

THE CALCULATION OF THE ERGODIC PROJECTION FOR MARKOV CHAINS AND PROCESSES WITH A COUNTABLE INFINITY OF STATES

BY

DAVID G. KENDALL and G. E. H. REUTER

*Magdalen College
Oxford*

*Victoria University
Manchester*

I. Introduction

1.1. Let $P \equiv \{p_{ij} : i, j = 0, 1, 2, \dots\}$ be the matrix of one-step transition probabilities for a temporally homogeneous Markov chain with a countably infinite set of states (labelled as $0, 1, 2, \dots$). The probability p_{ij}^n of a transition in n steps from state i to state j will then be the (i, j) th element of the matrix P^n , so that the specification of P (or equivalently of $\Delta \equiv P - I$) completely determines the system. It is known (Kolmogorov [24]) that the Cesàro limits

$$\pi_{ij} \equiv \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{r=1}^n p_{ij}^r \quad (1)$$

always exist. Let Π denote the matrix whose (i, j) th element is π_{ij} , and consider

Problem A: *Determine Π when P is given.*

This problem has obvious importance for practical applications. A number of special techniques are available for its solution in particular cases (see, e.g., Feller [12], Ch. 15, Foster [14], [15] and Jensen [18]); also Feller [12], pp. 332–4, has given a general iterative method of solution. We shall give another (non-iterative) general method in § 2: it will involve the non-negative solutions of

$$x_j = \sum_{\alpha} x_{\alpha} p_{\alpha j}$$

such that $\sum x_{\alpha} < \infty$, and the non-negative solutions of

$$y_i = \sum_{\alpha} p_{i\alpha} y_{\alpha}$$

such that $\sup_{\alpha} y_{\alpha} < \infty$.

1.2. Let $P_t \equiv \{p_{ij}(t) : i, j = 0, 1, 2, \dots\}$, for each $t \geq 0$, denote the matrix of transition probabilities for a temporally homogeneous Markov process with a countably infinite set of states, such that $p_{ij}(t) \rightarrow \delta_{ij}$ as $t \downarrow 0$. We may then define a one-parameter semigroup $\{P_t : t \geq 0\}$ of transition operators on the Banach space l by setting

$$(P_t x)_j \equiv \sum_{\alpha=0}^{\infty} x_{\alpha} p_{\alpha j}(t) \quad (2)$$

for each element $x \equiv (x_0, x_1, x_2, \dots)$ of l . In exceptionally simple cases we can write $P_t = \exp(\Omega t)$, where Ω is a bounded operator on l , but even for so simple an example as the birth-and-death process (for this see, e.g., Feller [12], pp. 371–5) this is no longer possible. However a generating operator Ω (analogous to Δ in the chain case) can always be defined by writing

$$\Omega x \equiv \text{strong } \lim_{t \downarrow 0} (P_t x - x)/t \quad (3)$$

for $x \in \mathcal{D}(\Omega)$, where $\mathcal{D}(\Omega)$ is by definition that set of elements $x \in l$ for which the limit in (3) exists. In general Ω will be an unbounded operator, but its domain $\mathcal{D}(\Omega)$ is always dense in l and Ω determines the system uniquely (hence it is called the *infinitesimal generator*).

It is also known (Lévy [27]) that the ordinary limits

$$\pi_{ij} \equiv \lim_{t \rightarrow \infty} p_{ij}(t) \quad (4)$$

exist (no Cesàro averaging being needed now); thus if we again write Π for the matrix whose (i, j) th element is π_{ij} , we are led to consider

Problem B: *Determine Π when Ω is given.*

A solution to this problem will be given in § 3: it will involve the non-negative elements in the nullspaces of Ω and Ω^* (the operator adjoint to Ω).

1.3. For any such Markov process, the limits

$$q_{ij} \equiv \lim_{t \downarrow 0} (p_{ij}(t) - \delta_{ij})/t \quad (5)$$

exist, and they are finite except perhaps when $i = j$ (Doob [7], Kolmogorov [25]). Always

$$\sum_{\alpha \neq i} q_{i\alpha} \leq -q_{ii},$$

and in many cases of practical interest

- (a) all q_{ij} are finite, and
- (b) $\sum_{\alpha} q_{i\alpha} = 0$ for each i .

Conversely, given a *conservative q-matrix*, i.e. a matrix $Q \equiv \{q_{ij}\}$ with non-negative elements off the main diagonal and satisfying both (a) and (b), then there exists *at least* one process for which $p'_{ij}(+0) = q_{ij}$, but in general this process is *not unique* (Feller [11], Doob [8]). When the process is unique the conservative q -matrix Q will be called *regular* and the associated process will be called the *Feller process* determined by Q . That this is the normal situation in practical problems is the sole justification for the common practice of specifying a Markov process merely by writing down a conservative q -matrix Q . Necessary and sufficient conditions for regularity have been given by Feller [11] and (in a somewhat different form) by Kato [20].

It will now be clear that for practical purposes the solution to Problem B will be insufficient, and that one must also consider

Problem C: Determine Π_F , the Π -matrix associated with the Feller process, when a regular conservative q -matrix Q is given.

A solution to this problem will be given in § 4: it will involve the non-negative solutions of

$$\sum_{\alpha} x_{\alpha} q_{\alpha j} = 0$$

such that $\sum x_{\alpha} < \infty$, and the non-negative solutions of

$$\sum_{\alpha} q_{i\alpha} y_{\alpha} = 0$$

such that $\sup_{\alpha} y_{\alpha} < \infty$.

1.4. Throughout this paper we shall confine ourselves to “honest” chains and processes, i.e. systems satisfying $\sum_{\alpha} p_{i\alpha} = 1$. However, a “dishonest” chain or process can always be imbedded in an honest one (obtained by adjoining a single “absorbing” state) and by means of this device our methods can easily be adapted to the general case.

We have also confined ourselves strictly to the problem of calculating the Π -matrix and it will be recalled that one cannot decide from an examination of the π_{ij} for a Markov chain whether or not a given dissipative state is *recurrent*. But this limitation also is only apparent; it has often been remarked that the question of recurrence can be settled by calculating the Π -matrix for a modified chain in which the given state is made absorbing.

1.5. In some respects this paper is a sequel to our paper [22] in which we considered the ergodic properties of one-parameter semigroups of operators on an ab-

stract Banach space, and readers of the earlier paper will find a discussion of the present problem from the standpoint of the general theory in § 6. Others who are mainly interested in probabilistic applications may prefer not to read beyond § 5, in which some examples are given to illustrate our methods. We have deliberately chosen the simplest examples which would serve this purpose, but we believe that the methods of this paper can usefully be applied to some more complicated systems which arise in practice. In particular, we intend to treat elsewhere the ergodic properties of two Markov processes which describe (i) the competition between two species, and (ii) the development of a stochastic epidemic.

2. Markov chains: the solution to Problem A

2.1. We first recall some results with which the reader will doubtless be familiar (perhaps in a different terminology).¹ The limits π_{ij} defined at (1) always exist and will clearly satisfy

$$\pi_{ij} \geq 0, \quad \sum_{\alpha} \pi_{i\alpha} \leq 1.$$

The structure of the matrix $\Pi \equiv \{\pi_{ij} : i, j \geq 0\}$ is most conveniently described by classifying the state j as *positive* when $\pi_{jj} > 0$ and as *dissipative* when $\pi_{jj} = 0$. The collection of positive states (if there are any) is then further divided into disjoint *positive classes*, the positive states j and k being in the same class if and only if $\pi_{jk} > 0$. (This relation can be shown to be reflexive, symmetric and transitive.) If we write $\pi_j \equiv \pi_{jj}$ for each positive state j , then the π_{ij} can be expressed in terms of the π_j and a set of numbers $\varpi(i, C)$, defined for each state i and each positive class C , where $0 \leq \varpi(i, C) \leq 1$. In fact

- (i) $\pi_{ij} = 0$ for all i , if j is dissipative.
- (ii) $\pi_{ij} = \varpi(i, C) \pi_j$ for all i , if j belongs to the positive class C .

Also π_j and $\varpi(i, C)$ have the following properties:

- (iii) $\sum_{j \in C} \pi_j = 1$ for each positive class C .
- (iv) If i is a positive state and C a positive class,

$$\varpi(i, C) = \begin{cases} 1 & \text{if } i \in C, \\ 0 & \text{if } i \notin C. \end{cases}$$

¹ For expository accounts of the theory, see CHUNG [4], FELLER ([12], Ch. 15), LOÈVE ([28], pp. 28-42).

(v) If i is dissipative and $\{C^{\varrho} : \varrho = 1, 2, \dots\}$ are the positive classes, then

$$\sum_{\varrho} \varpi(i, C^{\varrho}) \leq 1.$$

(vi) If C is a positive class, then

$$\sum_{\alpha \in C} \pi_{\alpha} p_{\alpha j} = \begin{cases} \pi_j & \text{if } j \in C, \\ 0 & \text{if } j \notin C, \end{cases} \quad (6)$$

and
$$\sum_{\alpha} p_{i\alpha} \varpi(\alpha, C) = \varpi(i, C) \text{ for all } i. \quad (7)$$

(vii) For all i and j ,

$$\sum_{\alpha} \pi_{i\alpha} \pi_{\alpha j} = \sum_{\alpha} \pi_{i\alpha} p_{\alpha j} = \sum_{\alpha} p_{i\alpha} \pi_{\alpha j} = \pi_{ij}. \quad (8)$$

(In (7) and (8) summations are over $\alpha = 0, 1, 2, \dots$; we shall adhere to this convention from now on.)

The classification of states and the description of π_{ij} have been given above in purely analytical terms: however, they have probabilistic meanings which (although they are not needed for what follows) the reader may usefully keep in mind. Positive states are precisely those which are *recurrent* with *finite* mean recurrence time μ_j , given by $\mu_j = 1/\pi_j$; dissipative states are either recurrent with infinite mean recurrence time, or non-recurrent (transient); finally $\varpi(i, C)$ is the probability that the system, starting at state i , will ultimately enter (and thereafter remain in) the positive class C . These interpretations depend on a detailed and deep analysis of the asymptotic properties of p_{ij}^n as $n \rightarrow \infty$, first made by Kolmogorov [24]; the properties of Π which we have stated lie less deep and have been proved more simply by Yosida & Kakutani [32] and Doob [7].

2.2. The calculation of Π when P is given involves two steps:

- (a) the classification of states and determination of the reciprocal mean recurrence times π_j ,
- (b) the calculation of the absorption probabilities $\varpi(i, C)$.

Step (b) will of course be superfluous (by (i) of 2.1) when there are no positive states, and will be trivial (by (iv)) when i is a positive state.

To perform step (a) we introduce the Banach space l whose elements are real sequences $x \equiv (x_0, x_1, x_2, \dots)$ such that

$$\|x\| \equiv \sum_{\alpha} |x_{\alpha}| < \infty.$$

Then P and Π determine bounded linear operators on l as follows:

$$(Px)_j \equiv \sum_{\alpha} x_{\alpha} p_{\alpha j}, \quad (\Pi x)_j \equiv \sum_{\alpha} x_{\alpha} \pi_{\alpha j}.$$

We have

$$Px \geq 0, \quad \Pi x \geq 0, \quad \|\Pi x\| \leq \|x\| = \|Px\|, \quad \text{whenever } x \geq 0,$$

and also (because of (8))

$$\Pi^2 = P\Pi = \Pi P = \Pi. \quad (9)$$

For each positive class C^e , we define $\pi^e \in l$ by

$$(\pi^e)_j = \begin{cases} \pi_j & \text{if } j \in C^e, \\ 0 & \text{if } j \notin C^e; \end{cases}$$

thus $\pi^e \geq 0$, $\|\pi^e\| = 1$ (by (iii) of 2.1), π^e has C^e as its *support*,¹ and from (6) we see that $P\pi^e = \pi^e$ so that π^e belongs to the nullspace $\mathcal{N}(\Delta)$ of the operator $\Delta \equiv P - I$. The following theorem is well known,² but we state it in a form which differs from the usual one and therefore sketch its proof.

THEOREM 1. *An l -vector x lies in $\mathcal{N}(\Delta)$ if and only if there exist real numbers $\{\lambda^e: e = 1, 2, \dots\}$ such that $\sum_e |\lambda^e| < \infty$ and $x = \sum_e \lambda^e \pi^e$. Also $x \geq 0$ if and only if $\lambda^e \geq 0$ for each e .*

Proof. From the fact that $\Delta\pi^e = 0$ it easily follows that $\Delta x = 0$ whenever x is of the stated form. On the other hand, if $x \in l$ and $\Delta x = 0$, then

$$\sum_{\alpha} x_{\alpha} \left(\frac{1}{n} \sum_{r=1}^n p_{\alpha j}^r \right) = x_j,$$

and so on letting $n \rightarrow \infty$ we obtain

$$\sum_{\alpha} x_{\alpha} \pi_{\alpha j} = x_j.$$

From (i) of 2.1 it now follows that $x_j = 0$ when j is dissipative, so that

$$\sum_e \sum_{\alpha \in C^e} x_{\alpha} (\pi^e)_j = x_j,$$

where the first summation is over all positive classes. The first assertion of the theorem now follows on setting

$$\lambda^e \equiv \sum_{\alpha \in C^e} x_{\alpha};$$

the second assertion follows from the fact that the vectors π^e are positive and have disjoint supports.

¹ The *support* of an l -vector x is the set of states j such that $x_j \neq 0$.

² See, e.g., CHUNG ([4], Th. 9), or LOÈVE ([28], p. 41).

We can deduce two important corollaries on using the fact that for each ϱ , the support of π^ϱ is precisely the corresponding positive class C^ϱ . Let us denote by $\mathcal{N}^+(\Delta)$ the set of *positive* vectors in $\mathcal{N}(\Delta)$. Then we have

COROLLARY 1.1. *A state j is positive if and only if it lies in the support of at least one l -vector $x \in \mathcal{N}^+(\Delta)$. Two positive states j and k lie in different positive classes if and only if there exists an $x \in \mathcal{N}^+(\Delta)$ whose support contains j but not k .*

COROLLARY 1.2. *Let j be a positive state. Then amongst the elements $x \in \mathcal{N}^+(\Delta)$ such that $x_j = 1$ there is a least, x^j , and $x^j / \|x^j\| = \pi^\varrho$ where C^ϱ is the positive class containing j .*

These corollaries show that step (a) can be performed *when the nullspace $\mathcal{N}(\Delta)$ has been found.*

2.3. We now have to carry out step (b) in so far as it is non-trivial, so that we must give a method for calculating $\varpi(k, C)$ when k is a dissipative state and C is a positive class. This will involve the nullspace $\mathcal{N}(\Delta^*)$ of the operator Δ^* adjoint to Δ .

Let m be the Banach space of real sequences $y \equiv (y_0, y_1, y_2, \dots)$ such that

$$\|y\| \equiv \sup_{\alpha} |y_{\alpha}| < \infty.$$

Then m is the adjoint space to l , and we shall write

$$(y, x) \equiv \sum_{\alpha} y_{\alpha} x_{\alpha} \quad \text{when } y \in m \text{ and } x \in l.$$

The operator P^* adjoint to P is given by

$$(y P^*)_i \equiv \sum_{\alpha} p_{i\alpha} y_{\alpha} \quad (y \in m),$$

and $(y P^*, x) = (y, P x)$ when $y \in m$ and $x \in l$. For each positive class C^ϱ , we define $\varpi^\varrho \in m$ by

$$(\varpi^\varrho)_i \equiv \varpi(i, C^\varrho) \quad (i = 0, 1, 2, \dots).$$

Properties (iii) and (iv) of 2.1 show that

$$\varpi^\varrho > 0, \quad \|\varpi^\varrho\| = 1, \quad (\varpi^\varrho, \pi^\varrho) = 1,$$

and $(\varpi^\varrho)_i = 0$ if $i \in C^\sigma$ ($\sigma \neq \varrho$). From (7) we have $\varpi^\varrho P^* = \varpi^\varrho$, so that $\varpi^\varrho \in \mathcal{N}^+(\Delta^*)$, the positive part of the nullspace $\mathcal{N}(\Delta^*)$ of the operator $\Delta^* \equiv P^* - I$. The nullspace of Δ^* is often much bigger than this result suggests,¹ but fortunately we can still

¹ The structure of $\mathcal{N}(\Delta^*)$ has been studied by BLACKWELL [2] and FELLER [13].

characterise ϖ^e by a minimal property similar to that involved in Corollary 1.2. We formulate this as

THEOREM 2. *Let C^e be a positive class and let π^e be the associated l -vector. Then ϖ^e is the least element y of $\mathcal{N}^+(\Delta^*)$ such that $(y, \pi^e) = 1$.*

Proof. We already know that $y \equiv \varpi^e$ satisfies the three conditions

$$y \geq 0, \quad y \Delta^* = 0, \quad (y, \pi^e) = 1.$$

Let y be any m -vector distinct from ϖ^e and satisfying these conditions. Then

$$\sum_{\alpha} p_{i\alpha}^n y_{\alpha} = y_i,$$

and so

$$\sum_{\alpha} \left(\frac{1}{n} \sum_{r=1}^n p_{i\alpha}^r \right) y_{\alpha} = y_i,$$

from which it follows (on using the positivity of y) that

$$\sum_{\alpha} \pi_{i\alpha} y_{\alpha} \leq y_i. \tag{10}$$

We can write (10) in the form

$$\sum_{\sigma} \varpi(i, C^{\sigma}) \sum_{\alpha \in C^{\sigma}} \pi_{\alpha} y_{\alpha} \leq y_i,$$

where the σ -summation is over all positive classes, and this is equivalent to

$$\sum_{\sigma \neq e} \varpi(i, C^{\sigma}) (y, \pi^{\sigma}) + \varpi(i, C^e) \leq y_i,$$

from which it follows that $\varpi^e < y$ as required.

We ought to mention here that Feller ([12], p. 332, (8.2) and (8.3)) has given a set of recurrence relations which uniquely determine the absorption probabilities $\varpi(i, C)$. This recurrent procedure may, however, be as difficult to carry out as the calculation of the π_{ij} directly from their definition (1). As Feller remarks, one has then to resort to his equation (8.4) and the solution to this equation is in general not unique. The non-trivial part of our Theorem 2 singles out the relevant solution.

2.4. We can now state our *solution to Problem A* as

THEOREM 3. *If the positive classes $\{C^e: e=1, 2, \dots\}$ and the associated l -vectors π^e and m -vectors ϖ^e are determined by Corollaries 1.1 and 1.2 and Theorem 2 above, then the matrix-elements of the "ergodic projection-operator" Π will be given by*

$$\pi_{ij} = 0 \quad (\text{all } i)$$

when j is not a positive state, and by

$$\pi_{ij} = (\varpi^e)_i (\pi^e)_j \quad (\text{all } i)$$

when j lies in the positive class C^e .

3. Markov processes: the solution to Problem B

3.1. If $\{p_{ij}(t) : i, j = 0, 1, 2, \dots; t \geq 0\}$ is the array of transition probabilities for a Markov process, and if (as we shall always assume) the continuity condition

$$p_{ij}(t) \rightarrow \delta_{ij} \quad \text{as } t \downarrow 0$$

holds, then the limits

$$\pi_{ij} \equiv \lim_{t \rightarrow \infty} p_{ij}(t) \quad (11)$$

exist. This result is due to Lévy [27], and can be proved by considering the chain defined by setting $p_{ij}^n \equiv p_{ij}(n\tau)$. This chain is aperiodic because $p_{ii}(n\tau) > 0$, and therefore $\lim_{n \rightarrow \infty} p_{ij}(n\tau)$ exists for each fixed $\tau > 0$. From the uniform continuity of $p_{ij}(\cdot)$ for fixed i and j , it follows that this limit is independent of τ and then that the limit at (11) exists.

Clearly we can calculate π_{ij} by applying the procedure of § 2 to the chain whose matrix of one-step transition probabilities is $P_\tau \equiv \{p_{ij}(\tau) : i, j = 0, 1, 2, \dots\}$, for any one $\tau > 0$, because

$$\pi_{ij} = \lim_{n \rightarrow \infty} p_{ij}(n\tau) = \lim_{n \rightarrow \infty} (P_\tau^n)_{ij}.$$

To do this, we must consider the nullspaces $\mathcal{N}(\Delta_\tau)$ and $\mathcal{N}(\Delta_\tau^*)$, where

$$\Delta_\tau \equiv P_\tau - I, \quad (P_\tau x)_j \equiv \sum_\alpha x_\alpha p_{\alpha j}(\tau) \quad (x \in l).$$

The dependence of this procedure on the choice of τ is, of course, only apparent, and this fact is exhibited most conveniently by introducing the one-parameter semigroup $\{P_t : t \geq 0\}$ of operators on l associated with the process. This has the properties:¹

- (a) $P_0 = I, \quad P_u P_v = P_{u+v} \quad (u \geq 0, v \geq 0).$
- (b) $P_t x \geq 0$ and $\|P_t x\| = \|x\|$ whenever $0 \leq x \in l$.
- (c) $\|P_t x - x\| \rightarrow 0$ as $t \downarrow 0$, for each $x \in l$.

Its infinitesimal generator Ω is defined by

$$\Omega x \equiv \text{strong } \lim_{t \downarrow 0} (P_t x - x)/t$$

¹ For the special properties of such "transition semigroups" on l , see KENDALL and REUTER [23] or KENDALL [21]; for the general theory of semigroups of operators, see HILLE [17].

whenever this limit exists, the domain $\mathcal{D}(\Omega)$ of Ω is dense in l , and

$$P_t x = x + \int_0^t P_u \Omega x \, du = x + \int_0^t \Omega P_u x \, du \quad (12)$$

for $x \in \mathcal{D}(\Omega)$ and $t \geq 0$. The adjoint operator Ω^* (on the adjoint space m of l) is defined by setting $y\Omega^* \equiv z$, where

$$(z, x) = (y, \Omega x) \quad \text{for all } x \in \mathcal{D}(\Omega),$$

whenever such an element z (necessarily unique) exists.

We now have the crucial

$$\text{LEMMA.} \quad \mathcal{N}(\Omega) = \mathcal{N}(\Delta_\tau) \quad \text{for each } \tau > 0; \quad (13)$$

$$\mathcal{N}(\Omega^*) = \bigcap_{\tau > 0} \mathcal{N}(\Delta_\tau^*). \quad (14)$$

Proof. If $x \in \mathcal{N}(\Omega)$, then (12) at once gives $P_\tau x = x$, so that $x \in \mathcal{N}(\Delta_\tau)$. Conversely if $x \in \mathcal{N}(\Delta_\tau)$ for one $\tau > 0$, then $P_\tau x = x$ and hence $\Pi x = x$ (as in the proof of Theorem 1). On using (9), we obtain that $P_t x = P_t \Pi x = \Pi x = x$ for all $t > 0$, and so finally that $\Omega x = 0$. This proves (13).

Next, if $y \in \mathcal{N}(\Omega^*)$, then (12) gives

$$\begin{aligned} (y, P_\tau x) &= (y, x) + \int_0^\tau (y, \Omega P_u x) \, du && \text{(all } x \in \mathcal{D}(\Omega)) \\ &= (y, x) + \int_0^\tau (y\Omega^*, P_u x) \, du = (y, x). \end{aligned}$$

Hence $(y P_\tau^*, x) = (y, x)$ for all $x \in \mathcal{D}(\Omega)$, and because $\mathcal{D}(\Omega)$ is dense this implies that $y P_\tau^* = y$, i.e. that $y \in \mathcal{N}(\Delta_\tau^*)$. Conversely if $y P_\tau^* = y$ for all $\tau > 0$, then $(y, P_\tau x) = (y, x)$ and

$$(y, \Omega x) = \lim_{\tau \downarrow 0} \left(y, \frac{P_\tau x - x}{\tau} \right) = 0 \quad \text{for all } x \in \mathcal{D}(\Omega),$$

so that $y \in \mathcal{N}(\Omega^*)$. This proves (14).

3.2. We now classify the states and describe the structure of Π . Let us say that the state j is *positive* if $\pi_{jj} > 0$, *dissipative* if $\pi_{jj} = 0$, and that two positive states j and k are in the same *positive class* if $\pi_{jk} > 0$ (this is an equivalence relation between positive states). Clearly this classification coincides with that of § 2 for each of the chains P_τ derived from the process. Using (13), we can at once transcribe Corollaries 1.1 and 1.2 in terms of $\mathcal{N}^+(\Omega)$, the positive part of $\mathcal{N}(\Omega)$, and obtain

THEOREM 4. A state is positive if and only if it lies in the support of some $x \in \mathcal{N}^+(\Omega)$; two positive states lie in different positive classes if and only if there exists an $x \in \mathcal{N}^+(\Omega)$ whose support contains one of the states but not the other.

If j is a positive state then amongst elements $x \in \mathcal{N}^+(\Omega)$ with $x_j = 1$ there is a least, x^j , and if $\pi^e \equiv x^j / \|x^j\|$ then π^e depends only on the class C^e containing j , and has C^e as its support.

To find the analogue of Theorem 2, we observe that the same l -vector π^e and m -vector ϖ^e are associated with a positive class C^e for each chain P_τ , and in each case ϖ^e is the least element y of $\mathcal{N}^+(\Delta_\tau^*)$ such that $(y, \pi^e) = 1$. Hence ϖ^e is also the least element y in

$$\bigcap_{\tau > 0} \mathcal{N}^+(\Delta_\tau^*)$$

or (by (14)) in $\mathcal{N}^+(\Omega^*)$, such that $(y, \pi^e) = 1$. Thus we have

THEOREM 5. If C^e is a positive class and π^e is the associated l -vector of Theorem 4, then amongst the elements y of $\mathcal{N}^+(\Omega^*)$ such that $(y, \pi^e) = 1$ there is a least, ϖ^e .

By combining Theorems 4 and 5 we obtain as a solution to Problem B:

THEOREM 6. If the positive classes C^e and the corresponding l -vectors π^e and m -vectors ϖ^e have been determined as in Theorems 4 and 5, then the matrix-elements of the ergodic projection-operator Π will be given by

$$\pi_{ij} = 0 \quad (\text{all } i)$$

when j is not a positive state, and by

$$\pi_{ij} = (\varpi^e)_i (\pi^e)_j \quad (\text{all } i)$$

when j lies in the positive class C^e .

It should be pointed out that whilst we have proved our results for processes by considering the chains P_τ and using the results of § 2, this is only a matter of convenience. Direct proofs can be given, by methods similar to those of § 2, based on the properties (analogous to (8))

$$\sum_{\alpha} \pi_{i\alpha} \pi_{\alpha j} = \sum_{\alpha} \pi_{i\alpha} p_{\alpha j}(t) = \sum_{\alpha} p_{i\alpha}(t) \pi_{\alpha j} = \pi_{ij} \quad (\text{for each } t \geq 0), \quad (15)$$

which were proved by Doob ([7], Theorems 6 and 7).

4. Feller processes: the solution to Problem C

4.1. Suppose that a matrix Q of finite real elements q_{ij} is given, such that

$$q_{ij} \geq 0 \quad \text{when } i \neq j, \quad \sum_{\alpha} q_{i\alpha} = 0 \quad \text{for all } i.$$

Suppose further that the conservative matrix Q is *regular*; i.e. that the “minimal” process constructed by Feller [11] is *honest* (satisfies $\sum_{\alpha} p_{i\alpha}(t) = 1$), or equivalently that there is *exactly one* Markov process such that $p'_{ij}(+0) = q_{ij}$. (See Doob [8], Reuter [31]).

We write $x \in \mathcal{D}(Q)$ whenever

- (a) $x \in l$,
- (b) $\sum_{\alpha} x_{\alpha} q_{\alpha j}$ is absolutely convergent for each j ,
- (c) $\sum_j \left| \sum_{\alpha} x_{\alpha} q_{\alpha j} \right| < \infty$,

and we define an operator Q with domain $\mathcal{D}(Q)$ by setting

$$(Qx)_j \equiv \sum_{\alpha} x_{\alpha} q_{\alpha j} \quad (x \in \mathcal{D}(Q)). \quad (16)$$

The set \mathcal{D}_0 of “finite” vectors (those with only finitely many non-zero components) is contained in $\mathcal{D}(Q)$, and we define Q_0 to be the restriction of Q to $\mathcal{D}(Q_0) \equiv \mathcal{D}_0$. Because \mathcal{D}_0 is dense in l we can define the adjoint Q_0^* of Q_0 , and this can be shown to be given by

$$(yQ_0^*)_i \equiv \sum_{\alpha} q_{i\alpha} y_{\alpha} \quad (y \in \mathcal{D}(Q_0^*)) \quad (17)$$

the domain $\mathcal{D}(Q_0^*)$ of Q_0^* consisting of those vectors $y \in m$ for which (17) defines an element yQ_0^* of m .

Now let Ω_F generate the Feller semigroup associated with the given q 's. Then¹

$$Q_0 \subseteq \Omega_F \subseteq Q, \quad (18)$$

and hence

$$Q^* \subseteq \Omega_F^* \subseteq Q_0^*; \quad (19)$$

further for each $\lambda > 0$ and $x \in l$, the equation

$$\lambda \xi - \Omega_F \xi = x \quad (20)$$

has exactly one solution $\xi \equiv \Phi_{\lambda} x$ in $\mathcal{D}(\Omega_F)$. When $x \geq 0$, then $\Phi_{\lambda} x \geq 0$ and $\xi \equiv \Phi_{\lambda} x$ is the least positive solution (in $\mathcal{D}(Q)$) of the equation

$$\lambda \xi - Q \xi = x. \quad (21)$$

¹ The properties of Ω_F which we state here can be proved without undue difficulty by examining FELLER's construction [11] of his “minimal” semigroup; see [31].

We now make use of our assumption that Q is regular, i.e. that Ω_F generates a *transition* semigroup. This implies that $\lambda\Phi_\lambda$ is a transition operator, and¹ that the nullspace $\mathcal{N}(\lambda I - Q_0^*)$ contains only the zero vector, for each $\lambda > 0$. Using these facts, we are able to eliminate the explicit references to Ω in Theorems 4 and 5, and replace them by statements involving the q 's alone. To do this, we prove

THEOREM 7. *If Q is regular, and Ω_F generates the (unique) associated transition semigroup, then*

$$\mathcal{N}^+(\Omega_F) = \mathcal{N}^+(Q), \quad (22)$$

$$\mathcal{N}^+(\Omega_F^*) = \mathcal{N}^+(Q_0^*). \quad (23)$$

Proof. Clearly (18) implies that $\mathcal{N}^+(\Omega_F) \subseteq \mathcal{N}^+(Q)$, and (22) will follow if we can prove the reverse inclusion. Suppose then that $x \geq 0$, $x \in \mathcal{D}(Q)$, and $Qx = 0$, and choose some fixed $\lambda > 0$. We shall then have $\lambda x - Qx = \lambda x \geq 0$, so that $\xi = x$ is a positive solution of the equation

$$\lambda\xi - Q\xi = \lambda x \geq 0 \quad (\xi \in \mathcal{D}(Q)). \quad (24)$$

But (cf. (21)) the least positive solution of (24) is $\xi = \Phi_\lambda(\lambda x)$, so that $x \geq \Phi_\lambda(\lambda x) = \lambda\Phi_\lambda x$; also we cannot have $x = \lambda\Phi_\lambda x$ because this would now give $\|x\| > \|\lambda\Phi_\lambda x\|$, contradicting the fact that $\lambda\Phi_\lambda$ is a transition operator. It follows that $x = \Phi_\lambda(\lambda x) \in \mathcal{D}(\Omega_F)$, so that $x \in \mathcal{N}^+(\Omega_F)$; this proves that $\mathcal{N}^+(Q) \subseteq \mathcal{N}^+(\Omega_F)$, and (22) follows.

We shall deduce (23) from the sharper statement that

$$\Omega_F^* = Q_0^*, \quad (25)$$

and to prove this it will suffice, by (19), to prove that

$$\mathcal{D}(\Omega_F^*) = \mathcal{D}(Q_0^*). \quad (26)$$

We shall need here one further fact from semigroup theory, namely that when $\lambda > 0$ and $x \in \mathcal{D}(\Omega_F)$, then

$$\Phi_\lambda(\lambda I - \Omega_F)x = x.$$

From this it follows at once that

$$w\Phi_\lambda^* \in \mathcal{D}(\Omega_F^*)$$

and that

$$w\Phi_\lambda^*(\lambda I - \Omega_F^*) = w$$

for all $w \in m$.

¹ The condition, $\mathcal{N}(\lambda I - Q_0^*) = \{0\}$, is given in KATO's paper [20]. The condition can be proved to be equivalent to the regularity of Q ; see [31].

Now if (26) were false, we could find $y \neq 0$ such that $y \in \mathcal{D}(Q_0^*)$ but $y \notin \mathcal{D}(\Omega_F^*)$. Fix $\lambda > 0$, and define

$$z \equiv y(\lambda I - Q_0^*) \Phi_\lambda^*.$$

Then $z \in \mathcal{D}(\Omega_F^*)$, so that $z \neq y$, but also

$$\begin{aligned} z(\lambda I - Q_0^*) &= z(\lambda I - \Omega_F^*) = y(\lambda I - Q_0^*) \Phi_\lambda^* (\lambda I - \Omega_F^*) \\ &= y(\lambda I - Q_0^*), \end{aligned}$$

so that $(z - y)$ would be a non-zero element of $\mathcal{N}(\lambda I - Q_0^*)$, contradicting the regularity of Q . Thus (26) must hold, and (23) follows. (For (26), see also [31], § 7.2.)

4.2. To obtain a solution to Problem C, we have merely to combine Theorem 7 with Theorems 4 to 6. This leads us to

THEOREM 8. Let $Q \equiv \{q_{ij} : i, j = 0, 1, 2, \dots\}$ be a regular conservative q -matrix, and let $\{p_{ij}(t) : t \geq 0\}$ be the unique process such that $p'_{ij}(+0) = q_{ij}$. Then:

- (i) A state is positive if and only if it lies in the support of some $x \in \mathcal{N}^+(Q)$; two positive states lie in different positive classes if and only if there exists an $x \in \mathcal{N}^+(Q)$ whose support contains one of the states but not the other.
- (ii) If j is a positive state, then the set of l -vectors $x \in \mathcal{N}^+(Q)$ with $x_j = 1$ has a least member x^j , and if $\pi^e \equiv x^j / \|x^j\|$ then π^e depends only on the positive class C^e containing j and has C^e as its support.
- (iii) If C^e is a positive class, then the set of m -vectors $y \in \mathcal{N}^+(Q_0^*)$ with $(y, \pi^e) = 1$ has a least member ϖ^e .

(iv) The limits
$$\pi_{ij} \equiv \lim_{t \rightarrow \infty} p_{ij}(t)$$

are given by
$$\pi_{ij} = 0 \quad (\text{all } i)$$

when j is not a positive state, and by

$$\pi_{ij} = (\varpi^e)_i (\pi^e)_j \quad (\text{all } i)$$

when j lies in the positive class C^e .

In order to apply Theorem 8, we have only to use the facts that $\mathcal{N}^+(Q)$ consists of all positive vectors $x \in l$ such that

$$\sum_{\alpha=0}^{\infty} x_\alpha q_{\alpha j} = 0 \quad (j = 0, 1, 2, \dots), \tag{27}$$

and $\mathcal{N}^+(Q_0^*)$ consists of all positive vectors $y \in m$ such that

$$\sum_{\alpha=0}^{\infty} q_{i\alpha} y_{\alpha} = 0 \quad (i = 0, 1, 2, \dots). \quad (28)$$

The relevance to the ergodic problem of the positive convergent solutions to (27) and of the positive bounded solutions to (28) has, of course, long been recognised by statisticians; the importance of Theorem 8 is that it allows the appropriate solutions to be identified in cases of non-uniqueness.

5. Examples

5.1. The random walk. We begin with an example illustrating the solution to Problem A; it is a familiar one and here the ergodic behaviour is well understood (cf. Foster [14], Harris [16], Jensen [18], Karlin & McGregor [19]). We label the states of a Markov chain as $\dots, -1, 0, 1, \dots$ and then put

$$p_{ij} \equiv \begin{cases} p_i & \text{if } j = i + 1, \\ q_i & \text{if } j = i - 1, \\ 0 & \text{if } j \neq i \pm 1, \end{cases}$$

where $p_i > 0$, $q_i > 0$ and $p_i + q_i = 1$, for all i . To find the nullspace of Δ we must solve the difference equations

$$x_j = p_{j-1} x_{j-1} + q_{j+1} x_{j+1} \quad (\text{all } j),$$

and we write these as

$$u_j \equiv p_j x_j - p_{j-1} x_{j-1} = q_{j+1} x_{j+1} - q_j x_j.$$

Evidently¹ $u \in l$ if $x \in l$, so that we must have

$$\begin{aligned} p_j x_j &= u_j + u_{j-1} + u_{j-2} + \dots, \\ q_j x_j &= -(u_j + u_{j+1} + u_{j+2} + \dots), \end{aligned} \quad (29)$$

and we therefore put

$$v_j \equiv \sum_{-\infty}^j u_{\alpha}, \quad \sigma \equiv \sum_{-\infty}^{\infty} u_{\alpha},$$

so that (29) becomes

$$p_j x_j = v_j, \quad q_j x_j = v_{j-1} - \sigma.$$

Because both x and u are in l , we know that x_j and v_j tend to zero when $j \rightarrow -\infty$; thus $\sigma = 0$. The last pair of equations now gives $v_j = (p_j/q_j) v_{j-1}$, and hence

¹ A typical l -vector now takes the form $x \equiv (\dots, x_{-1}, x_0, x_1, \dots)$, with $\|x\| \equiv \sum_{-\infty}^{\infty} |x_{\alpha}| < \infty$.

$$v_j = \begin{cases} \frac{p_1 p_2 \dots p_j}{q_1 q_2 \dots q_j} v_0 & (j > 0), \\ \frac{q_{j+1} q_{j+2} \dots q_0}{p_{j+1} p_{j+2} \dots p_0} v_0 & (j < 0). \end{cases}$$

Let us set

$$R \equiv \sum_{-\infty}^{-1} \frac{1}{p_j} \frac{q_{j+1} \dots q_0}{p_{j+1} \dots p_0} + \frac{1}{p_0} + \sum_1^{\infty} \frac{1}{p_j} \frac{p_1 \dots p_j}{q_1 \dots q_j}. \quad (30)$$

If $R = \infty$, then $v_0 = 0$, the nullspace of Δ contains the zero vector only, and every state is dissipative. If $R < \infty$, then the nullspace of Δ is spanned by a single positive vector, every state is positive, and the states form a single positive class. In both cases, we shall have

$$\pi_{ij} = \begin{cases} \frac{1}{p_j} \frac{q_{j+1} \dots q_0}{p_{j+1} \dots p_0} \frac{1}{R} & \text{if } j < 0, \\ 1/(p_0 R) & \text{if } j = 0, \\ \frac{1}{p_j} \frac{p_1 \dots p_j}{q_1 \dots q_j} \frac{1}{R} & \text{if } j > 0, \end{cases} \quad (31)$$

for all i (these expressions all being zero when $R = \infty$).

The nullspace of Δ^* can readily be found for the present example, but (cf. 2.2) this is never required when, as here, $\mathcal{N}(\Delta)$ is *either* zero-dimensional *or* is one-dimensional with a strictly positive generating vector.

5.2. The "flash of flashes". We next give two examples illustrating the solution to problem B, and we have chosen for this purpose the two most pathological processes which we know. The first of these has a conservative q -matrix which is so highly non-regular as to be associated with a continuum of processes all satisfying the same set of "backward" and "forward" differential equations. Problem C would here be quite meaningless and the process will be specified by giving its infinitesimal generator Ω .

The space l will now be so labelled that a typical l -vector becomes

$$x \equiv (\dots, x^{-1}, x^0, x^1, \dots)$$

where

$$x^s \equiv (\dots, x_{-1}^s, x_0^s, x_1^s, \dots)$$

and the x_n^s are real numbers such that

$$\|x\| \equiv \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} |x_{\alpha}^{\sigma}| < \infty.$$

Now let a_n^s ($s, n = \dots, -1, 0, 1, \dots$) be positive real numbers such that $\sum \sum 1/a_\alpha^\alpha < \infty$, and define the conservative q -matrix Q by

$$q_{mn}^{rs} = 0 \quad (r \neq s); \quad q_{mn}^{ss} = \begin{cases} -a_m^s & \text{if } n = m, \\ +a_m^s & \text{if } n = m + 1, \\ 0 & \text{otherwise.} \end{cases}$$

The operator Ω is then defined to be the restriction of the matrix-operator Q (defined as at (16)) to the domain of l -vectors x such that

- (i) $\sum \sum |a_{\alpha-1}^\alpha x_{\alpha-1}^\alpha - a_\alpha^\alpha x_\alpha^\alpha| < \infty$,
- (ii) $U^s x = L^{s+1} x \quad (s = \dots, -1, 0, 1, \dots)$,
- (iii) $\lim_{\sigma \rightarrow +\infty} U^\sigma x = \lim_{\sigma \rightarrow -\infty} L^\sigma x$,

where

$$U^s x \equiv \lim_{n \rightarrow +\infty} a_n^s x_n^s, \quad L^s x \equiv \lim_{n \rightarrow -\infty} a_n^s x_n^s.$$

(It is shown in Kendall [21] that Ω generates a transition semigroup.)

We must now find the nullspace of Ω . If $x \in \mathcal{N}(\Omega)$, then

$$a_{n-1}^s x_{n-1}^s - a_n^s x_n^s = 0$$

for all s and n , and so

$$x_n^s = c^s / a_n^s, \quad \text{where } c^s = U^s x = L^s x.$$

Condition (ii) then shows that c^s is independent of s , and the other conditions are now satisfied automatically. Thus $\mathcal{N}(\Omega)$ is one-dimensional and is spanned by the vector whose $\binom{s}{n}$ th component is $1/a_n^s$. It follows that all the states are positive, that they form a single positive class, and that

$$\pi_{ij} \equiv \pi_{mn}^{rs} = \frac{1}{a_n^s} \bigg/ \sum \sum \frac{1}{a_\alpha^\alpha} \tag{32}$$

for all values of r, s, m and n . As in 5.1 there is no need to calculate $\mathcal{N}(\Omega^*)$.

5.3. A sequence of flashes communicating via an instantaneous state. In this example (also taken from Kendall [21]) the Markov process has a highly non-conservative q -matrix; one of states, labelled 0, is "instantaneous" (has $q_{00} = -\infty$) and the corresponding row of the q -matrix can be written as

$$(-\infty, 0, 0, 0, \dots).$$

As in 5.2 the process will be specified by giving its infinitesimal generator Ω . This time l is so labelled that a typical l -vector becomes

$$x \equiv (x^0, x^1, x^2, \dots)$$

where x^0 is a real number and x^s is a real vector

$$x^s \equiv (\dots, x_{-1}^s, x_0^s, x_1^s, \dots).$$

The norm is defined by

$$\|x\| \equiv |x^0| + \sum_{\sigma=1}^{\infty} \sum_{\alpha=-\infty}^{\infty} |x_{\alpha}^{\sigma}| < \infty.$$

Now choose positive numbers a_n^s ($s = 1, 2, 3, \dots; n = \dots, -1, 0, 1, \dots$) such that $\sum \sum 1/a_{\alpha}^{\sigma} < \infty$ and define $\mathcal{D}(\Omega)$ to be the set of vectors x which satisfy

$$(i) \quad \sum \sum |a_{\alpha-1}^{\sigma} x_{\alpha-1}^{\sigma} - a_{\alpha}^{\sigma} x_{\alpha}^{\sigma}| < \infty$$

and

$$(ii) \quad L^s x \equiv \lim_{n \rightarrow -\infty} a_n^s x_n^s = x^0 \quad (\text{all } s \geq 1)$$

(such vectors necessarily belong to l). Finally define Ωx for $x \in \mathcal{D}(\Omega)$ by

$$(\Omega x)^0 \equiv \sum_{\sigma=1}^{\infty} (U^{\sigma} x - x^0),$$

$$(\Omega x)_n^s \equiv a_{n-1}^s x_{n-1}^s - a_n^s x_n^s \quad (s \geq 1, \text{ all } n),$$

where

$$U^s x \equiv \lim_{n \rightarrow +\infty} a_n^s x_n^s$$

(this last limit necessarily exists and the series equated to $(\Omega x)^0$ is necessarily absolutely convergent). Evidently $\Omega x = 0$ if and only if $a_n^s x_n^s = c = x^0$, and so $\mathcal{N}(\Omega)$ is one-dimensional and is spanned by the vector whose components are

$$x^0 = 1, \quad x_n^s = 1/a_n^s.$$

Thus all the states are positive and form a single positive class, and

$$\left. \begin{aligned} \pi^{00} &= A, & \pi_{.n}^{0s} &= A/a_n^s, \\ \pi_m^r &= A, & \pi_{mn}^s &= A/a_n^s, \end{aligned} \right\} \quad (33)$$

where $1/A \equiv 1 + \sum \sum 1/a_{\alpha}^{\sigma}$. Once again there is no need to find $\mathcal{N}(\Omega^*)$.

5.4. A general Markovian queuing process. We now turn to Problem C: the calculation of the ergodic projection for Feller processes. As a first example we consider the Markov process having the q -matrix

$$\left| \begin{array}{cccccc} -b_0 & b_0 & 0 & 0 & \dots \\ a_1 & -(a_1 + b_1) & b_1 & 0 & \dots \\ 0 & a_2 & -(a_2 + b_2) & b_2 & \dots \\ 0 & 0 & a_3 & -(a_3 + b_3) & \dots \\ \dots & \dots & \dots & \dots & \dots \end{array} \right| \quad (34)$$

where $b_0, a_1, b_1, a_2, \dots$ are given *positive* real numbers.¹ In the classical queuing process of Erlang we have $a_1 = a_2 = \dots = a$ and $b_0 = b_1 = b_2 = \dots = b$; the state-label r is equal to the number of persons waiting or being served and so a can be identified as the reciprocal of the mean service time and b as the reciprocal of the mean time between arrivals, each of these times having a negative-exponential distribution. In the more general system considered here each of a and b is allowed to depend on the number r of persons present in the queue. A similar extension of the classical queuing situation has been considered previously by Jensen [18].

The q -matrix at (34) is conservative: it will be regular and so define a unique Markov process if and only if the set of equations

$$\begin{aligned} (\lambda + b_0) y_0 &= b_0 y_1, \\ (\lambda + a_r + b_r) y_r &= a_r y_{r-1} + b_r y_{r+1} \quad (r \geq 1), \end{aligned}$$

has no bounded solution other than $y_0 = y_1 = y_2 = \dots = 0$ for any one (and then for all) $\lambda > 0$. In looking for a non-null solution we may clearly assume that $y_0 = 1$ and then the relations

$$\left. \begin{aligned} y_1 &= 1 + \lambda/b_0, \\ b_r (y_{r+1} - y_r) &= \lambda y_r + a_r (y_r - y_{r-1}), \quad (r \geq 1) \end{aligned} \right\} \quad (35)$$

show inductively that $1 = y_0 < y_1 < y_2 < \dots$. We therefore put $Y \equiv \lim y_r \leq \infty$. From (35) we find that

$$y_{r+1} - y_r = \lambda \left\{ \frac{y_r}{b_r} + \frac{a_r y_{r-1}}{b_r b_{r-1}} + \dots + \frac{a_r \dots a_1 y_0}{b_r \dots b_1 b_0} \right\} \quad (36)$$

for all $r \geq 0$. Now if

$$c_r \equiv \frac{1}{b_r} + \frac{a_r}{b_r b_{r-1}} + \dots + \frac{a_r \dots a_1}{b_r \dots b_1 b_0} \quad (r \geq 0),$$

then (36) gives

$$\lambda c_r \leq y_{r+1} - y_r \leq \lambda c_r y_r \quad (r \geq 0),$$

and so

$$1 + \lambda \sum_0^{\infty} c_r \leq Y \leq \prod_0^{\infty} (1 + \lambda c_r).$$

We conclude that *the q -matrix (34) is regular if and only if*

$$\sum_{r=0}^{\infty} \left\{ \frac{1}{b_r} + \frac{a_r}{b_r b_{r-1}} + \dots + \frac{a_r \dots a_1}{b_r \dots b_1 b_0} \right\} = \infty. \quad (37)$$

This result is due to Dobrušin [6], who deduced it from Feller's regularity criterion; the advantages of the alternative analytical criterion for regularity are very well shown by this example.

¹ We now revert to 0, 1, 2, ... as the state-labels.

We now suppose that (37) is satisfied, so that (34) is associated with a unique Markov process. The solution to Problem C requires that we find all positive l -vectors x such that

$$-b_0 x_0 + a_1 x_1 = 0, \quad (38)$$

$$b_{r-1} x_{r-1} - (a_r + b_r) x_r + a_{r+1} x_{r+1} = 0 \quad (r \geq 1). \quad (39)$$

These equations have a one-dimensional set of solutions generated by

$$x_0 = 1, \quad x_1 = \frac{b_0}{a_1}, \quad x_2 = \frac{b_0 b_1}{a_1 a_2}, \quad \dots, \quad x_r = \frac{b_0 \dots b_{r-1}}{a_1 \dots a_r}, \quad \dots,$$

so that we must distinguish between two cases, depending on the nature of

$$S \equiv 1 + \sum_{r=1}^{\infty} \frac{b_0 b_1 \dots b_{r-1}}{a_1 a_2 \dots a_r}. \quad (40)$$

If $S = \infty$, then every state is dissipative, whilst if $S < \infty$ then all the states are positive and form a single positive class. In both cases

$$\pi_{i0} = \frac{1}{S}, \quad \pi_{ij} = \frac{b_0 \dots b_{j-1}}{a_1 \dots a_j} \frac{1}{S} \quad (j \geq 1), \quad (41)$$

for all i .

It should be noted that $S = \infty$ and $S < \infty$ are each consistent with the regularity condition (37). If $a_r = a$ and $b_r = b$ for all values of r then (37) will hold and we find that $S < \infty$ if and only if $b < a$ (a well-known result). The importance of the series (40) in queuing problems has been previously noted by Jensen [18].

5.5. The general birth-and-death process. We now consider the Markov process associated with the conservative q -matrix (34) when $b_0 = 0$ and $a_r > 0$, $b_r > 0$, for $r \geq 1$. Naturally we must again require the q -matrix to be regular, and the regularity condition will be derived in a moment. The state-label r will now be identified with the number of individuals in a population for which the chances of a birth or death occurring in the short time interval δt are

$$b_r \delta t + o(\delta t) \quad \text{and} \quad a_r \delta t + o(\delta t),$$

respectively. The condition $b_0 = 0$ ensures that the population cannot recover if it once becomes extinct. The familiar "simple" birth-and-death process corresponds to the choice

$$a_r \equiv r a, \quad b_r \equiv r b$$

for the parameters.

The q -matrix will be regular if and only if, for some one (and then for all) $\lambda > 0$, the set of equations

$$\lambda y_0 = 0,$$

$$(\lambda + a_r + b_r) y_r = a_r y_{r-1} + b_r y_{r+1} \quad (r \geq 1),$$

has no non-null bounded solution. Without loss of generality we may put $y_0 = 0$ and $y_1 = 1$, and then we shall have

$$y_2 = 1 + (\lambda + a_1)/b_1,$$

$$b_r (y_{r+1} - y_r) = \lambda y_r + a_r (y_r - y_{r-1}) \quad (r \geq 2),$$

so that $1 = y_1 < y_2 < \dots$. Again it is convenient to put $Y \equiv \lim_{r \rightarrow \infty} y_r$, and very much as before we find that

$$\lambda d_r + e_r \leq y_{r+1} - y_r \leq (\lambda d_r + e_r) y_r \quad (r \geq 1),$$

where

$$d_r \equiv \frac{1}{b_r} + \frac{a_r}{b_r b_{r-1}} + \dots + \frac{a_r \dots a_2}{b_r \dots b_2 b_1}, \quad e_r \equiv \frac{a_r \dots a_2 a_1}{b_r \dots b_2 b_1},$$

and so

$$1 + \sum_1^{\infty} (e_r + \lambda d_r) \leq Y \leq \prod_1^{\infty} (1 + e_r + \lambda d_r).$$

It follows that the q -matrix (34), with $b_0 = 0$, is regular if and only if

$$\sum_{r=1}^{\infty} \left\{ \frac{1}{b_r} + \frac{a_r}{b_r b_{r-1}} + \dots + \frac{a_r \dots a_2}{b_r \dots b_2 b_1} \right\} = \infty. \quad (42)$$

This result also is to be attributed to Dobrušin [6].¹ Note that (37) is in effect a condition on b_1, a_2, b_2, \dots only, and that it is equivalent to (42), so that the regularity of the system does not depend on whether the state $r=0$ is or is not absorbing.

We now assume that (42) holds and find the solution to Problem C. First we must find all positive l -vectors x such that

$$a_1 x_1 = 0,$$

$$-(a_1 + b_1) x_1 + a_2 x_2 = 0,$$

$$b_{r-1} x_{r-1} - (a_r + b_r) x_r + a_{r+1} x_{r+1} = 0 \quad (r \geq 2),$$

and we see at once that the general solution is an arbitrary non-negative multiple of the vector

$$u^0 \equiv [1, 0, 0, 0, \dots].$$

Thus the zero-state is positive and forms by itself the only positive class C ; all other states are dissipative. Accordingly we shall have

¹ See also KARLIN & MCGREGOR [19].

$$\pi_{ij} = 0 \quad (j \geq 1, \text{ all } i)$$

and

$$\pi_{i0} = \varpi(i, C) \quad (\text{all } i),$$

so that the missing column of the Π -matrix consists of the extinction-probabilities

$$\varpi(i, C) \equiv \lim_{t \rightarrow +\infty} p_{i0}(t) \quad (i = 0, 1, 2, \dots);$$

it is clear that this limit is attained monotonically, and that it could be interpreted as the chance that the population, initially of size i , will ultimately become extinct.

Our solution to Problem C shows that the vector ϖ whose i th component is $\varpi(i, C)$ is the least non-negative bounded solution y to the equations

$$a_r y_{r-1} - (a_r + b_r) y_r + b_r y_{r+1} = 0 \quad (r = 1, 2, 3, \dots), \quad (43)$$

subject to the requirement that

$$(y, u^0) \equiv y_0 = 1.$$

The general solution to the difference equation (43) is

$$y_0 = A, \quad y_r = A + B \left(1 + \sum_{s=1}^{r-1} \frac{a_1 \dots a_s}{b_1 \dots b_s} \right) \quad (r \geq 1),$$

where A and B are arbitrary constants, and the nature of its bounded solutions will depend on

$$T \equiv 1 + \sum_{s=1}^{\infty} \frac{a_1 \dots a_s}{b_1 \dots b_s}. \quad (44)$$

If $T = \infty$, we must have $\varpi \equiv [1, 1, 1, \dots]$

and extinction is almost certain whatever the initial state. If however $T < \infty$, then $\mathcal{N}(Q_0^*)$ will be two-dimensional and the minimality condition comes into play; we find then that the vector ϖ and the individual extinction probabilities are given by

$$\left. \begin{aligned} \varpi_0 &= \pi_{00} = 1, \\ \varpi_i &= \pi_{i0} = \frac{1}{T} \sum_{s=i}^{\infty} \frac{a_1 \dots a_s}{b_1 \dots b_s} \quad (i \geq 1). \end{aligned} \right\} \quad (45)$$

As in 5.4 it should be noted that $T = \infty$ and $T < \infty$ are both consistent with the regularity condition (42): for instance if $a_r = ra$ and $b_r = rb$ (the "simple" birth-and-death process) then (42) always holds, whilst $T < \infty$ if and only if $a < b$. It is also worth observing that the conditions $S = \infty$ (in 5.4) and $T = \infty$ (in 5.5) are entirely different in character. Roughly speaking a queuing process will be completely dissipative ($S = \infty$) when the ratios b_r/a_r are too large, whilst the corresponding birth-and-death process will have the universal extinction property ($T = \infty$) when these

ratios are too small. This is of course quite reasonable from the intuitive standpoint. Finally it should also be mentioned that the Π -matrix for the examples in 5.4 and 5.5 has been previously found by Ledermann and Reuter [26] assuming the q -matrix to be regular (the regularity conditions (37) and (42) were not known to them); the present method is however much simpler.

6. The calculation of $\mathcal{R}(\Pi)$ and $\mathcal{N}(\Pi)$

6.1. A different approach to Problem B. In [22] we treated a generalisation of Problem B: to determine the subspace Γ of "ergodic" vectors $x \in X$ and the associated "ergodic projection operator" Π (defined on Γ), for a one-parameter semigroup $\{P_t : t \geq 0\}$ of operators on a Banach space X . A vector x was there called ergodic whenever $\Pi x \equiv \lim_{t \rightarrow \infty} P_t x$ existed as some kind of generalised limit and with regard to some suitable topology for X . This work formed a natural continuation of earlier investigations in general ergodic theory by Dunford [10], Hille ([17], Ch. XIV) and Phillips [30], but for us it was also prompted by a desire to solve Problem B of the present paper. If we set $X = l$ in [22], the resulting theorems are not always applicable to Problem B, but the methods of [22] can be adapted to yield a partial solution. This has an entirely different character from the solution given in § 3, where the lattice properties of l were exploited: we shall now have to study the interplay of a variety of weak topologies. It should be emphasised that the solution of § 3 is superior in that it is always available; the methods to be described now may sometimes fail altogether but have their own merits whenever they can be applied. Problem A can also be treated by similar methods; the analysis is then simpler and we leave the details to the reader.

As in § 3, we consider a Markov process whose array of transition probabilities is $\{p_{ij}(t) : i, j = 0, 1, 2, \dots\}$. Let Ω be the infinitesimal generator of the associated transition semigroup $\{P_t : t \geq 0\}$ on l ; also, with $\pi_{ij} \equiv \lim_{t \rightarrow \infty} p_{ij}(t)$ as before, define the ergodic projection operator Π by

$$(\Pi x)_j \equiv \sum_{\alpha} x_{\alpha} \pi_{\alpha j} \quad (x \in l).$$

We recall (cf. (9) and (15)) that

$$\Pi^2 = P_t \Pi = \Pi P_t = \Pi \quad \text{for all } t \geq 0. \quad (46)$$

It is important to note that the operator Π is in general *not* the same as the operator Π of [22].

Because Π is idempotent we can write any $x \in l$ uniquely in the form $x = x_1 + x_2$, where $x_1 \in \mathcal{R}(\Pi)$ and $x_2 \in \mathcal{N}(\Pi)$, and $x_1 = \Pi x$. Thus we have the direct decomposition

$$l = \mathcal{R}(\Pi) \oplus \mathcal{N}(\Pi). \quad (47)$$

If we can find the summands $\mathcal{R}(\Pi)$ and $\mathcal{N}(\Pi)$ in (47), then we can find Πx , for any $x \in l$, by taking the component of x in $\mathcal{R}(\Pi)$. It will now be seen that we may rephrase Problem B as follows.

Problem B₁: *Determine $\mathcal{R}(\Pi)$ and $\mathcal{N}(\Pi)$ when Ω is given.*

We shall see that $\mathcal{R}(\Pi)$ is easy to identify, and coincides with $\mathcal{N}(\Omega)$. The determination of $\mathcal{N}(\Pi)$ is more troublesome; it turns out that $\mathcal{R}(\Omega) \subseteq \mathcal{N}(\Pi)$, and we shall try to reach $\mathcal{N}(\Pi)$ by closing $\mathcal{R}(\Omega)$ with regard to several weak topologies for l .

6.2. Weak topologies. It will be convenient for future reference to state some results which we shall need concerning weak topologies for a Banach space X . (Cf. [5] and [3], Ch. IV.)

Let G be a linear set of functionals $g \in X^*$ which is "total" for X , i.e. such that the vanishing of (g, x) for all $g \in G$ implies $x = 0$. Then the G -weak topology for X is the (Hausdorff) topology generated by the sub-basic open sets

$$\{x : x \in X, \alpha < (g, x) < \beta\} \quad (\alpha < \beta),$$

where α, β range over the real numbers and g ranges over G . If $E \subseteq X$ then we shall write \bar{E} for the strong and $[E]_G$ for the G -weak closures of E . Also if $A \subseteq X$ and $B \subseteq X^*$ we shall write A^\perp and B^\top for the annihilators of A in X^* and B in X :

$$A^\perp \equiv \{y : y \in X^* \text{ and } (y, x) = 0 \text{ for all } x \in A\}, \quad (48)$$

$$B^\top \equiv \{x : x \in X \text{ and } (y, x) = 0 \text{ for all } y \in B\}. \quad (49)$$

The following fact (Dieudonné [5], Th. 5) will often be used:

when L is a linear subset of X , then

$$[L]_G = (L^\perp \cap G)^\top. \quad (50)$$

In particular, strongly closed linear subsets of X being also weakly closed (i.e. closed in the weak topology obtained by taking $G \equiv X^*$):

when L is a linear subset of X , then

$$\bar{L} = (L^\perp)^\top. \quad (51)$$

The preceding considerations apply equally well, of course, to the Banach space X^* . We shall need only the “weak*” topology (cf. Loomis [29], § 9): this is the weak topology induced on X^* by X (regarded as a subspace of X^{**}). The statement at (50) then becomes:

when M is a linear subset of X^ , then
the weak* closure of M is $(M^\top)^\perp$.*

Finally we shall need the fact that every finite-dimensional linear subset of X^* is weak* closed (see Bourbaki [3], Ch. I, § 2, No. 3).

The preceding results will be applied chiefly to the space $X \equiv l$ and its adjoint $X^* \equiv m$, but at one point it will be useful to note that they apply equally to the space $X \equiv c_0$ and its adjoint $X^* \equiv l$; some care will then be necessary in manipulating the symbols $(\)^\perp$ and $(\)^\top$. (We assume that the reader is familiar with the elementary properties of the spaces c_0 , l and m ; these can be found in Banach [1]. As usual we write c_0 for that subspace of m which consists of sequences $z \equiv (z_0, z_1, \dots)$ such that $z_\alpha \rightarrow 0$ when $\alpha \rightarrow \infty$.)

In what follows we shall often require the G -weak closure of the linear subset $\mathcal{R}(\Omega)$ of $X \equiv l$ for various total linear subsets G of $X^* \equiv m$. It is readily seen ([22], p. 167, eq. (42)) that $(\mathcal{R}(\Omega))^\perp = \mathcal{N}(\Omega^*)$; this shows incidentally that $\mathcal{N}(\Omega^*)$, being of the form E^\perp (where $E \subseteq l$), is always a weak* closed subset of $m = l^*$. It follows from (51) and (50) that

$$\overline{\mathcal{R}(\Omega)} = (\mathcal{N}(\Omega^*))^\top \tag{52}$$

and

$$[\mathcal{R}(\Omega)]_G = (\mathcal{N}(\Omega^*) \cap G)^\top. \tag{53}$$

6.3. A partial solution to Problem B_1 . We shall write u^i and v^j for the i th and j th unit vectors in l and m , so that $(u^i)_j = (v^j)_i = \delta_{ij}$ and $p_{ij}(t) = (v^j, P_t u^i)$. Then, for each i and j , $(v^j, (P_t - \Pi) u^i) \rightarrow 0$ as $t \rightarrow \infty$. But $\|P_t - \Pi\| \leq 2$ so that a double application of the Banach-Steinhaus theorem,¹ together with the facts that the linear sets spanned by the u^i and the v^j are dense in l and in c_0 , yields the important relation

$$\lim_{t \rightarrow \infty} (z, P_t x) = (z, \Pi x) \quad (x \in l, z \in c_0). \tag{54}$$

We now introduce the transition operator λJ_λ defined for all $\lambda > 0$ by

$$\lambda J_\lambda x \equiv \int_0^\infty \lambda e^{-\lambda t} P_t x dt \quad (x \in l) \tag{55}$$

¹ See HILLE [17], Th. 2.12.1.

and enjoying the properties

$$J_\lambda x \in \mathcal{D}(\Omega), \quad (\lambda I - \Omega) J_\lambda x = x \quad (x \in l), \quad (56)$$

$$J_\lambda (\lambda I - \Omega) x = x \quad (x \in \mathcal{D}(\Omega)). \quad (57)$$

From (54) and (55) we easily find that

$$\lim_{\lambda \downarrow 0} (z, \lambda J_\lambda x) = (z, \Pi x) \quad (x \in l, z \in c_0). \quad (58)$$

This makes it plain that we are here concerned with “ c_0 -weak ergodicity” in the terminology of [22]; the results of [22], however, cannot be applied because c_0 may fail to be invariant under the operators λJ_λ^* acting on m . The methods of [22] will now be adapted to prove the following theorem: this is as close as we can get to a general solution of Problem B₁ without resorting to the methods of § 3.

THEOREM 9.

- (i) *The range $\mathcal{R}(\Pi)$ and nullspace $\mathcal{N}(\Pi)$ of the ergodic projection operator Π always satisfy the relations*

$$\mathcal{R}(\Pi) = \mathcal{N}(\Omega) \quad (59)$$

and
$$\overline{\mathcal{R}(\Omega)} \subseteq \mathcal{N}(\Pi) \subseteq [\mathcal{R}(\Omega)]_{c_0}. \quad (60)$$

- (ii) *There is an example in which*

$$\overline{\mathcal{R}(\Omega)} \subset \mathcal{N}(\Pi) \subset [\mathcal{R}(\Omega)]_{c_0}.$$

Proof. The Lemma in § 3.1 shows that when $x \in \mathcal{N}(\Omega)$, then $P_t x = x$ for all $t \geq 0$ and so, from (54), $x = \Pi x \in \mathcal{R}(\Pi)$. On the other hand if $x \in \mathcal{R}(\Pi)$ then $x = \Pi x = P_t \Pi x = P_t x$ and so $\Omega x = \lim t^{-1}(P_t x - x) = 0$. Thus $\mathcal{R}(\Pi) = \mathcal{N}(\Omega)$.

Next, if $x \in \mathcal{R}(\Omega)$ then $x = \Omega y = \lim t^{-1}(P_t y - y)$ and so $\Pi x = \lim t^{-1}(\Pi P_t y - \Pi y) = \lim t^{-1}(\Pi y - \Pi y) = 0$. This shows that $\mathcal{R}(\Omega) \subseteq \mathcal{N}(\Pi)$ and because $\mathcal{N}(\Pi)$ is strongly closed it follows that $\overline{\mathcal{R}(\Omega)} \subseteq \mathcal{N}(\Pi)$.

Finally the relation (56) gives

$$(z, x) = (z, \lambda J_\lambda x) - (z, \Omega J_\lambda x)$$

for all $\lambda > 0$, all $x \in l$ and all $z \in c_0$; from this and from (58) it follows that

$$(z, x) = \lim_{\lambda \downarrow 0} (z, \Omega(-J_\lambda x))$$

whenever $\Pi x = 0$. This shows that every $x \in \mathcal{N}(\Pi)$ is the c_0 -weak limit of a sequence

of elements of $\mathcal{R}(\Omega)$, so that $\mathcal{N}(\Pi) \subseteq [\overline{\mathcal{R}(\Omega)}]_c$, and the proof of (i) is now complete. The example referred to at (ii) will be given in § 6.7.

The facts set out in Theorem 9 at once suggest the following questions.

(1°) How, using a knowledge of Ω alone, can one recognise the class of processes for which

$$\mathcal{N}(\Pi) = \overline{\mathcal{R}(\Omega)},$$

and the class of processes for which

$$\mathcal{N}(\Pi) = [\mathcal{R}(\Omega)]_c,$$

and how can one describe each class of processes in probabilistic terms?

(2°) Can we find a weak topology for l with regard to which the closure of $\mathcal{R}(\Omega)$ is *always* equal to $\mathcal{N}(\Pi)$?

We shall answer both questions, but our answer to (2°) will be somewhat unsatisfactory since the specification of the topology will involve prior knowledge of the ergodic behaviour of the system.

6.4. *Markov processes for which $\mathcal{N}(\Pi) = \overline{\mathcal{R}(\Omega)}$.* We know that $l = \mathcal{R}(\Pi) \oplus \mathcal{N}(\Pi)$, that $\mathcal{R}(\Pi) = \mathcal{N}(\Omega)$, and that $\mathcal{N}(\Pi) \supseteq \overline{\mathcal{R}(\Omega)}$. The last inclusion can sometimes be replaced by equality:

$$\mathcal{N}(\Pi) = \overline{\mathcal{R}(\Omega)}, \quad (61)$$

and we shall now find various conditions which ensure the truth of (61).

First there is a simple necessary and sufficient condition which can be stated in terms of Ω alone. It is clear that $\mathcal{N}(\Omega)$ and $\overline{\mathcal{R}(\Omega)}$ have only the element 0 in common, so that it is meaningful to consider the direct sum $\mathcal{N}(\Omega) \oplus \overline{\mathcal{R}(\Omega)}$. (It is shown in [22] that this direct sum is always strongly closed, but we shall not need to use this fact.) We now assert that (61) *holds if and only if*

$$l = \mathcal{N}(\Omega) \oplus \overline{\mathcal{R}(\Omega)}. \quad (62)$$

It is obvious that (61) implies (62). Conversely if (62) holds, let $x \in \mathcal{N}(\Pi)$. We can write $x = u + v$, where $u \in \mathcal{N}(\Omega) = \mathcal{R}(\Pi)$ and $v \in \overline{\mathcal{R}(\Omega)} \subseteq \mathcal{N}(\Pi)$; thus $x - v = u$ lies in $\mathcal{N}(\Pi)$ and $\mathcal{R}(\Pi)$ and hence $x - v = 0$, so that $x = v \in \overline{\mathcal{R}(\Omega)}$. It follows that $\mathcal{N}(\Pi) \subseteq \overline{\mathcal{R}(\Omega)}$, which implies (61). Recalling equation (52), we obtain the following partial solution¹ to Problem B₁:

¹ Like the solution of § 3.2, it demands only a knowledge of $\mathcal{N}(\Omega)$ and $\mathcal{N}(\Omega^*)$.

Calculate $\mathcal{N}(\Omega)$ and $\overline{\mathcal{R}(\Omega)} = (\mathcal{N}(\Omega^*))^\top$. Then if (62) holds, resolve $x \in l$ into two components corresponding to the direct decomposition (62); the component in $\mathcal{N}(\Omega)$ will be Πx .

Processes for which (61), or equivalently (62), holds will be called *strongly ergodic*, for a reason which will shortly become apparent.

We saw during the proof of Theorem 9 that when $x \in \mathcal{N}(\Omega)$, then $P_t x = x$ and hence $\lambda J_\lambda x = x$ for all $\lambda > 0$. Also, when $x \in \overline{\mathcal{R}(\Omega)}$ so that $x = \Omega y$ (say), then $\lambda J_\lambda x = \lambda J_\lambda \Omega y = \lambda(\lambda J_\lambda y - y)$ and so $\|\lambda J_\lambda x\| \leq 2\lambda\|y\| \rightarrow 0$ as $\lambda \downarrow 0$. Now $\|\lambda J_\lambda\| = 1$ for all $\lambda > 0$, and so we deduce by using the Banach-Steinhaus theorem that $\|\lambda J_\lambda x\| \rightarrow 0$ as $\lambda \downarrow 0$ whenever $x \in \overline{\mathcal{R}(\Omega)}$. Now assume that the process is strongly ergodic; then for each $x \in l$ we can write $x = \Pi x + (x - \Pi x)$ where $\Pi x \in \mathcal{N}(\Omega)$ and $x - \Pi x \in \overline{\mathcal{R}(\Omega)}$. We can then conclude from the preceding remarks that $\lambda J_\lambda x \rightarrow \Pi x$ strongly as $\lambda \downarrow 0$.

Conversely, suppose that $\lambda J_\lambda x$, for each $x \in l$, converges strongly as $\lambda \downarrow 0$ to an element which from (58) must coincide with Πx . Then if $x \in \mathcal{N}(\Pi)$, $\lambda J_\lambda x = x + \Omega J_\lambda x$ tends strongly to $\Pi x = 0$, so that $\Omega(-J_\lambda x)$ tends strongly to x and $x \in \overline{\mathcal{R}(\Omega)}$. Thus $\mathcal{N}(\Pi)$ is contained in, and therefore equal to, $\overline{\mathcal{R}(\Omega)}$; the process is strongly ergodic.

We have now shown that strong ergodicity of the process is equivalent to the strong convergence of $\lambda J_\lambda x$ as $\lambda \downarrow 0$ for all x . A similar result holds for general semigroups of operators acting on an arbitrary Banach space (see Hille [17], Th. 14.7.1, Kendall & Reuter [22], Th. 8), but in the present special case of a transition semigroup acting on l much more is true: we shall now show that strong ergodicity is also equivalent to the strong convergence of $P_t x$ as $t \rightarrow \infty$.

To prove this, suppose first that $\lambda J_\lambda x \rightarrow \Pi x$ (strongly) as $\lambda \downarrow 0$ for all x . Writing e for the element of m given by $(e)_i \equiv 1$ ($i = 0, 1, 2, \dots$), we have

$$(e, \Pi x) = \lim_{\lambda \downarrow 0} (e, \lambda J_\lambda x) = (e, x),$$

because λJ_λ is a transition operator and so $(e, \lambda J_\lambda x) = \|\lambda J_\lambda x\| = \|x\| = (e, x)$ whenever $x \geq 0$, whence $(e, \lambda J_\lambda x) = (e, x)$ for all x . On setting $x = u^i$ we obtain

$$\sum_{\alpha} \pi_{i\alpha} = 1 \quad (i = 0, 1, 2, \dots). \quad (63)$$

Next, suppose merely that (63) holds and put

$$\sigma_i(t) \equiv \sum_{\alpha} |p_{i\alpha}(t) - \pi_{i\alpha}|.$$

Choose any $J \geq 0$; then

$$\begin{aligned} \sigma_i(t) &\leq \sum_{\alpha \leq J} |p_{i\alpha}(t) - \pi_{i\alpha}| + \sum_{\alpha > J} p_{i\alpha}(t) + \sum_{\alpha > J} \pi_{i\alpha} \\ &= \sum_{\alpha \leq J} |p_{i\alpha}(t) - \pi_{i\alpha}| + 1 - \sum_{\alpha \leq J} p_{i\alpha}(t) + \sum_{\alpha > J} \pi_{i\alpha}. \end{aligned}$$

On letting $t \rightarrow \infty$ we find that

$$0 \leq \limsup_{t \rightarrow \infty} \sigma_i(t) \leq 1 - \sum_{\alpha \leq J} \pi_{i\alpha} + \sum_{\alpha > J} \pi_{i\alpha} = 2 \sum_{\alpha > J} \pi_{i\alpha},$$

and for each fixed i the right-hand side may be made as small as we please by suitable choice of J . Thus (63) implies that

$$\sigma_i(t) = \|P_t u^i - \Pi u^i\| \rightarrow 0 \quad \text{as } t \rightarrow \infty, \quad (64)$$

for each i . Now $\|P_t - \Pi\| \leq 2$ for all t and the u^i span a linear set dense in l , so that the Banach–Steinhaus theorem gives

$$\|P_t x - \Pi x\| \rightarrow 0 \quad \text{as } t \rightarrow \infty \quad (\text{all } x \in l). \quad (65)$$

Finally (65) implies, by a standard Abelian argument, that $\lambda J_\lambda x \rightarrow \Pi x$ strongly as $\lambda \downarrow 0$, and this completes the proof that (65) is equivalent to the strong ergodicity of the process.

In the course of the above proof we have shown that (63) is also equivalent to strong ergodicity. If we use the description of the Π -matrix given in § 3.2, we find that (63) is equivalent to

$$\sum_{\rho} \varpi(\rho, C^e) = 1 \quad (e = 0, 1, 2, \dots), \quad (66)$$

where the summation is over the positive classes.

Our last characterisation of strongly ergodic processes will involve the relation between $\mathcal{N}(\Pi)$ and the m -vectors ϖ^e . Clearly

$$\Pi x = \sum_{\rho} (\varpi^e, x) \pi^e \quad (67)$$

where the summation is over the positive classes; the series in (67) is absolutely convergent because

$$\|(\varpi^e, x) \pi^e\| = |(\varpi^e, x)| \leq \sum_{\alpha} |x_{\alpha}| \varpi(\alpha, C^e)$$

and so

$$\sum_{\rho} \|(\varpi^e, x) \pi^e\| \leq \sum_{\rho} \sum_{\alpha} |x_{\alpha}| \varpi(\alpha, C^e) \leq \sum_{\alpha} |x_{\alpha}| < \infty.$$

Accordingly $x \in \mathcal{N}(\Pi)$ if and only if $(\varpi^e, x) = 0$ for all ρ ; i.e. $\mathcal{N}(\Pi) = V^\perp$, where V is the linear set in m spanned by $\{\varpi^e : \rho = 1, 2, \dots\}$. If the process is strongly ergodic then

$$V^{\perp\perp} = (\mathcal{N}(\Pi))^{\perp} = (\overline{\mathcal{R}(\Omega)})^{\perp} = (\mathcal{R}(\Omega))^{\perp} = \mathcal{N}(\Omega^*).$$

Conversely if $V^{\perp\perp} = \mathcal{N}(\Omega^*)$ then

$$(\mathcal{N}(\Pi))^\perp = V^{\perp\perp} = \mathcal{N}(\Omega^*) = \overline{(\mathcal{R}(\Omega))}^\perp$$

and so

$$(\mathcal{N}(\Pi))^{\perp\perp} = \overline{(\mathcal{R}(\Omega))}^{\perp\perp}.$$

But $\mathcal{N}(\Pi)$ and $\overline{\mathcal{R}(\Omega)}$ are strongly closed linear subsets of l , and so it follows (using (51)) that $\mathcal{N}(\Pi) = \overline{\mathcal{R}(\Omega)}$, i.e. that the process is strongly ergodic.

From the remarks in § 6.2 it will be seen that $V^{\perp\perp}$ is the weak* closure of V , and so we have shown that strong ergodicity is equivalent to the requirement that $\mathcal{N}(\Omega^*)$ be the weak* closure of the linear set spanned by the vectors ϖ^e in m . (Note that $\mathcal{N}(\Omega^*)$ is weak* closed and contains the vectors ϖ^e , so that in any case $\mathcal{N}(\Omega^*) \supseteq V^{\perp\perp}$.)

If the number of positive classes is finite then V will have finite dimension and so will be weak* closed. The necessary and sufficient condition for strong ergodicity then becomes: $\mathcal{N}(\Omega^*) = V$, i.e. $\mathcal{N}(\Omega^*)$ is spanned by the vectors ϖ^e . If we write d and d^* for the dimensions of $\mathcal{N}(\Omega)$ and $\mathcal{N}(\Omega^*)$ (so that d is the number of positive classes) then the last result can be expressed as follows: if $d < \infty$ then the process is strongly ergodic if and only if $d^* = d$. The condition (66) also simplifies when d is finite: it then becomes $e = \sum_{\varrho} \varpi^e$.

We now collect all the preceding results.

THEOREM 10. *The process will be strongly ergodic, i.e. will have the property that $\mathcal{N}(\Pi) = \overline{\mathcal{R}(\Omega)}$, if and only if any one of the following equivalent conditions is satisfied.*

- (1°) For each $x \in l$, $P_t x$ converges strongly to a limit (necessarily equal to Πx) as $t \rightarrow \infty$.
- (2°) For each $x \in l$, $\lambda J_\lambda x$ converges strongly to a limit (necessarily equal to Πx) as $\lambda \downarrow 0$.
- (3°) For each $i = 0, 1, 2, \dots$,

$$\sum_{\alpha} |p_{i\alpha}(t) - \pi_{i\alpha}| \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

- (4°) For each $i = 0, 1, 2, \dots$, $\sum_{\alpha} \pi_{i\alpha} = 1$.

- (5°) For each initial state $i = 0, 1, 2, \dots$,

$$\sum_{\varrho} \varpi(i, C^{\varrho}) = 1.$$

- (6°) The vectors $\{\varpi^e : \varrho = 1, 2, \dots\}$ span a linear set V in m whose weak* closure is $\mathcal{N}(\Omega^*)$.

COROLLARY. Let $d \equiv \dim \mathcal{N}(\Omega)$ be the number of positive classes and let $d^* \equiv \dim \mathcal{N}(\Omega^*)$. Then, in the special case when d is finite, either of the following equivalent conditions is necessary and sufficient for strong ergodicity.

$$(7^\circ) \quad d^* = d.$$

$$(8^\circ) \quad e = \sum_e \varpi^e.$$

Condition (7^o) of the Corollary is very useful in practice; like the earlier condition (62), it allows us to detect strong ergodicity and thence to calculate Π , by inspecting $\mathcal{N}(\Omega)$ and $\mathcal{N}(\Omega^*)$. (Condition (7^o) has an analogue for general semigroups; see Theorem 9 of [22].) The remaining conditions, (1^o)–(6^o) and (8^o) cannot by their very nature assist us in calculating Π , but they do indicate to what extent the ergodic theory of general semigroups can be applied to the present special case. For instance (3^o) and (4^o) both imply that processes with finitely many states are always strongly ergodic; also (5^o) shows that processes without dissipative states are strongly ergodic, whereas processes in which all states are dissipative are not strongly ergodic. In § 6.5 we shall give a probabilistic formulation of (5^o) which will show how strong ergodicity is controlled by the behaviour of the system in relation to the dissipative states.

Theorem 10 will now be illustrated by “the random walk in continuous time with an absorbing barrier at the origin”. This is the process of § 5.5 with

$$a_1 = a_2 = a_3 = \dots = a > 0,$$

$$b_1 = b_2 = b_3 = \dots = b > 0,$$

and with $b_0 = 0$ as before. The process is regular ((42) holds) and the operator Q is bounded; because $Q_0 \subseteq \Omega_F \subseteq Q$ and Ω_F is a closed operator, it follows that $\Omega \equiv \Omega_F = Q$ and $\Omega^* = Q^*$. The calculations in § 5.5 show that $\mathcal{N}(\Omega)$ is one-dimensional and is spanned by the vector

$$u^0 \equiv [1, 0, 0, \dots];$$

there is just one positive class, consisting of the zero-state alone. As for $\mathcal{N}(\Omega^*)$, there are two possibilities:

Case 1. $a \geq b$ (so that $T = \infty$). Then $d^* = 1$ and $\mathcal{N}(\Omega^*)$ is spanned by $e \equiv [1, 1, 1, \dots]$.

Case 2. $a < b$ (so that $T = b/(b-a) < \infty$). Then $d^* = 2$ and $\mathcal{N}(\Omega^*)$ is spanned by

e and

$$\varpi \equiv [1, r, r^2, \dots],$$

where $r = a/b$.

From (7°) of the Corollary we see that there is strong ergodicity in Case 1 but not in Case 2. In Case 1, $\mathcal{R}(\Pi)$ consists of all multiples of u^0 and $\mathcal{N}(\Pi)$ is

$$(\mathcal{N}(\Omega^*))^T = \{x : x \in l \text{ and } \sum_{\alpha} x_{\alpha} = 0\};$$

hence Π is given by

$$\Pi x = \left(\sum_{\alpha} x_{\alpha}\right) u^0.$$

The reader will find it instructive to verify the various assertions of Theorem 10 for this example.

6.5. The probabilistic significance of strong ergodicity. The condition

$$\sum_{e} \varpi(i, C^e) = 1 \quad (i = 0, 1, 2, \dots) \quad (66)$$

for strong ergodicity will now be expressed in probabilistic terms.

If $(\Omega, \mathcal{F}, \text{pr})$ is a probability space, i.e. if Ω is a non-empty set of points ω , if \mathcal{F} is a Borel field of subsets of Ω and if pr is a countably additive non-negative measure on \mathcal{F} such that (\mathcal{F}, pr) is complete¹ and $\text{pr}(\Omega) = 1$, then a family $\{X_t^{(i)} : t \geq 0\}$ of random variables (\mathcal{F} -measurable ω -functions) taking values in the compactified set $\{0, 1, 2, \dots; \infty\}$ will be called a *representation* of the Markov process with $i (= 0, 1, 2, \dots)$ as initial state when

$$(i) \quad X_0^{(i)}(\omega) = i \quad \text{for all } \omega \in \Omega$$

and

$$(ii) \quad \text{pr} \{X_{t_1}^{(i)} = j_1, \dots, X_{t_n}^{(i)} = j_n\} = p_{i j_1}(t_1) p_{j_1 j_2}(t_2 - t_1) \dots p_{j_{n-1} j_n}(t_n - t_{n-1})$$

$$\text{for all } n \geq 1, \quad 0 < t_1 < t_2 < \dots < t_n$$

$$\text{and } j_1, j_2, \dots, j_n = 0, 1, 2, \dots$$

The representation is called *separable* (relative to the class of closed sets) when there exist a countable dense subset S of $(0, \infty)$ and an ω -set $\Lambda \in \mathcal{F}$ with $\text{pr}(\Lambda) = 0$, with the following property:

- (iii) if $\omega \notin \Lambda$, if J is any open subinterval of $(0, \infty)$ and if C is any compact subset of $\{0, 1, 2, \dots; \infty\}$ (i.e. any finite set of integers or any infinite set of integers with the compactification point ∞ adjoined) then $X_t^{(i)}(\omega) \in C$ for all $t \in J$ if and only if $X_s^{(i)}(\omega) \in C$ for all $s \in J \cap S$.

¹ (\mathcal{F}, pr) is said to be "complete" when $A \subseteq Z \in \mathcal{F}$ and $\text{pr}(Z) = 0$ together imply that $A \in \mathcal{F}$. There is no loss of generality in assuming this. It will be recalled that the symbol Ω has also been used to denote the infinitesimal generator, but this should cause no confusion.

It is known (Doob [9], Ch. II, Th. 2.4) that a separable representation can always be found, and it is even possible¹ to choose S in such a way that Λ can be taken to be the empty set. In practice, however, it is often desirable to identify S with some particular countable dense set such as the set of positive rationals, and to justify this step we need the further result (a consequence of the continuity-in-probability of the process;² see Doob [7], Th. 8, and Doob [9], Ch. II, Th. 2.2) that if $\{X_t^{(i)}: t \geq 0\}$ satisfies (i), (ii) and (iii) in relation to a *particular* pair (S_0, Λ_0) and if S is *any* countable dense subset of $(0, \infty)$ then a null set $\Lambda = \Lambda(S)$ can always be found so that $\{X_t^{(i)}: t \geq 0\}$ satisfies (i), (ii) and (iii) in relation to (S, Λ) .

The advantage of using a separable representation depends on the facts that (a) only the joint distributions of *finite* sets $\{X_{t_1}^{(i)}, X_{t_2}^{(i)}, \dots, X_{t_n}^{(i)}\}$ of random variables are of practical importance, and (b) for theoretical purposes we often want to assign probabilities to events (ω -sets) whose specification imposes restrictions on an *uncountable* collection of the $X_t^{(i)}$. Thus all the representations satisfying (i) and (ii) are equally acceptable for *practical* purposes, while it will be *mathematically* convenient to select one of these which also satisfies (iii). Let us suppose that this has been done for some given initial state i (there will then be no ambiguity in dropping the index i from $X_t^{(i)}$). We shall prove that

$$\varpi(i, C^q) = \text{pr} \{X_t \in C^q \text{ for all sufficiently large } t\}$$

where

$$C^q_+ \equiv \begin{cases} C^q & \text{when } C^q \text{ is finite,} \\ C^q \cup \{\infty\} & \text{when } C^q \text{ is infinite,} \end{cases}$$

and $q = 1, 2, 3, \dots$

We start from the fact (see § 3.1) that if $\tau > 0$ is fixed then the classes C^q and the quantities $\varpi(i, C^q)$ are the same for the process $\{p_{jk}(t): t \geq 0\}$ as for the chain $\{p_{jk}(n\tau): n = 0, 1, 2, \dots\}$, and $\{X_{n\tau}: n = 0, 1, 2, \dots\}$ is a representation of this chain with i as initial state. It is known (see Chung [4], p. 26) that for this chain $\varpi(i, C^q)$ is an "absorption probability": it is the probability that the system will ultimately enter (and thereafter remain in) the positive class C^q . Thus

$$\varpi(i, C^q) = \text{pr} \{\omega: X_{n\tau}(\omega) \in C^q \text{ for all sufficiently large } n\}.$$

Now, having already fixed the initial state i and chosen $\tau > 0$, choose S to consist of all positive rational multiples of τ and then consider the ω -set

¹ This is clear from Doob's proof.

² This in turn is a consequence of the fact that $p_{jj}(t) \rightarrow 1$ as $t \downarrow 0$ for each j .

$$E_\varrho \equiv \{\omega : X_t(\omega) \in C_\varrho^+ \text{ for all } t \geq \text{some } T(\omega)\}.$$

We must show that $E_\varrho \in \mathcal{F}$ and that $\text{pr}(E_\varrho) = \varpi(i, C^e)$.

Clearly E_ϱ can also be defined by

$$E_\varrho \equiv \bigcup_{N \geq 1} \{\omega : X_t(\omega) \in C_\varrho^+ \text{ for all } t > N\tau\} \equiv \bigcup_{N \geq 1} E_\varrho(N),$$

say, and $E_\varrho(N)$ is a subset of and differs from

$$E'_\varrho(N) \equiv \{\omega : X_s(\omega) \in C_\varrho^+ \text{ for all } s \in S \cap (N\tau, \infty)\}$$

by a subset of Λ . Also, as a consequence of $\sum_\alpha p_{i\alpha}(t) = 1$, we have $\text{pr}\{\omega : X_t(\omega) = \infty\} = 0$ for each fixed t , and so

$$E''_\varrho(N) \equiv \{\omega : X_s(\omega) \in C^e \text{ for all } s \in S \cap (N\tau, \infty)\}$$

is a subset of and differs from $E'_\varrho(N)$ by a set of zero probability. Finally $E''_\varrho(N)$ is a subset of

$$E'''_\varrho(N) \equiv \{\omega : X_{n\tau}(\omega) \in C^e \text{ for all integers } n > N\},$$

and

$$\bigcup_{N \geq 1} E'''_\varrho(N) - \bigcup_{N \geq 1} E''_\varrho(N) \subseteq \bigcup_{n \geq 2} \bigcup_{\substack{s > n\tau \\ s \in S}} F_{n,s}$$

where

$$F_{n,s} \equiv \{\omega : X_{n\tau}(\omega) \in C^e, X_s(\omega) \notin C^e\} \quad (n\tau < s \in S).$$

But now

$$\text{pr}(F_{n,s}) = \sum_{\alpha \in C^e} \sum_{\beta \notin C^e} p_{i\alpha}(n\tau) p_{\alpha\beta}(s - n\tau) = 0$$

because $\beta \notin C^e$ is inaccessible from $\alpha \in C^e$ in the Markov chain associated with the time-interval $\tau' \equiv s - n\tau > 0$. It follows that $E_\varrho \in \mathcal{F}$ and that

$$\begin{aligned} \text{pr}(E_\varrho) &= \text{pr}\left(\bigcup_N E_\varrho(N)\right) = \text{pr}\left(\bigcup_N E'''_\varrho(N)\right) \\ &= \text{pr}\{\omega : X_{n\tau}(\omega) \in C^e \text{ for all sufficiently large } n\} \\ &= \varpi(i, C^e), \end{aligned}$$

as required.

The above proof also shows that

$$\begin{aligned} \text{pr}\left(\bigcup_\varrho E_\varrho\right) &= \text{pr}\left(\bigcup_\varrho \bigcup_N E_\varrho(N)\right) = \text{pr}\left(\bigcup_\varrho \bigcup_N E'''_\varrho(N)\right) \\ &= \sum_\varrho \text{pr}\left(\bigcup_N E'''_\varrho(N)\right) = \sum_\varrho \varpi(i, C^e) \end{aligned}$$

(because the positive classes are disjoint), and so on combining this with (5°) of Theorem 10 we obtain

THEOREM 11. Let $\{X_t^{(i)} : t \geq 0\}$ be a separable representation of the Markov process with state-space $\{0, 1, 2, \dots; \infty\}$ and initial state i ($= 0, 1, 2, \dots$) associated with the array $\{p_{jk}(t) : j, k = 0, 1, 2, \dots; t \geq 0\}$ of transition probabilities satisfying the conditions

$$\sum_{\alpha} p_{j\alpha}(t) = 1,$$

$$\sum_{\alpha} p_{j\alpha}(u) p_{\alpha k}(v) = p_{jk}(u+v),$$

$$\lim_{t \downarrow 0} p_{jk}(t) = p_{jk}(0) = \delta_{jk}.$$

Then

- (1°) $\varpi(i, C^e)$ is the probability that the system will ultimately enter and thereafter remain in the positive class C^e (augmented, if not finite, by the adjunction of the state ∞).
- (2°) $\sum_e \varpi(i, C^e) = 1$ if and only if the system, with probability one, ultimately enters and remains in some one of the augmented positive classes.

COROLLARY. The array $\{p_{jk}(t)\}$ of transition probabilities will have the property of strong ergodicity defined in § 6.4 if and only if, for every initial state i and for some one (and then for every) associated separable process $\{X_t^{(i)} : t \geq 0\}$, the system with probability one ultimately enters and remains in some one of the augmented positive classes.

It will have been noticed that the augmentation of the infinite positive classes was forced upon us by the use of the separability theory: one cannot guarantee to find a separable representation unless the state-space is first compactified. We shall now show that this apparent blemish in Theorem 11 is not due to a defect in method and that it can be associated with an essential feature of the stochastic motion. This can be seen, for instance, in the example analysed in § 5.2 (the “flash of flashes”). Here all states form a single positive class; the π_{ij} satisfy (4°) of Theorem 10, so that the system is strongly ergodic and the Corollary to Theorem 11 applies. Because the whole (uncompactified) state-space forms a single positive class it might be thought that the assertion of the Corollary *without* the word “augmented” would be a truism: in fact it would be *false*. An examination of the matrix-elements in the representation of the resolvent operator J_λ (for this, see Kendall [21]) reveals that the stochastic motion cannot be described for all $t \geq 0$ without the introduction of a countable infinity of fictitious states¹ each of which will be visited infinitely

¹ For the definition of “fictitious state”, see LÉVY [27], p. 348.

often whatever the initial state may be. The separability theory correctly anticipates this possibility, the fictitious states being lumped together to form a single compactification point ∞ .

6.6. *Markov processes for which $\mathcal{N}(\Pi) = [\mathcal{R}(\Omega)]_c$.* The examples at the end of § 6.4 show that $\mathcal{N}(\Pi)$ can be larger than $\overline{\mathcal{R}(\Omega)}$. However, Theorem 9 provides an upper as well as a lower bound to $\mathcal{N}(\Pi)$, and we shall now study those processes for which the upper bound is attained. In accordance with the main theme of this paper our first task must be to find a necessary and sufficient condition for the validity of

$$\mathcal{N}(\Pi) = [\mathcal{R}(\Omega)]_c, \quad (68)$$

the condition being of such a form that it can be checked using only a knowledge of Ω .

Now $l = \mathcal{N}(\Omega) \oplus \mathcal{N}(\Pi)$ and $\mathcal{N}(\Pi) \subseteq [\mathcal{R}(\Omega)]_c$, so that each element of l can be expressed as the sum of an element of $\mathcal{N}(\Omega)$ and an element of $[\mathcal{R}(\Omega)]_c$, in general in more than one way. The decomposition will be unique if and only if (68) holds, and alternatively if and only if

$$\mathcal{N}(\Omega) \cap [\mathcal{R}(\Omega)]_c = \{0\}. \quad (69)$$

This last condition, which can also be written as

$$\mathcal{N}(\Omega) \cap (\mathcal{N}(\Omega^*) \cap c_0)^\top = \{0\}, \quad (70)$$

is of the required form, and we have obtained another partial solution to Problem B₁:

Calculate $\mathcal{N}(\Omega)$ and $(\mathcal{N}(\Omega^*) \cap c_0)^\top$. Then if (70) holds we shall have

$$l = \mathcal{N}(\Omega) \oplus (\mathcal{N}(\Omega^*) \cap c_0)^\top,$$

and Πx will be the component of x in $\mathcal{N}(\Omega)$.

Note that once again the procedure requires only a knowledge of the nullspaces $\mathcal{N}(\Omega)$ and $\mathcal{N}(\Omega^*)$.

Now consider the following tableau:

Space :	c_0	l	m
	$\mathcal{N}(\Omega^*) \cap c_0$	$[\mathcal{R}(\Omega)]_c$	
		$\mathcal{N}(\Pi)$	$V^{\top 1}$
		$\overline{\mathcal{R}(\Omega)}$	$\mathcal{N}(\Omega^*)$

Note that the Banach spaces in the first row obey the relations $c_0^* \equiv l$, $l^* \equiv m$, and that each space contains the strongly closed linear sets listed in the same column. Also in each of the three situations of the form

$$\begin{pmatrix} X & X^* \\ E & F \end{pmatrix},$$

E is the annihilator in X of F and F is the annihilator in X^* of E . Notice lastly that because c_0 can be imbedded in m , $\mathcal{N}(\Omega^*) \cap c_0$ can also be considered as a subspace of m . We omit the proofs of the preceding assertions; they follow at once from the facts about weak topologies stated in § 6.2 and from the equality $\mathcal{N}(\Pi) = V^\top$ derived in § 6.4, V being the linear set in m spanned by the vectors ϖ^ϱ ($\varrho = 1, 2, 3, \dots$).

Next, observe that the annihilator in c_0 of $[\mathcal{R}(\Omega)]_c$ will be contained in the annihilator in the *larger* space m of the *smaller* set $\mathcal{N}(\Pi)$; that is,

$$\mathcal{N}(\Omega^*) \cap c_0 \subseteq V^{\top\perp}. \quad (71)$$

Let the weak* closure of $\mathcal{N}(\Omega^*) \cap c_0$, considered as a subset of m , be denoted by W , so that

$$W \equiv \{y : y \in m \text{ and } (y, x) = 0 \text{ for all } x \in [\mathcal{R}(\Omega)]_c\}. \quad (72)$$

Because $V^{\top\perp}$ is a weak* closed subset of m , (71) gives

$$W \subseteq V^{\top\perp}. \quad (73)$$

If (68) holds, then

$$\begin{aligned} W &= \{y : y \in m \text{ and } (y, x) = 0 \text{ for all } x \in \mathcal{N}(\Pi)\} \\ &= (\mathcal{N}(\Pi))^\perp = V^{\top\perp}. \end{aligned}$$

Conversely if $W = V^{\top\perp}$ then $\mathcal{N}(\Pi)$ ($= V^\top$) and $[\mathcal{R}(\Omega)]_c$ have the same annihilator in m ; it follows from (51) that they are identical. Thus we have shown that (68) holds if and only if $W = V^{\top\perp}$. Now this last condition is equivalent to the requirement that each of the vectors ϖ^ϱ should lie in W . For if $W = V^{\top\perp}$, then

$$\varpi^\varrho \in V \subseteq V^{\top\perp} = W;$$

conversely if each ϖ^ϱ lies in W then $V \subseteq W$ and so, taking weak* closures in m , $V^{\top\perp} \subseteq W \subseteq V^{\top\perp}$.

As before the criterion simplifies when the number of positive classes is finite, for then V has finite dimension and so coincides with $V^{\top\perp}$; (71) then shows that $\mathcal{N}(\Omega^*) \cap c_0$ is also finite-dimensional and so coincides with its own weak* closure as subset of m . Thus (68) will now hold if and only if every vector ϖ^ϱ lies in c_0 .

These results are summarised in

THEOREM 12. *A necessary and sufficient condition for*

$$\mathcal{N}(\Pi) = [\overline{\mathcal{R}(\Omega)}]_{c_0} \equiv (\mathcal{N}(\Omega^*) \cap c_0)^\tau \quad (68)$$

to hold is that each vector ϖ^ϱ ($\varrho = 1, 2, 3, \dots$) should lie in the weak* closure of $\mathcal{N}(\Omega^*) \cap c_0$, when this is considered as a subspace of m .

In particular (68) will hold if each vector ϖ^ϱ lies in c_0 .

When the number, $\dim \mathcal{N}(\Omega)$, of positive classes is finite, then (68) will hold if and only if each ϖ^ϱ lies in c_0 .

The condition, $\varpi^\varrho \in c_0$, can be rephrased thus: for each $\varepsilon > 0$ there exists a finite set $A^\varepsilon(\varepsilon)$ of states such that

$$\varpi(i, C^\varrho) < \varepsilon \quad \text{unless } i \in A^\varepsilon(\varepsilon).$$

Thus ϖ^ϱ cannot lie in c_0 if the class C^ϱ is infinite (because $\varpi(i, C^\varrho) = 1$ whenever $i \in C^\varrho$), and therefore we shall have $\mathcal{N}(\Pi) \subset [\overline{\mathcal{R}(\Omega)}]_{c_0}$, whenever there is only a finite number of positive classes and one at least contains an infinity of states.

As an illustration of the preceding theory let us examine again the example at the end of § 6.4. It will be recalled that there $\mathcal{N}(\Omega)$ consists of all multiples of $u^0 \equiv [1, 0, 0, \dots]$ and that

in Case 1 ($a \geq b$), $\mathcal{N}(\Omega^*) \cap c_0 = \{0\}$;

in Case 2 ($a < b$), $\mathcal{N}(\Omega^*) \cap c_0$ consists of all multiples of the vector $\varpi \equiv [1, r, r^2, \dots]$, where $r \equiv a/b$.

Accordingly $\mathcal{N}(\Omega)$ and $(\mathcal{N}(\Omega^*) \cap c_0)^\tau$ intersect in

$$\{0\} \quad \text{in Case 2,}$$

$$\{\lambda u^0 : \lambda \text{ real}\} \quad \text{in Case 1;}$$

thus (68) holds in Case 2 but not in Case 1. In Case 2 it follows at once that the ergodic projection operator is given by

$$\Pi x = \left(\sum_{\alpha} x_{\alpha} r^{\alpha} \right) u^0.$$

6.7. *A Markov process for which $\overline{\mathcal{R}(\Omega)} \subset \mathcal{N}(\Pi) \subset [\overline{\mathcal{R}(\Omega)}]_{c_0}$.* In the example discussed at the end of §§ 6.4, 6.6 we saw that

$$\overline{\mathcal{R}(\Omega)} = \mathcal{N}(\Pi) \subset [\overline{\mathcal{R}(\Omega)}]_{c_0} \quad \text{when } a \geq b,$$

$$\overline{\mathcal{R}(\Omega)} \subset \mathcal{N}(\Pi) = [\overline{\mathcal{R}(\Omega)}]_{c_0} \quad \text{when } a < b;$$

and the ergodic projection operator could be found in all cases by using one or other of the two partial solutions to Problem B₁ which were given in §§ 6.4 and 6.6. We now give an example which shows that one can have

$$\overline{\mathcal{R}(\Omega)} \subset \mathcal{N}(\Pi) \subset [\mathcal{R}(\Omega)]_c; \tag{74}$$

whenever this happens the methods of § 6 break down and the general method of § 3 must be used.

It will be convenient in the example to label the states as $(\dots, -2, -1; 0; 1, 2, \dots)$, components of vectors in l or m being labelled accordingly. We specify a conservative q -matrix as follows:

$$\begin{array}{cccc|ccc}
 & & & & (j=0) & & & \\
 \vdots & & & & & & & \vdots \\
 \cdots & -1 & 0 & 0 & 0 & 0 & 0 & 0 \cdots \\
 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\
 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\
 \hline
 (i=0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\
 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\
 \cdots & 0 & 0 & 0 & 0 & 0 & 1 & -1 \cdots \\
 \vdots & & & & & & & \vdots
 \end{array}$$

Regularity is easily checked and the boundedness of Q ensures as before that $\Omega \equiv \Omega_F = Q$ and $\Omega^* = Q^*$. Simple calculations show that $\mathcal{N}(\Omega)$ is spanned by the vector $u^0 \equiv [\dots, 0, 0; 1; 0, 0, \dots]$ and that $\mathcal{N}(\Omega^*)$ is spanned by the two vectors

$$\begin{aligned}
 v &\equiv [\dots, 1, 1; 0; 0, 0, \dots] \\
 w &\equiv [\dots, 0, 0; 1; 1, 1, \dots].
 \end{aligned}$$

Evidently $d = 1 \neq 2 = d^*$ and the process is not strongly ergodic. There is just one positive class consisting of the absorbing state labelled 0, and u^0 is the associated π -vector. The general element of $\mathcal{N}(\Omega^*)$ has the form $\lambda v + \mu w$; the condition $(\lambda v + \mu w, u^0) = 1$ requires that $\mu = 1$ and positivity requires that $\lambda \geq 0$. The minimal element ϖ is therefore given by $\mu = 1, \lambda = 0$, i.e. $\varpi = w$. Now the number of positive classes is finite and $\varpi \notin c_0$, so that (68) cannot hold. This completes the proof of (74).

It is instructive to identify the three sets occurring in the inequalities (74); they are

$$\begin{aligned}\overline{\mathcal{R}(\Omega)} &= \{x : x \in l \text{ and } \sum_{-\infty}^{-1} x_\alpha = 0 = \sum_0^{\infty} x_\alpha\}, \\ \mathcal{N}(\Pi) &= \{x : x \in l \text{ and } \sum_0^{\infty} x_\alpha = 0\}, \\ [\mathcal{R}(\Omega)]_{c_0} &= l.\end{aligned}$$

If we had used the c -weak (stronger than the c_0 -weak) topology to close $\mathcal{R}(\Omega)$, the subspace so obtained would have been

$$[\mathcal{R}(\Omega)]_c = \{x : x \in l \text{ and } \sum_{-\infty}^{\infty} x_\alpha = 0\},$$

which is not even comparable with $\mathcal{N}(\Pi)$. (As usual c denotes the subspace of m spanned by c_0 and the vector e all of whose components equal 1.)

6.8. *G*-weak topologies for which $\mathcal{N}(\Pi) = [\mathcal{R}(\Omega)]_G$. Our final task will be to answer the question (2°) at the end of § 6.3. We will show that one can find linear subsets G of m , total for l and such that

$$\mathcal{N}(\Pi) = [\mathcal{R}(\Omega)]_G. \quad (75)$$

Unfortunately our specifications of G will not be given in terms of Ω alone, so that our result will not provide a solution to Problem B₁.

We define the linear subsets A and B of m by

$$\begin{aligned}A &\equiv \{g : g \in m \text{ and } (g, \lambda J_\lambda x) \rightarrow (g, \Pi x) \text{ as } \lambda \downarrow 0, \text{ all } x \in l\}, \\ B &\equiv \{g : g \in m \text{ and } (g, P_t x) \rightarrow (g, \Pi x) \text{ as } t \rightarrow \infty, \text{ all } x \in l\}.\end{aligned}$$

From (54) and (55) it follows that

$$A \supseteq B \supseteq c_0, \quad (76)$$

so that A and B are both total for l .

We next observe that the three statements

- (i) $g \in \mathcal{N}(\Omega^*)$,
- (ii) $g P_t^* = g$ for all $t \geq 0$,
- (iii) $g \lambda J_\lambda^* = g$ for all $\lambda > 0$,

are equivalent. The equivalence of (i) and (ii) is asserted by the second part of the lemma in § 3.1; the equivalence of (ii) and (iii) follows from (55) and Lerch's theorem. It can now be inferred that both $\mathcal{N}(\Omega^*) \cap A$ and $\mathcal{N}(\Omega^*) \cap B$ coincide with the set

$$\{g : g \in m \text{ and } (g, x) = (g, \Pi x) \text{ for all } x \in l\}, \quad (77)$$

on using (46) and the similar relations involving λJ_λ in place of P_t .

But the set in (77) is $(\mathcal{R}(I - \Pi))^{\perp} = (\mathcal{N}(\Pi))^{\perp}$, and hence when G is either A or B we have

$$[\mathcal{R}(\Omega)]_G = (\mathcal{N}(\Omega^*) \cap G)^{\top} = (\mathcal{N}(\Pi))^{\perp\top} = \mathcal{N}(\Pi),$$

which establishes (75) for these two choices of G .

References

- [1]. S. BANACH, *Théorie des opérations linéaires*. Warsaw, 1932.
- [2]. D. BLACKWELL, On transient Markov processes with a countable number of states and stationary transition probabilities. *Ann. Math. Statistics*, 26 (1955), 654–658.
- [3]. N. BOURBAKI, *Espaces vectoriels topologiques*. Actualités Scientifiques et Industrielles, 1189 & 1229, Paris, 1953–55.
- [4]. K. L. CHUNG, *Notes on Markov Chains*. Columbia Graduate Math. Stat. Soc., 1951.
- [5]. J. DIEUDONNÉ, La dualité dans les espaces vectoriels topologiques. *Ann. Sci. École Norm. Sup. Paris* (3), 59 (1942), 107–139.
- [6]. R. L. DOBRUŠIN, On conditions of regularity of Markov processes which are stationary in time and have a denumerable set of possible states. *Uspehi Matem. Nauk.*, 7 (1952), 185–191.
- [7]. J. L. DOOB, Topics in the theory of Markoff chains. *Trans. Amer. Math. Soc.*, 52 (1942), 37–64.
- [8]. —, Markoff chains—denumerable case. *Trans. Amer. Math. Soc.*, 58 (1945), 455–473.
- [9]. —, *Stochastic Processes*. New York, 1953.
- [10]. N. DUNFORD, A mean ergodic theorem. *Duke Math. J.*, 5 (1939), 635–646.
- [11]. W. FELLER, On the integro-differential equations of purely discontinuous Markoff processes. *Trans. Amer. Math. Soc.*, 48 (1940), 488–515; *ibid.*, 58 (1945), 474.
- [12]. —, *An Introduction to Probability Theory and Its Applications*, I. New York, 1950.
- [13]. —, Boundaries induced by stochastic matrices. *Trans. Amer. Math. Soc.*, 83 (1956), 19–54.
- [14]. F. G. FOSTER, On Markov chains with an enumerable infinity of states. *Proc. Cambridge Phil. Soc.*, 48 (1952), 587–591.
- [15]. —, On the stochastic matrices associated with certain queuing processes. *Ann. Math. Statistics*, 24 (1953), 355–360.
- [16]. T. E. HARRIS, First passage and recurrence distributions. *Trans. Amer. Math. Soc.*, 73 (1952), 471–486.
- [17]. E. HILLE, *Functional Analysis and Semigroups*. New York, 1948.
- [18]. A. JENSEN, *A Distribution Model*. Copenhagen, 1954.
- [19]. S. KARLIN & J. MCGREGOR, Representation of a class of stochastic processes. *Proc. Nat. Acad. Sci. Wash., U.S.A.*, 41 (1955), 387–391.
- [20]. T. KATO, On the semigroups generated by Kolmogoroff's differential equations. *J. Math. Soc. Japan*, 6 (1954), 1–13.
- [21]. D. G. KENDALL, Some further pathological examples in the theory of denumerable Markov processes. *Quart. J. Math. Oxford (Ser. 2)*, 7 (1956), 39–56.
- [22]. D. G. KENDALL & G. E. H. REUTER, Some ergodic theorems for one-parameter semigroups of operators. *Phil. Trans. Roy. Soc. London (Ser. A)*, 249 (1956), 151–177.
- [23]. —, Some pathological Markov processes with a denumerable infinity of states and the associated semigroups of operators on l . *Proc. Intern. Congr. Math.*, Amsterdam, 1954, Vol. III, pp. 377–415.

- [24]. A. N. KOLMOGOROV, Anfangsgründe der Theorie der Markoffschen Ketten mit unendlich vielen möglichen Zuständen. *Mat. Sbornik (N.S.)*, 1 (1936), 607–610; *Bull. Univ. État Moscou (A)*, 1 (1937), 1–15.
- [25]. —, On some problems concerning the differentiability of the transition probabilities in a temporally homogeneous Markov process having a denumerable set of states. *Moskov. Gos. Univ. Učenyje Zapiski Matematika* (4), 148 (1951), 53–59.
- [26]. W. LEDERMANN & G. E. H. REUTER, Spectral theory for the differential equations of simple birth and death processes. *Phil. Trans. Roy. Soc. London (Ser. A)*, 246 (1954), 321–369.
- [27]. P. LÉVY, Systèmes markoviens et stationnaires : cas denombrable. *Ann. Sci. École Norm. Sup. Paris* (3), 68 (1951), 327–381.
- [28]. M. LOÈVE, *Probability Theory*. New York, 1955.
- [29]. L. H. LOOMIS, *An Introduction to Abstract Harmonic Analysis*. New York, 1953.
- [30]. R. S. PHILLIPS, A note on ergodic theory. *Proc. Amer. Math. Soc.*, 2 (1951), 663–670.
- [31]. G. E. H. REUTER, Denumerable Markov processes and the associated contraction semi-groups on l . *Acta Math.*, 97 (1957), 1–46.
- [32]. K. YOSIDA & S. KAKUTANI, Markov process with an enumerable infinite number of possible states. *Jap. J. Math.*, 16 (1939), 47–55.