, i.e. Y_v is contained in the set T_0+F_G. Hence $\dim Y_v\leqq\dim F_G$. Now let U be a sufficiently small neighborhood of 0 in F_G. Then $T_0+U\subset Y_v$. Thus $\dim U\leqq\dim Y_v$, i.e. $\dim F_G\leqq\dim Y_v$. Therefore we have

$$f(G)=\dim Y_v=\dim T(n)_G-\dim\Delta_G=|G|-n-d(G)\,, \qquad \text{q.e.d.}$$

EXAMPLE. Let $n=8$ and G be given by the following graph;



Then, $f(G)$ is easily seen to be 5. Hence one gets $d(G)=|G|-n-f(G)=20-8-5=7$. But here $c(G)=23$ and the direct computation for $d(G)$ is rather awkward. ($d(G)$ is the rank of a 23×8 matrix).

Universivy of California, Berkeley

References

[1] R. Bott and J. P. Mayberry, Matrices and trees, Economic activity analysis, edited by O. Morgenstern, Wiley, New York, 1954, pp. 391-400 see Math. Review, **16**, p. 665.
[2] N. Bourbaki, Eléments de mathématique, Livre V, Espaces vectoriels topologiques, Paris, 1953.
[3] J. L. Doob, Stochastic Processes, Wiley, New York, 1953.
[4] F. P. Gantmacher, Applications of the theory of matrices, Interscience, New York, 1959.
[5] J. G. Kemeny and J. L. Snell, Mathematical models in the social sciences, Ginn, New York, 1962.
[6] O. Ore, Theory of graphs, Colloq. Pub. 38, Amer. Math. Soc. 1962.

(Received August 15, 1966)