EQUILIBRIUM PROCESSES

BY
SIDNEY C. PORT

1. **Introduction.** Let $(E, \mathfrak{F})$ be a set with a σ-field of subsets $\mathfrak{F}$ containing all one point sets, and let $P(x, B)$ be a transition function of a Markov process with states in E. Assume that P has at least one σ-finite invariant measure λ which we take as fixed throughout the discussion. In §2 we describe precisely how to construct a system of denumerably many independent Markov processes all having the same transition law P and whose initial positions are given by the Poisson process on E with mean $\lambda(B)$ on B. There we establish the fundamental fact that this system maintains itself in macroscopic equilibrium; thus we call such a process an equilibrium process. This property was first established for systems of this type by Doob for Brownian motion and by Derman for countable state space Markov chains. Our purpose in this paper will be to investigate (1) the number of processes, $M_n(B; r)$, whose occupation time in B is exactly r by time n; (2) the number of distinct processes, $L_n(B)$, which are in B at least once by time n; (3) the number of processes, $J_n(B)$, which are in B for a last time at time n; and (4) the number of processes, $A_n(B)$, which are in B at time n, where B is always a transient set of finite, positive λ measure. Previously these quantities were investigated for this system in the countable state space case by the author in [6], and the results we obtain here will be extensions of those in [6] to the more general setting of this paper(1).

In summary, then, we do the following. In §2 we describe the construction of the basic system(2), and in §3 we give some preliminary material on Markov processes having nontrivial dissipative part which is needed in the sequel. In §4 we show that $M_n(B; r)/n$ converges with probability one to a constant $C_r(B)$ and that if $C_r(B) > 0$, then $[M_n(B, r) - EM_n(B, r)](nC_r(B))^{-1/2}$ is asymptotically normally distributed. As a corollary we show that $L_n(B)/n$ converges with probability one to a constant $C(B) > 0$, and that

$$[L_n(B) - EL_n(B)](nC(B))^{-1/2}$$

Presented to the Society, January 25, 1966 under the title *A system of denumerably many transient Markov chains*; received by the editors February 8, 1966.

(1) In the countable case the results were established by probabilistic arguments using the λ-dual process. For the processes considered here, there is, in general, no dual process, and a more analytic approach is needed.

(2) While in the countable case the construction of an equilibrium process is evident, this is no longer true in the general state space case. Some care is required to guarantee that the desired quantities are measurable.

is asymptotically normally distributed. Similar facts are shown to hold for $\sum_{i \leq n} J_i(B)$ with the same constant $C(B)$. We then show that this constant $C(B)$ is the capacity of B. In §5 we show that $S_n(B)/n = [A_1(B) + \cdots + A_n(B)]/n$ converges with probability one to $\lambda(B)$, and that whenever $\int_B E_x N(B)^2 d\lambda(x) < \infty$ (where $N(B)$ is the total occupation time in B for a single process with transition law P), then there is a constant $0 < \sigma^2(B) < \infty$, such that $[S_n(B) - n\lambda(B)](\sigma^2(B)n)^{-1/2}$ is asymptotically normally distributed. The paper concludes, in §6, with some examples.

2. **The equilibrium process.** Let $(E, \mathfrak{F}, \lambda)$ be a σ-finite measure space, where $\mathfrak{F}$ is a σ-field which contains all one point subsets of E. A Poisson process $A(\cdot)$, as defined by Moyal [5], is a stochastic counting measure (or equivalently, a symmetric random point process) on the sets of $\mathfrak{F}$, which is uniquely determined by its probability generating functional,

$$E\left(\exp\left[-\int s(x)dA(x)\right]\right) = \exp\left[\int \left(e^{-s(x)} - 1\right) d\lambda(x)\right], \tag{1}$$

where $s(x) \geq 0$ is a measurable function on E. The salient facts about this process (and the only facts we shall need) are the following:

(1) If $\{E_k\}$ is any denumerable partition of E into disjoint, measurable sets of finite measure, then $\{A(E_k)\}$ are independent, Poisson random variables with $EA(E_k) = \lambda(E_k)$.That is, the number of points laid down in E_k by the process has a Poisson distribution with mean $\lambda(E_k)$.

(2) Given $A(E_k) = r$, and if $\gamma_1, \gamma_2, \cdots, \gamma_r$ represent the position of the r-points in E_k, then the γ_i are independent random variables in E_k, each with distribution

$$P(\gamma_i \in dx) = \frac{\lambda(dx)}{\lambda(E_k)} \quad \text{for} \quad x \in E_k$$

and

$$P(\gamma_i \in dx) = 0 \quad \text{for} \quad x \notin E_k.$$

In other words the γ_i are independent and uniformly distributed on E_k.

In intuitive terms the system we have in mind can be described as follows: At time 0 we place particles in E according to a Poisson process. Subsequently, we allow each particle to move, independently of the others, according to the laws of the same Markov process. At a later time n, we are interested in various facts about this system (e.g., the number of particles which visit a set B, etc.).

In the general state space considered here, the rigorous construction of such a system requires some thought, since trying to follow the intuitive picture too closely may lead to nonmeasurable quantities. For our purposes an adequate model may be made using facts (1) and (2). Let $\{E_k\}$ be an arbitrary (but fixed) partition of E into disjoint measurable sets of finite measure. For each k, let $\{X_i(t,k)\}$ be a sequence of independent Markov processes on $(E, \mathfrak{F})$ with the same transition operator $P(x, B)$, and having initial distribution $\lambda(dx)/\lambda(E_k)$ on E_k. Also let A_k be a Poisson random variable, $EA_k = \lambda(E_k)$, and A_k, $\{X_i(t,k)\}$ are independent. For different k, the collections $\{A_k, X_i(t,k)\}$ are independent. Then the system of Markov processes is the collection

$$\{X_i(t,k), 1 \leqq i \leqq A_k, 1 \leqq k < \infty\}. \tag{2}$$

For ease in language we may think of each $X_i(t,k)$ as the location of the i,kth particle at time t. We shall be concerned with the number of particles, $M_n(B; r)$, which visit a set $B \in \mathfrak{F}$ exactly r times by time n, and several other quantities which can be defined in terms of the $M_n(B; r)$. For completeness we sketch below a proof that $M_n(B; r)$ is a random variable on the probability space of the process given in (2).

PROPOSITION. *Let (Ω, B, P) be a probablity space upon which the process $\{X_i(t,k), A_k\}$ is defined. Then, $M_n(B,r)$ is a measurable function on this space.*

Proof. Clearly $M_n(B,r) = \sum_{k=1}^{\infty} M_n(B,r,k)$, where

$$M_n(B,r,k) = \sum_{i=1}^{A_k} \delta\left(\sum_{l=1}^{n} 1_B(X_i(l,k)), r\right).$$

Here and in the following $1_B(x)$ is the characteristic function of B, and $\delta(x,y) = 1$ if $x = y$ and $\delta(x,y) = 0$ if $x \neq y$. Now for any integer $j \geqq 0$,

$$[M_n(B,r,k) = j] = \bigcup_{m=1}^{\infty} [A_k = m] \cap \left[\sum_{i=1}^{m} \delta\left(\sum_{l=1}^{n} 1_B(X_i(l,k)), r\right) = j\right]$$

Since $[A_k = m] \in \mathrm{B}$, and since $\sum_{l=1}^{n} 1_B(X_i(l,k))$ is a measurable function, we see that

$$[A_k = m] \cap \left[\sum_{i=1}^{m} \delta\left(\sum_{l=1}^{n} 1_B(X_i(l,k)), r\right) = j\right] \in \mathrm{B}.$$

Consequently each $M_n(B,r,k)$ is a measurable function, and thus so is $M_n(B,r)$.

We obtain particularly interesting results when the measure λ is an invariant measure of the transition operator P. That is, if for any $B \in \mathfrak{F}$, and any positive integer n,

$$\lambda(B) = \int_E P^n(x,B)\, d\lambda(x),$$

where P^n is the nth power of the transition operator P. For continuous-time processes, this was first pursued by Doob for Brownian motion, while for discrete-time Markov chains, it was done by Derman. They both established the corresponding version of Theorem 2.1 given below.

DEFINITION. *An equilibrium process is a collection of processes, as described in* (2), *where* λ *is an invariant measure for the transition law* P.

Our first result will be to justify the above name by showing that such a process maintains itself in macroscopic equilibrium in the following sense.

THEOREM 2.1. *Let* $\{F_k\}$ *be any partition of* E *into disjoint measurable sets of finite measure. Let* $A_n(F_k)$ *be the number of particles in* F_k *at time* n. *Then for fixed* n, *the* $\{A_n(F_k)\}$ *are independent Poisson distributed random variables with means* $EA_n(F_k) = \lambda(F_k)$, *respectively.*

Proof. Let

$$A_n(F_i, k) = \sum_{j=1}^{A_k} 1_{F_i}(X_j(n,k)).$$

Then for $0 < s_i < 1$ and any integer $r \geqq 1$,

$$E\left(\prod_{i=1}^{r} s_i^{A_n(F_i,k)} \Big| A_k\right) = \left[1 + \sum_{i=1}^{r} (s_i - 1) \int_{E_k} \frac{d\lambda(x)}{\lambda(E_k)} P^n(x, F_k)\right]^{A_k},$$

and thus

$$E\left(\prod_{i=1}^{r} s_i^{A_n(F_i,k)}\right) = \exp\left[\sum_{i=1}^{r} (s_i - 1) \int_{E_k} d\lambda(x) P^n(x, F_i)\right].$$

By the construction of an equilibrium process, the sequences $\{A_n(F_i, k), 1 \leqq i < \infty\}$ are independent. Consequently the $\{A_n(F_i, k)\}$ are independent Poisson variables with means $\int_{E_k} d\lambda(x) P^n(x, F_i)$ respectively. But

$$A_n(F_i) = \sum_{k=1}^{\infty} A_n(F_i, k),$$

and since

$$EA_n(F_i) = \sum_{k=1}^{\infty} \int_{E_k} d\lambda(x)\ P^n(x,F_i) = \int_E d\lambda(x) P^n(x, F_i) = \lambda(F_i),$$

we see that the $\{A_n(F_i)\}$ are independent, Poisson variables with means $\lambda(F_i)$, respectively. This completes the proof.

3. **Preliminaries.** Assume that $P(x, B)$ is the transition function of a Markov process $\{X_n\}$ with states in the measurable space $(E, \mathfrak{F})$, where $\mathfrak{F}$ is a σ-field of subsets of E containing all one point sets. Assume also that $P(x, B)$ has at least one σ-finite invariant measure λ, which henceforth is taken as fixed. For a function $f \in L_1(\lambda)$, let $Pf(x) = \int P(x, dy) f(y)$.

PROPOSITION 1. *If* $f \in L_1$, *then* $Pf \in L_1$ *and* $\|Pf\|_1 \leqq \|f\|_1$, *so that* P *is a positive, contraction operator on* L_1. *Moreover, if* $f \geqq 0$, *then* $\|Pf\|_1 = \|f\|_1$. *Let* $\hat{P}$ *be the adjoint of* P *on* $L_\infty(\lambda)$. *Then* $\hat{P}g \geqq 0$ *a.e. if* $g \geqq 0$, *and* $\hat{P}1 = 1$ *a.e.*

Proof. Let 1_A be the characteristic function of A. Then $\| P1_A \|_1 = \| 1_A \|_1$ whenever $\lambda(A) < \infty$. If f_n is a nonnegative simple function, then, by linearity of P,

$$\| Pf_n \|_1 = (1, Pf_n) = \| f_n \|_1,$$

where for $g \in L_\infty$ and f in L_1, $(g,f) = \int gf d\lambda$. Hence if $0 \leqq f_n \uparrow f$, $\| Pf \|_1 = \lim_n \| Pf_n \|_1 = \lim_n \| f_n \|_1 = \| f \|_1$. This establishes the assertions about P. Let f be an arbitrary function in L_1; then $f = f^+ - f^-$ where f^+ and f^- are $\geqq 0$. Thus

$$\begin{aligned}(\hat{P}1 - 1, f) &= (\hat{P}1 - 1, f^+) - (\hat{P}1 - 1, f^-) \\ &= (1, Pf^+) - \| f^+ \|_1 - (1, Pf^-) + \| f^- \|_1 = 0.\end{aligned}$$

Thus $\hat{P}1 = 1$ a.e. Finally let $g \geqq 0 \in L_\infty$. Then for any $f \geqq 0$ in L_1, $(\hat{P}g, f) = (g, Pf) \geqq 0$.

If $\hat{P}g < 0$ on a set of positive measure, then (since λ is σ-finite) $\hat{P}g < 0$ on a set A of finite, positive measure. Hence $(\hat{P}g, 1_A) < 0$, a contradiction.

Let $G = \sum_{n=0}^{\infty} P^n$, where here and in the following any operator to the 0th power is the identity. A set $B \in \mathfrak{F}$ is called *transient* if $G1_B(x) < \infty$ a.e.

The *hitting time of B* after time 0, V_B, is the random variable,

$$V_B = \min\{n > 0: X_n \in B\} (= \infty \text{ if } X_n \notin B \text{ for all } n > 0).$$

Since λ is an invariant measure, we see that $\lambda(B) > 0$ implies that $\lambda\{x: P_x(V_B < \infty) > 0\} > 0$. (Indeed, $\lambda\{x: P(x,B) > 0\} > 0$.)

As is well known (see [2]) the operator P splits E into two disjoint sets, C and D, the conservative and dissipative parts, such that, for any $f \geqq 0$ in $L_1(\lambda)$, $(Gf)(x) = \infty$, a.e. on $C \cap \{x: (Gf)(x) > 0\}$, while $(Gf)(x) < \infty$ for a.e. $x \in D$. The set C is stochastically closed (see Theorem 2.2 of [2]). That is, $P(x, C) = 1$ a.e., $x \in C$. Hence $G(x, D) = 0$, a.e. $x \in C$, and thus as λ is σ-finite, D is a denumerable union of transient sets of finite measure. On the other hand, if $B = \bigcup_n B_n$, where B_n are transient, then $\lambda(B \cap C) = 0$. For if $\lambda(B \cap C) > 0$, then $\lambda(B_n \cap C) > 0$ for some n; and consequently for some $F \subset B_n \cap C$ having finite positive measure. $(G1_F)(x) \leqq (G1_{B_n})(x) < \infty$ a.e., $x \in F$. But $F \subset C$, and thus $\lambda(F) > 0$ implies $(G1_F)(x) = \infty$ a.e. $x \in F$, a contradiction. In particular, if $\lambda(C) = 0$ (i.e., P is dissipative), then E is a countable union of transient sets of finite measure, and every set of finite measure is transient. *Let $\mathfrak{T}$ be the class of all transient sets of finite, positive measure.*

Define, for sets $A, B \in \mathfrak{T}$, and $n \geqq 1$,

$$(3.1) \qquad {}_BP^n(x, A) = P_x(X_n \in A, V_B \geqq n).$$

Then if I_B is the operator $(I_B f)(x) = f(x)$, $x \in B$, and $(I_B f)(x) = 0$, $x \notin B$, we see at once that,

$$_BP^n(x,A) = {}_BP^n 1_A(x) = P(I_{B'}P)^{n-1}1_A(x) = (PI_{B'})^{n-1}P1_A(x),$$

where $B' = E - B$.

The adjoint $_B\hat{P}^n$ of $_BP^n$ is the operator on L_∞ given by

$$_B\hat{P}^n = (\hat{P}I_{B'})^{n-1}\hat{P} = \hat{P}(I_{B'}\hat{P})^{n-1}. \tag{3.2}$$

The following facts will be used in the sequel.

Proposition 2. *If $B \in \mathfrak{X}$, then*

$$C(B) = \int_B P_x(V_B = \infty)\,d\lambda(x) > 0. \tag{3.3}$$

Proof. By (2.3) of [1] we see that for $x \in B$,

$$1 = \int_B G(x,dy)P_y(V_B = \infty).$$

Hence, if we set $e_B(x) = P_x(V_B = \infty)1_B(x)$, then

$$\begin{aligned}\lambda(B) &= \int_B \int_B G(x,dy)P_y(V_B = \infty)\,d\lambda(x) = \sum_{k=0}^{\infty}(1_B, P^k e_B)\\ &= \sum_{k=0}^{\infty}(\hat{P}^k 1_B, e_B).\end{aligned}$$

If $C(B) = 0$, then $e_B = 0$ a.e. on B and the last term on the right is 0, which is impossible since $\lambda(B) > 0$.

Proposition 3. *Assume $B \in \mathfrak{X}$; then*

$$_B\hat{P}^n 1(x) \downarrow \hat{e}_B(x) \geqq 0 \text{ a.e.} \tag{3.4}$$

$$\int_B \hat{e}_B(x)\,d\lambda(x) = C(B), \tag{3.5}$$

where $C(B)$ is as in Proposition 2.

Proof. By Proposition 1, $\hat{P}1 \leqq 1$, a.e. Assume

$$_B\hat{P}^n 1 \leqq {}_B\hat{P}^{n-1}1, \text{ a.e.}$$

Then a.e.,

$$_B\hat{P}^{n+1}1 = \hat{P}I_{B'}({}_B\hat{P}^n 1) \leqq (\hat{P}I_{B'}){}_B\hat{P}^{n-1}1 = {}_B\hat{P}^n 1.$$

Hence, by induction, $_B\hat{P}^n 1$ is nonincreasing a.e., and thus

$$\lim_{n\to\infty} {}_B\hat{P}^n 1(x) = \hat{e}_B(x)$$

exists a.e. To establish (3.5) we proceed as follows:

$$\int_E P_x(V_B = n)\,d\lambda(x) = (1, {}_BP^n 1_B) = ({}_B\hat{P}^n 1, 1_B) = \int_B {}_B\hat{P}^n 1(x)\,d\lambda(x).$$

Consequently, by bounded convergence, we see that

$$\lim_{n\to\infty} \int_B P_x(V_B = n)\,d\lambda(x) = \int_B \hat{e}_B(x)\,d\lambda(x)$$

and thus

$$\lim_{n\to\infty} \frac{1}{n}\int_E P_x(V_B \leqq n)\,d\lambda(x) = \int_B \hat{e}_B(x)\,d\lambda(x).$$

On the other hand, a last entrance decomposition yields

$$\begin{aligned}\int_E P_x(V_B \leqq n)\,d\lambda(x) &= \int_E \left[\int_B \sum_{m=1}^n P^m(x,dy)\,P_y(V_B > n-m)\right] d\lambda(x)\\ &= \int_B \sum_{m=1}^n d\lambda(y)\,P_y\,(V_B > n-m),\end{aligned}$$

and by bounded convergence,

$$\lim_{n\to\infty} \int_B \frac{1}{n}\sum_{m=1}^n d\lambda(y) P_y(V_B > n-m) = \int_B P_y(V_B = \infty)\,d\lambda(y) = C(B).$$

Hence

$$C(B) = \int_B \hat{e}_B(x)\,d\lambda(x).$$

We conclude this section by introducing the terminology to be used throughout the remainder of the paper. Let $\{X_n\}$ be a Markov process having the function $P(x,B)$ for its transition law. Then the *successive visits to* B take place at times V_B^r, which are defined as follows: For convenience set $V_B^0 = 0$. Then $V_B^1 = V_B$, and

$$V_B^r = \min\{n > V_B^{r-1} : X_n \in B\}$$

$$(= \infty \text{ if } V_B^{r-1} = \infty \text{ or } X_n \notin B \text{ for all } n > V_B^{r-1}).$$

The occupation time of B by time n, $N_n(B) = \sum_{j=1}^n 1_B(X_j)$, and $N(B) = \lim_{n\to\infty} N_n(B)$. If B is transient, then $E_x N(B) = G(x,B) < \infty$, a.e.; thus $P_x(N(B) < \infty) = 1$, a.e. *The time of last visit to* B, $T_B = \min\{n \geqq 0 : X_j \notin B \text{ all } j > n\}$. Then whenever B is transient, $P_x(T_B < \infty) = 1$, a.e.

4. **Multiplicity of particles.** Throughout this section we will assume that we have a given equilibrium process as described in §2, and that P is nonconservative, i.e., $\lambda(D) > 0$, where D is the dissipative part of E.

THEOREM 4.1. *Let* $B \in \mathfrak{X}$, *and let* $M_n(B;r)$ *denote the number of particles which have visited* B *exactly* r *times by time* n. *Then*

$$P\left(\lim_{n\to\infty}\frac{M_n(B;r)}{n}=C_r(B)\right)=1, \tag{4.1}$$

where $C_r(B)=\int_B d\lambda(x)\hat{e}_B(x)P_x(N(B)=r-1)$. *Moreover, for fixed* n *and* B, $M_n(B;j)$, $1\leqq j\leqq n$, *are independent, Poisson distributed random variables with means*

$$EM_n(B,r) = \int_E d\lambda(x)P_x(N_n(B)=r).$$

If $C_r(B)>0$, then

$$\lim_{n\to\infty} P\left(\frac{M_n(B;r)-EM_n(B;r)}{(nC_r(B))^{1/2}}\leqq x\right) = \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{x}\exp(-t^2/2)\,dt. \tag{4.2}$$

Proof. Let $I_n(B,r)$ be the number of particles which visit B for the rth time at time n(3).

LEMMA 4.1. *For fixed* B *and* r, $I_n(B,r)$ *are independent Poisson distributed random variables with means*

$$EI_n(B;r)=\int_E P_x(V_B^r=n)\,d\lambda(x).$$

Proof. Let $I_n(B,r,k)$ denote the number of particles which start in E_k and visit B for the rth time at time n.

Now for any $n\geqq 1$,

$$E\prod_{i=1}^{n} s_{ii}^{I_i(B,r,k)} = E\left[1+\sum_{i=1}^{n}(s_i-1)\int_{E_k}\frac{d\lambda(x)}{\lambda(E_k)}P_x(V_B^r=i)\right]^{A_k}$$
$$=\exp\left[\sum_{i=1}^{n}(s_i-1)\int_{E_k}d\lambda(x)P_x(V_B^r=i)\right].$$

Thus for fixed k, $I_n(B,r,k)$ are independent Poisson variables with means $\int_{E_k}d\lambda(x)P_x(V_B^r=n)$. Since

$$I_n(B,r)=\sum_{k=1}^{\infty} I_n(B,r,k),$$

and for $k=1,2,\cdots$, the sequences $\{I_n(B,r,k),\ 1\leqq n<\infty\}$, are independent, we see that the $I_n(B,r)$ are independent Poisson with means, $\sum_k\int_{E_k}d\lambda(x)P_x(V_B^r=n)$, respectively.

LEMMA 4.2. *For fixed* $B\in\mathfrak{X}$ *and* $r\geqq 1$,

$$\lim_{n\to\infty} EI_n(B,r) = \int_B d\lambda(x)\hat{e}_B(x)P_x(V_B^{r-1}<\infty) = C_r^*(B).$$

(3) $I_n(B,r)$ can be defined in terms of $M_n(B,r)$; see equation (4.3).

Proof. Let $\phi_n(y) = P_y(V_B^{r-1} = n)1_B(y)$. Then

$$EI_n(B,r) = \int_E P_x(V_B^r = n)d\lambda(x)$$

$$= \int_E \sum_{j=1}^n \int_B {}_BP^j(x,dy)P_y(V_B^{r-1} = n-j)\,d\lambda(x)$$

$$= \sum_{j=1}^n (1, {}_BP^j\phi_{n-j}) = \sum_{j=1}^n ({}_B\hat{P}^j1, \phi_{n-j}).$$

Now by Proposition 3, ${}_B\hat{P}^n1(y) \to \hat{e}_B(y)$ a.e. Moreover,

$$\sum_n \phi_n(y) = P_y(V_B^{r-1} < \infty)1_B(y).$$

Hence a.e.

$$\lim_{n\to\infty} \sum_{j=1}^n {}_B\hat{P}^j1(y)\phi_{n-j}(y) = \hat{e}_B(y)P_y(V_B^{r-1} < \infty)1_B(y).$$

But ${}_BP^j1(y) \leqq 1$ a.e., and thus

$$\sum_{j=1}^n {}_B\hat{P}^j1(y)\phi_{n-j}(y) \leqq \sum_{j=1}^n \phi_{n-j}(y) \leqq P_y(V_B^{r-1} \leqq n)1_B(y) \leqq 1_B(y), \text{ a.e.}$$

Thus by dominated convergence,

$$\lim_{n\to\infty} \int_E \sum_{j=1}^n {}_B\hat{P}^j1(y)\phi_{n-j}(y)d\lambda(y) = \int_B \hat{e}_B(y)P_y(V_B^{r-1} < \infty)d\lambda(y).$$

We may now establish (4.1). From Lemmas 4.1 and 4.2 we see at once that $\sup EI_n(B,r) < \infty$ and that

$$\sum_n \frac{\operatorname{Var} I_n(B,r)}{n^2} = \sum_n \frac{EI_n(B,r)}{n^2} < \infty.$$

Consequently, by the strong law of large numbers,

$$\lim_{n\to\infty} \sum_{j=1}^n \left[\frac{I_j(B,r) - EI_j(B,r)}{n}\right] = 1$$

with probability one. But Lemma 4.2 implies that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n EI_j(B,r) = C_r^*(B)$$

and thus,

$$P\left(\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n I_j(B,r) = C_r^*(B)\right) = 1.$$

However,

$$\sum_{j=1}^{n} I_j(B,r) = \sum_{k=r}^{\infty} M_n(B;k) \tag{4.3}$$

and thus (after a slight computation) we find that (4.1) holds.

To show that $M_n(B;r)$, $1 \leqq r \leqq n$, are independent Poisson variates, it suffices to show that $M_n(B;r,k)$ are, where $M_n(B;r,k)$ are the number of particles, intially in E_k, which visit B exactly r times by time n. This may be accomplished by a computation similar to that used in the proof of Lemma 4.1, which also shows that

$$EM_n(B;r,k) = \int_{E_k} P_x(N_n(B) = r)\,d\lambda(x).$$

Finally if $C_r(B) > 0$, then $EM_n(B;r) \sim nC_r(B)$ and thus $EM_n(B;r) \to \infty$, $n \to \infty$. Using the well-known fact that a normalized Poisson variate is asymptotically normally distributed when its mean tends to infinity, we find that (4.2) holds. This completes the proof.

Examples show that $C_r(B)$ may vanish for infinitely many values of r. However, since $\sum_{r=1}^{\infty} C_r(B) = C(B) > 0$, there must be at least one value of r such that $C_r(B) > 0$. Under some irreducibility assumptions, we may establish that for every $r > 0$, $\sum_{j=r}^{\infty} C_j(B) = C_r^*(B) > 0$.

LEMMA 4.3. *Assume that either* (i) *for every A having positive measure, $P_x(V_A < \infty) > 0$, a.e., or* (ii) *that for the set $B \in \mathfrak{X}$ in question, $P_x(V_B < \infty) > 0$ for all $x \in B$. Then for every $r \geqq 1$,*

$$\int_B \hat{e}_B(x) P_x(V_B^r < \infty)\,d\lambda(x) > 0. \tag{4.4}$$

Proof. Assume (i) holds and assume $\lambda(B) > 0$. Define $A_r = \{x : P_x(V_B^r < \infty) > 0\}$. By assumption, $\lambda(E - A_1) = 0$. Assume $\lambda(E - A_r) = 0$. Then $\lambda(A_r \cap B) > 0$ and thus $P_x(V_{B \cap A_r} < \infty) > 0$, a.e. However, for each such x, $P_x(V_B^{r+1} < \infty) > 0$ and thus $\lambda(E - A_{r+1}) = 0$. Hence, by induction, $P_x(V_B^r < \infty) > 0$ a.e. Since $\int_B \hat{e}_B(x)\,d\lambda(x) > 0$ we see that (4.4) holds. Now assume (ii) holds. Let $\pi_B(x,dy) = P_x(V_B < \infty;\, X_{V_B} \in dy)$. Then for every $x \in B$,

$$P_x(V_B^{r+1} < \infty) = \int_B \pi_B(x,dy) P_y(V_B^r < \infty) > 0,$$

since otherwise $\pi_B(x,B) = 0$ for some $x \in B$, a contradiction. From this, (4.4) follows as before.

From Lemmas 4.1 through 4.3 we easily obtain the following.

THEOREM 4.2. *Under the conditions of Lemma* 4.3, (4.5) $$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n} I_j(B,r)=C_r^*(B)\right)=1.$$

Moreover, $I_n(B,r)$ converges in law to a Poisson random variable $I(B,r)$ having mean $C_r^*(B)$, and finally

(4.6) $$\lim_{n\to\infty} P\left(\sum_{j=1}^{n}[I_j(B;r) - EI_j(B,r)](nC_r^*(B))^{-1/2}\leqq x\right) = \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{x}\exp(-t^2/2)\,dt.$$

Now for $r=1$, $C_1^*(B)=\sum_{r=1}^{\infty}C_r(B)=C(B)$, which is always positive by Proposition 2. If $L_n(B)$ is the number of distinct particles to visit B by time $\leqq n$, then clearly $L_n(B)=\sum_{j=1}^{n}I_j(B,1)$, and thus by Theorem 4.2 we obtain the following.

COROLLARY 4.1. *If $L_n(B)$ is the number of distinct particles to visit B by time $\leqq n$, and $B\in\mathfrak{T}$, then*

(4.7) $$P\left(\lim_{n\to\infty}\frac{L_n(B)}{n}=C(B)\right)=1$$

and

(4.8) $$\lim_{n\to\infty} P\left(\frac{L_n(B)-EL_n(B)}{(nC(B))^{1/2}}\leqq x\right)=\frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{x}\exp(-t^2/2)\,dt.$$

The last quantities to be investigated are the number of particles, $J_n(B)$, which are in B for a last time at time n.

THEOREM 4.3. *Let $B\in\mathfrak{T}$. Then $\{J_n(B)\}$ are independent, Poisson distributed, random variables with the common mean $C(B)$. Moreover,*

(4.9) $$P\left(\frac{J_1(B)+\cdots+J_n(B)}{n}\to C(B)\right) = 1$$

and

(4.10) $$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n}J_i(B)-nC(B)}{(nC(B))^{1/2}}\leqq x\right) = \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{x}\exp(-t^2/2)\,dt.$$

Proof. Clearly (4.9) and (4.10) are obvious consequences of the fact that the $J_n(B)$ are independent and identically distributed Poisson variables. To establish this fact, let $J_n(B;k)$ be the number of particles in E_k at time 0 which last visit B at time n. Then for $k=1,2,\cdots$, the sequences $\{J_n(B,k), 1\leqq n<\infty\}$ are independent. Moreover for any $n\geqq 1$,

$$E\left(\prod_{i=1}^{n} s_i^{J_i(B,k)}\right) = E\left[1+\sum_{i=1}^{n}(s_i-1)\int_{E_k}\frac{d\lambda(x)}{\lambda(E_k)}P_x(T_B=i)\right]^{A_k}$$

$$= \exp\left[\sum_{i=1}^{n}(s_i-1)\int_{E_k} d\lambda(x)P_x(T_B=i)\right].$$

Hence $J_n(B,k)$ are independent Poisson variables with means $\int_{E_k} d\lambda(x)P_x(T_B=n)$. Consequently, $J_n(B)$ are independent Poisson variables with means

$$EJ_n(B) = \int_E d\lambda(x)P_x(T_B=n) = \int_E d\lambda(x)\int_B P^n(x,dy)P_y(V_B=\infty)$$

$$= \int_B d\lambda(y)P_y(V_B=\infty) = C(B).$$

We conclude this section by sketching the interpretation of $C(B)$ as the *capacity* of B.

A measurable function f is the potential of a *charge* ϕ if $\phi \in L_1(\lambda)$ and $G\phi = f$. A set B is called an *equilibrium set* if there is a potential $\leqq 1$ everywhere, equal to 1 on B, and having nonnegative charge ψ with support on B. This charge is uniquely determined by such a potential (since $PG\psi \leqq 1$) and its total charge $\int_E \psi d\lambda$ is the *capacity* of B.

If $B \in \mathfrak{T}$, then B is an equilibrium set, since

$$\phi_B(x) = \int_E G(x,dy)e_B(y),$$

where $e_B(y)$ is $1_B(y)P_y(V_B=\infty)$ and $V_B(x)=1$, $x \in B$, and $\phi_B(x) = P_x(V_B<\infty)$, $x \notin B$. Thus $C(B)$ is the capacity of B.

REMARK. One of the main points of this section was to show that capacity may be interpreted probabilistically for sets in $\mathfrak{T}$ as the number of new particles per unit time which enter B and also as the number of particles per unit time which leave B, never to return. That these two quantities are equal is again testimony to the macroscopic equilibrium nature of the process. It also follows from Lemma 4.2 (for $r=1$) that for $B \in \mathfrak{T}$

$$C(B) = \lim_{n\to\infty}\frac{EL_n(B)}{n} = \lim_{n\to\infty}\frac{1}{n}\int_E P_x(V_B \leqq n)d\lambda(x). \tag{4.11}$$

Choquet has axiomatized capacities as nonnegative, nondecreasing, set functions which are alternating in the sense that for any $r \geqq 1$,

$$C(A_1\cap\cdots\cap A_r) \leqq \sum_{i=1}^{r} C(A_i) - \sum_{i\neq j} C(A_i\cup A_j) + \cdots + (-1)^r C(A_1\cup\cdots\cup A_r). \tag{4.12}$$

Using (4.11) it is easy to show that our set function $C(\cdot)$ defined on the sets of $\mathfrak{T}$, satisfies these requirements. Indeed if $A \subset B$, then $P_x(V_A \leqq n) \leqq P_x(V_B \leqq n)$ and thus $EL_n(A) \leqq EL_n(B)$. Hence $C(A) \leqq C(B)$. Moreover,

$$P_x(V_{A_1 \cap \cdots \cap A_r} \leqq n) \leqq P_x(V_{A_i} \leqq n, 1 \leqq i \leqq r) = \sum_{i=1}^{r} P_x(V_{A_i} \leqq n)$$
$$= \sum_{i \neq j} P_x(V_{A_i \cup A_j} \leqq n) + \cdots + (-1)^r P_x(V_{A_1 \cup \cdots \cup A_r} \leqq n),$$

and consequently,

$$EL_n(A_1 \cap \cdots \cap A_r)$$
$$\leqq \sum_{i=1}^{r} EL_n(A_i) - \sum_{i \neq j} EL_n(A_i \cup A_j) + \cdots + (-1)^r EL_n(A_1 \cup \cdots \cup A_r).$$

Hence (4.12) holds.

Now observe that (4.7) and the fact that $EJ_n(B) = C(B)$ yield some interesting facts about $J_n(\cdot)$ and EJ_n as a *set* function, which are not at all apparent. For example, since $C(\cdot)$ is a nondecreasing set function, so is $EJ_n(\cdot)$. Thus the expected number of permanent departures at any time from a larger set is at least as great as that from a smaller.

Finally we note that (4.1) asserts that the percentage of particles, $M_n(B;r)/n$, to visit B exactly r times by time n converges to the integral of $P_x(N(B) = r-1)$ by the "dual capacitory measure" of B. This makes sense in view of the interpretation of capacity given before.

5. **Number of particles.** In this section the same basic assumptions as in §4 will prevail. Let $S_n(B)$ be the nth partial sum, $A_1(B) + \cdots + A_n(B)$, of the $A_i(B)$, where $A_i(B)$ is the number of particles in B at time i.

THEOREM 5.1. *If* $B \in \mathfrak{T}$, *then*

$$P\left(\lim_{n\to\infty} \frac{S_n(B)}{n} = \lambda(B)\right) = 1. \tag{5.1}$$

Moreover, if $\int_B E_x N(B)^2 d\lambda(x) < \infty$(4), *then*

$$\lim_{n\to\infty} P\left(\frac{S_n(B) - n\lambda(B)}{(n\sigma^2(B))^{1/2}} \leqq x\right) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{x} \exp(-u^2/2)\, du, \tag{5.2}$$

where

$$\sigma^2(B) = \lambda(B) + 2\int_B d\lambda(x) E_x N(B). \tag{5.3}$$

(4) It is easy to show that this condition is satisfied if and only if $\int_B \int_B G(x, dy) G(y, B) d\lambda(x) < \infty$.

Proof. A direct (but slightly tedious) computation shows that the sequence $\{A_n(B)\}$ is strictly stationary, and thus by the pointwise ergodic theorem,

$$P\left(\lim_{n\to\infty} \frac{S_n(B)}{n} = S^*\right) = 1.$$

To establish (5.1) we must show that $P(S^* = \lambda(B)) = 1$. This may be accomplished by showing

$$\lim_{n\to\infty} P\left(\left|\frac{S_n(B)}{n} - \lambda(B)\right| > \varepsilon\right) = 0. \tag{5.4}$$

By Chebyshev's inequality,

$$P\left(\left|\frac{S_n(B)}{n} - \lambda(B)\right| > \varepsilon\right) \leqq \frac{\operatorname{Var} S_n(B)}{\varepsilon^2 n^2};$$

so to establish (5.4), all we need do is show

$$\lim_{n\to\infty} \frac{\operatorname{Var} S_n(B)}{n^2} = 0. \tag{5.5}$$

Since $S_n(B) = \sum_{r=1}^n rM_n(B,r)$, and the $M_n(B;r)$ for $1 \leqq r \leqq n$ are independent Poisson variates, we see that

$$\begin{aligned}\operatorname{Var} S_n(B) &= \sum_{r=1}^n r^2 EM_n(B;r) = \int_E \sum_{r=1}^n r^2 P_x(N_n(B) = r)\, d\lambda(x) \\ &= \int_E E_x N_n(B)^2 d\lambda(x).\end{aligned} \tag{5.6}$$

Now

$$E_x N_n(B)^2 = E_x N_n(B) + 2 \sum_{1 \leqq i < j \leqq n} \int_B P^i(x,dy) P^{j-i}(y,B),$$

and thus

$$\operatorname{Var} S_n(B) = n\lambda(B) + 2 \sum_{i=1}^{n-1} \int_B E_y N_{n-i}(B)\, d\lambda(y). \tag{5.7}$$

Since B is a transient set, $\lim_{n\to\infty} P^n(y,B) = 0$ a.e., and as B has finite measure, bounded convergence implies $\int_B P^n(y,B)\, d\lambda(y) \to 0$, and thus $1/n \int_B E_y N_n(B)\, d\lambda(y) \to 0$. From (5.7) we see that (since $E_y N_n(B)$ is nondecreasing)

$$\operatorname{Var} S_n(B) \leqq n\lambda(B) + 2n \int_B E_y N_n(B) d\lambda(y),$$

and thus (5.5) holds.

Now if $\int_B E_y N(B)\, d\lambda(y) < \infty$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n-1} \int_B E_y N_i(B) d\lambda(y) = \int_B E_y N(B) d\lambda(y) < \infty.$$

Consequently from (5.7) we obtain

$$\sum_{r=1}^{n} r^2 EM_n(B;r) = \operatorname{Var} S_n(B) \sim n\left[\lambda(B) + 2\int_B E_y N(B)\,d\lambda(y)\right]$$

(5.8)

$$= n\sigma^2(B).$$

Recall that if $\phi(\theta)$ is the characteristic function of an infinitely divisible law with finite variance, then the Kolmogorov representation of $\phi(\theta)$ is (see [3, p. 307])

$$\log \phi(0) = i\theta\lambda + \int_{-\infty}^{\infty} (e^{i\theta x} - 1 - i\theta x)x^{-2}G(dx).$$

For the random variable

$$Y_n(B;r) = r[M_n(B;r) - EM_n(B;r)](n\sigma^2(B))^{-1/2}$$

λ and G are, respectively, $\lambda_{nr} = 0$ and

$$G_{nr}(x) = 0, \qquad x < r(n\sigma^2)^{-1/2}$$

$$= \frac{r^2 EM_n(B;r)}{n\sigma^2}, \quad x > r(n\sigma^2)^{-1/2}.$$

Hence for the random variable $\sum_{r=1}^{n} Y_n(B;r)$ we find that λ and G are, respectively, $\lambda = 0$ and

$$G_n(x) = \sum_{r=1}^{n} G_{nr}(x) = \sum_{r=1}^{j} \frac{r^2 EM_n(B;r)}{n\sigma^2}, \tag{5.9}$$

if

$$j(n\sigma^2)^{-1/2} \leqq x < (j+1)(n\sigma^2)^{-1/2}.$$

Clearly,

$$G_n(x) \leqq G_n(\infty) = \sum_{r=1}^{n} \frac{r^2 EM_n(B;r)}{n\sigma^2}.$$

Hence, by (5.8),

$$\limsup_{n\to\infty} G_n(x) \leqq \lim_{n\to\infty} G_n(\infty) = 1.$$

On the other hand, by Theorem 4.1 and Fatou's Lemma, we obtain from (5.9) that for any $x > 0$,

$$\liminf_{n\to\infty} G_n(x) \geqq \sum_{r=1}^{\infty} \frac{r^2 C_r(B)}{\sigma^2(B)}$$

$$= \frac{1}{\sigma^2(B)} \sum_{r=1}^{\infty} r^2 \int_B \hat{e}_B(x) P_x(N(B) = r-1)\,d\lambda(x)$$

$$= \frac{1}{\sigma^2(B)} \int_B E_x(1 + N(B))^2 \hat{e}_B(x)\,d\lambda(x).$$

However, from (5.6) we see that

$$\begin{aligned}\operatorname{Var} S_n(B) &= \int_E E_x N_n(B)^2 d\lambda(x)\\ &= \int_E d\lambda(x) \int_B \sum_{j=1}^{n} {}_BP^j(x,dy) E_y\,(1 + N_{n-j}(B))^2\\ &= \int_B \left[\sum_{j=1}^{n} {}_B\hat{P}^j 1(y) E_y(1 + N_{n-j}\,(B)^2 \right] d\lambda(y);\end{aligned}$$

and since the integrand in this last expression is dominated by $nE_y(1 + N(B))^2$, we obtain by the dominated convergence theorem and a simple summability argument that

$$\sigma^2(B) = \lim_{n\to\infty} \frac{1}{n} \operatorname{Var} S_n(B) = \int_B E_x(1 + N(B))^2 \hat{e}_B(x) d\lambda(x).$$

Thus for $x > 0$, $\lim_{n\to\infty} \inf G_n(x) \geqq 1$. Since $G_n(x) = 0$ for $x < 0$, we have shown that

$$\begin{aligned}\lim_{n\to\infty} G_n(x) &= 1, \quad x > 0,\\ &= 0, \quad x < 0.\end{aligned}$$

But this is precisely the G needed to represent the standard normal distribution and thus, by an appeal to a well-known result in the theory of i. d. laws (see [3, p. 312]), we obtain (5.2).

REMARK. The result in (5.1) shows that the sequence $\{A_n(B)\}$ is ergodic, and thus $\lambda(B)$ may be interpreted as the number of particles per unit time in B. The central limit theorem for $S_n(B)$ enables one to approximate the probability distribution of the discrepancy between the time average, $S_n(B)/n$, and the ensemble average, $\lambda(B)$.

6. Examples.

EXAMPLE 1. *Denumerable state space.* Let E be a countable set, $\mathfrak{F}$, all subsets of E, and $p(x, y)$ the transition matrix of a transient Markov chain on E, having λ as an invariant measure. Then E is a denumerable union of transient sets and thus P is dissipative. The class $\mathfrak{T}$ contains at least all sets of finite λ measure. If the chain is irreducible, then the conditions of Lemma 4.3 are also satisfied. This case was treated in detail in [6] for finite sets, so the results we obtain here are slightly more general than previously, since we demand only that $\lambda(B) < \infty$.

EXAMPLE 2. *Random walk on a locally compact group.* Let $\mathfrak{G}$ be a locally compact group, $\mathfrak{F}$ the σ-field generated by the open sets, λ a left Haar measure, and $X_1, X_2, \cdots$ be independent $\mathfrak{G}$-valued random variables, having a common regular probability distribution μ on $\mathfrak{G}$. Set $S_n = X_n X_{n-1} \cdots X_1 S_0$. The transition operator is

$$P(x, B) = \mu(Bx^{-1}) = P(S_n \in B \,|\, S_{n-1} = x).$$

A value $x \in \mathfrak{G}$ is possible if for any neighborhood N of the identity e,

$$P_e(S_n \in Nx) > 0, \quad \text{for some } n.$$

The set $\mathfrak{P}$ of all possible values is a closed semigroup [4]. We will assume that $\mathfrak{G} = \mathfrak{P}$, i.e., the random walk on $\mathfrak{G}$ is "irreducible," and that the walk is transient, i.e., that there is a neighborhood N of e such that $G(e, N) < \infty$.

Although, in general, a locally compact group need not be σ-compact, under the assumption that $\mathfrak{P} = \mathfrak{G}$ it is. The following proof of this fact is due to J. Folkman. Since μ is regular, there is a sequence of compact sets C_n whose union $\bigcup_n C_n = A$ and $\mu(G - A) = 0$. Now there is an open-closed σ-compact subgroup H of $\mathfrak{G}$ (the group generated by a symmetric, compact neighborhood of e will do). Let $B = \bigcup_{n=0}^{\infty} A^n H$. Since H is open, $A^n H$ is open and thus B is open. B is also closed, since for $x \in \mathfrak{G} - B$, $xH \cap B = \emptyset$. Indeed $y \in xH \cap B$ implies $y = xh_1 = \gamma h_2$ and thus $x = \gamma h_2 h_1^{-1} \in B$. Thus B is closed, open, and σ-compact. If $x \in B$, then there is a neighborhood Nx of x such that $Nx \cap B = \emptyset$. For all n, $P(S_n \in B) = 1$ and thus for all $n, P(S_n \in Nx) = 0$; hence $x \in \mathfrak{P}$. Thus $\mathfrak{P} \subset B$; and since B is closed and σ-compact and $\mathfrak{P}$ is closed, we see that $\mathfrak{P}$ is σ-compact.

For any $x \in \mathfrak{G}$, if $BB^{-1} \subset N$ and B open, then

$$G(x, B) \leqq G(e, N).$$

Hence if the random walk is transient, $G(x, B) < \infty$ all x, and thus $G(x, By) < \infty$ for all x, y. Since any compact set C can be covered by finite number of sets By, we see that $G(x, C) < \infty$ for all x and every compact set. Hence $\mathfrak{G}$ is a countable union of compact sets and thus $\mathfrak{P}$ is dissipative. A trivial computation shows that the left Haar measure is an invariant measure for the left random walk. Since Haar measure is finite on compacts, λ is σ-finite in our case. Thus the class $\mathfrak{T}$ contains at least all relatively compact sets having nonempty interior. Moreover, any such set satisfies (ii) of Lemma 4.3, and for each such B,

$$\int_B d\lambda(x) \int_B G(x, dy)\, G(y, B) < \infty,$$

so that the central limit theorem for $S_n(B)$ is applicable.

References

1. J. L. Doob, *Discrete potential theory and boundaries*, J. Math. Mech. **8** (1959), 433–458.
2. J. Feldman, *Subinvariant measures for Markoff operators*, Duke Math. J. **29** (1962), 71–98.
3. B. Gnedenko, *Probability theory*, Chelsea, New York, 1962.
4. R. M. Loynes, *Products of independent random elements in a topological group*, Z. Warscheinlichkeitstheorie und Verw. Gebiete **1** (1962), 446–455.
5. J. E. Moyal, *The general theory of stochastic population processes*, Acta Math. **108** (1962), 1–31.
6. S. C. Port, *A system of denumerably many transient Markov chains*, Ann. Math. Statist. **37** (1966), 406–411.

RAND CORPORATION
SANTA MONICA, CALIFORNIA