Z. Wahrscheinlichkeitstheorie 3, 227–246 (1964)

On the Construction Problem for Markov Chains

By

DAVID WILLIAMS*

1. Introduction

Let $P = \{P(t) : t \geqq 0\}$ be a substochastic transition function on a countable set E. Then the (componentwise) derivative $P'(0) = Q$ exists and satisfies the conditions

$$-\infty \leqq q_{ii} \leqq 0, \quad q_{ij} \geqq 0, \quad (i \in E, j \in E, i \neq j) \qquad (1)$$
$$\sum_{k \in E} q_{ik} \leqq 0, \quad (i \in E).$$

Conversely, suppose given a countable set E and an $E \times E$ matrix Q satisfying the conditions (1). Denote by $\mathscr{I}_Q$ the class of all substochastic transition functions P on E such that $P'(0) = Q$. One of the basic problems of Markov chain theory is: given Q, construct $\mathscr{I}_Q$. In this paper, the problem is solved under the following two assumptions.

Assumption A. *Q is finite and conservative i. e.*

$$-\infty < q_{ii} < 0, \quad q_{ij} \geqq 0, \quad (i \in E, j \in E, i \neq j);$$
$$\sum_{k \in E} q_{ik} = 0, \quad (i \in E).$$

Under Assumption A, the dimension d of the space of bounded vectors x on E such that $(\lambda - Q)x = 0$ (λ a positive number) is independent of λ.

Assumption B. *The dimension d is finite.*

Suppose that $P \in \mathscr{I}_Q$ and that $\{X(t) : t \geqq 0\}$ is a (stopped) Markov chain with transition matrix P. In general, $\{X(t)\}$ will reach infinity at some point of the exit boundary induced by Q and our problem is essentially to analyse its possible modes of return. Assumption A precludes pathological behaviour on the part of $\{X(t)\}$ when $\{X(t)\}$ is not at the boundary and also guarantees that $\mathscr{I}_Q$ is not empty. The dimension d determines the number of "escape routes to infinity" available to $\{X(t)\}$. Assumption B, by restricting the exit boundary to be finite (in fact, of cardinality d) ensures that $\mathscr{I}_Q$ can not be too large.

It goes without saying that our analysis will rely very heavily on the results of FELLER's paper [*3*]. The method of proof of Theorem 2 was inspired by that used by REUTER ([*8*]) in this complete solution for the case "$d = 1$".

As explained by the author in [*10*], all of the processes constructed by FELLER may be obtained by the method of first "extending the minimal process to the boundary" and then "deleting the time spent on the boundary by the extended

* This work was done at Stanford University, Stanford, California and was supported in part by Grant AFOSR 62–243 from United States Air Force Office of Scientific Research Mathematical Sciences Directorate.

process", a method first used by NEVEU ([6]) in this "absolute dominance construction". The same method yields all processes constructed here. It becomes possible therefore to ignore probabilistic considerations until after Theorem 3 and then suddenly derive a very complete and rigorous probabilistic interpretation.

We shall however discuss informally the heuristic arguments which provide the motivation for Theorem 1. CHUNG ([2]) has recently explained how to make these ideas precise and has succeeded in giving purely probabilistic proofs of many of FELLER's main results. The situation considered here is more general than that considered by FELLER in two respects: the forward equations are not assumed and there is no restriction on the size of the entrance boundary. It therefore seems unlikely that purely probabilistic methods will lead to a solution *unless some deeper reason is found for why the "absolute dominance construction" works.*

I wish to thank Professors K. L. CHUNG and D. ORNSTEIN for some helpful comments on this work.

Notes. 1. The relation between the solution presented here and that announced by JURKAT is not clear.

2. The most important case of our analytic construction is summarized in § 3.8 and the reader is advised to read that subsection before the earlier part of Section 3.

3. Throughout the paper, "Markov chain" will mean what is, in strict terminology, "Markov chain *with stationary transition probabilities*".

2. Notation and prerequisites

A (standard) *substochastic transition function* $P \equiv \{P(t): t \geqq 0\}$ on E is characterized by the following relations:

$$p_{ij}(t) \geqq 0\,, \quad \sum_{k \in E} p_{ik}(t) \leqq 1\,, \quad \sum_{k \in E} p_{ik}(s)\, p_{kj}(t) = p_{ij}(s+t)\,,$$

$$\lim_{u \to 0} p_{ij}(u) = p_{ij}(0) = \delta_{ij}\,, \quad (s, t \geqq 0;\ i, j \in E)\,.$$

P is called *stochastic* if $\sum_{k \in E} p_{ik}(t) = 1$, $(i \in E,\ t \geqq 0)$.

Throughout the paper, $\{P(t): t \geqq 0\}$ will denote a substochastic transition function with $P'(0) = Q$, Q satisfying Assumptions A and B, and $\{X(t): t \geqq 0\}$ will denote a Markov chain with $\{P(t): t \geqq 0\}$ as its transition matrix. The shorthand $\mathrm{P}_i\{M\}$ will denote the probability of the event M conditional on the event $\{X(0) = i\}$ and, as usual, q_i will denote $-\,q_{ii}(i \in E)$.

The matrix Q provides $\{X(t)\}$ with a certain set of instructions: on entering any state i, $\{X(t)\}$ is to stay there for a random time α with $P\{\alpha > t\} = \exp(-q_i t)$; at the end of this time it is to jump to a randomly chosen state Y with $\mathrm{P}\{Y = j\} = q_{ij}/q_i$. The sequence $\{Y_0, Y_1, Y_2, \ldots\}$ of successive states visited by $\{X(t): t \geqq 0\}$ therefore forms a discrete parameter Markov chain (the "jump chain" of $\{X(t)\}$) with one-step transition matrix $[(1-\delta_{ij})\, q_{ij}/q_i]$.

From this information it is possible to calculate the probability $f_{ij}(t)$ that, starting at i, $\{X(\cdot)\}$ is at j at time t having made in the meantime only finitely many jumps. The functions $f_{ij}(t)$ are the elements of a substochastic transition function $F \equiv \{F(t) : t \geqq 0\}$, the *minimal transition function* associated with Q. It is clear that $p_{ij}(t) \geqq f_{ij}(t)$ for $i, j \in E$ and $t \geqq 0$; hence the name "minimal" for F. If F is stochastic, then it is the only element of $\mathscr{I}_Q$.

If $P \neq F$, then, with a positive probability, there will be a finite first time T_∞ by which $\{X(t)\}$ has made an infinite number of jumps. FELLER's theory allows us to assert that at such a time, $\{X(t)\}$ is at a certain point a of the *exit boundary* A induced by Q and, as one would hope, the set $E + A$ may be given a Hausdorff topology in such a way that a is the limit of the sequence $\{Y_n\}$. As already stated, A is of cardinality d.

Not all limit points of the sequence $\{Y_n\}$ are points of the exit boundary. For the purpose of boundary theory, each recurrent class for $\{Y_n\}$, equivalently, each recurrent class under the minimal transition function $\{F(t)\}$, must be identified with a limit point of $\{Y_n\}$. These recurrent classes, on which $\{F(t)\}$ is evidently stochastic, play but a trivial rôle in the theory and as they can be of nuisance value, we shall adopt the usual course and eliminate them. *We therefore assume that all states are transient under the minimal transition function* $\{F(t)\}$, or, in other terms,

$$\int_0^\infty f_{ii}(t)\,dt < \infty, \qquad (i \in E).$$

It is important to realize that *we do not exclude the possibility that some or all states are recurrent under the transition matrix* $\{P(t)\}$.

Under the assumption of transience, the chain $\{Y_n\}$ must drift towards infinity. The exit boundary is precisely the set of limit points of paths $J \equiv (j_0, j_1, j_2, \ldots)$ followed by $\{Y_n\}$ which may be traversed in a finite time by $\{X(t)\}$. We recall that the time taken by $\{X(t)\}$ to traverse the path J is (almost certainly) finite or infinite according as the mean path time $\sum q_{j_n}^{-1}$ is finite or infinite. The *passive boundary* consists of limit points of those paths followed by $\{Y_.\}$ which would take $\{X(t)\}$ an infinite time to traverse.

Let us now define:

$$L_i^a = \mathrm{P}_i\{X(T_\infty) = a\}, \qquad (i \in E,\ a \in A);$$
$$L_i^a = \mathrm{P}_i\{T_\infty \leqq t;\quad X(T_\infty) = a\}, \qquad (i \in E,\quad a \in A,\ t \geqq 0).$$

We shall write L_i^0 for the probability that, starting at i, no point of the exit boundary is reached i. e. that the process $\{X(t)\}$ approaches the passive boundary asymptotically.

The equations:
$$L_i^0 = 1 - \sum_{a \in A} L_i^a, \tag{2}$$

$$\sum_{j \in E} f_{ij}(t) + \sum_{a \in A} L_i^a(t) = 1, \tag{3}$$

$$L_i^a(s+t) - L_i^a(s) = \sum_{j \in E} f_{ij}(s)\, L_j^a(t), \tag{4}$$

$$L_i^a - L_i^a(s) = \sum_{j \in E} f_{ij}(s)\, L_j^a, \tag{5}$$

$$L_i^0 = \sum_{j \in E} f_{ij}(s)\, L_j^0, \tag{6}$$

express intuitively obvious facts. Also, it is clear from our discussion of the jump chain $\{Y_n\}$ that

$$\sum_{j \neq i} (q_{ij}/q_i)\, L_j^a = L_i^a, \qquad (i \in E,\ a \in A),$$

i. e.

$$\sum_{j \in E} q_{ij} L_j^a = 0, \qquad (i \in E,\ a \in A). \tag{7}$$

Equation (7) may also be derived from equations (2) — (6) and the fundamental result:

$$F'(t) = Q\,F(t) = F(t)\,Q\,,^{\star} \tag{8}$$

which properly belongs earlier in the exposition.

For our purposes, the above equations are more useful in their Laplace transformed versions. The following notation will be used:

$$\Phi(\lambda) \equiv \int_0^\infty \exp(-\lambda t)\, F(t)\, dt, \qquad (\lambda > 0);$$

$$x_i^a(\lambda) \equiv \int_0^\infty \exp(-\lambda t)\, dL_i^a(t), \qquad (i \in E,\ a \in A,\ \lambda > 0).$$

The symbol $x^a(\lambda)$ will denote the vector on E with i^{th} component $x_i^a(\lambda)$ and, from now on, we shall write x_i^a for L_i^a. The constant vector $(1, 1, 1, \ldots)$ on E will be denoted by 1 and the identity matrix on E by I. We now have:

(from (2)) $$x^0 = 1 - \sum_{a \in A} x^a; \tag{9}$$

(from (3)) $$\sum_{a \in A} x^a(\lambda) = [I - \lambda\Phi(\lambda)]\,1; \tag{10}$$

(from (4)) $$x^a(\lambda) - x^a(\mu) = (\mu - \lambda)\,\Phi(\lambda)\,x^a(\mu) = (\mu - \lambda)\,\Phi(\mu)\,x^a(\lambda); \tag{11}$$

(from (5)) $$x^a(\lambda) = [I - \lambda\Phi(\lambda)]\,x^a; \tag{12}$$

(from (6)) $$x^0 = \lambda\Phi(\lambda)\,x^0; \tag{13}$$

(from (7)) $$Q\,x^a = 0; \tag{14}$$

(from (8)) $$(\lambda - Q)\,\Phi(\lambda) = I = \Phi(\lambda)\,(\lambda - Q). \tag{15}$$

It is easily deduced from (12), (14) and (15) that

$$(\lambda - Q)\,x^a(\lambda) = 0\,.^{\star\star} \tag{16}$$

Much more than this is known. For each $\lambda > 0$, the vectors $x^a(\lambda)$ form an *extreme base* for the solutions of the equation $(\lambda - Q)\,y = 0$. More exactly:

(Lemma 1) *for* $\lambda > 0$, *the extreme points of the convex set of those vectors* y *satisfying*

$$0 \leqq y \leqq 1, \quad (\lambda - Q)\,y = 0$$

---

* Unless otherwise stated, matrix equations are to be interpreted in a componentwise sense.

** Note. Some care must be shown in the manipulation of the above identities because multiplication of infinite matrices is not always associative. Thus, for example,

$$[\Phi(\lambda)Q]x^a = -x^a(\lambda) \quad \text{whereas} \quad \Phi(\lambda)[Q x^a] = 0\,.$$

are precisely the $x^a(\lambda)$ $(a \in A)$, *and, moreover, every bounded solution of the equation* $(\lambda - Q)y = 0$ *is a linear combination of the* $x^a(\lambda)$.

The extremal property of the $x^a(\lambda)$ rests on the following fact:

(Lemma 2) *as* i *"converges" to the boundary point* b

$$x_i^a(\lambda) \to \delta^{ab} \quad \textit{and} \quad x_i^a \to \delta^{ab}. \tag{17}$$

(From a state i "near" the boundary point b, $X(t)$ will, with high probability, soon reach b.)

As stated in the Introduction, the task of adding precision to the above heuristic approach has been carried out by CHUNG. The same ground may be covered by adopting an analytic approach in the manner of FELLER's paper.

Notice that Lemma 1 provides a *completely algebraic characterization of the* $x^a(\lambda)$, independent of the notion of "boundary". Notice also that, once the $x^a(\lambda)$ are defined, Lemma 2 suggests the way to adjoin the set A to the state-space E, namely, by making sets on which $\{x^a(\lambda) > 1 - \eta\}$ $(0 < \eta < 1)$ the typical neighbourhoods of the point a. It was by using these considerations that FELLER first defined the exit boundary.

The information required in the sequel is that *there exist*:

(i) *a minimal transition function* $F(t)$ *with resolvent matrix* $\Phi(\lambda)$;

(ii) *a finite set* A *adjoined to* E *in such a manner that* $E + A$ *is a Hausdorff space and that each point of* A *is a limit point of* E;

(iii) *vectors* $x^a(\lambda)$, x^a, x^0;

such that equations (9)—(16) *and Lemmas* 1 *and* 2 *hold*. This is proved both in FELLER's paper and in CHUNG's.

3. The analytic construction*

3.1. Some further prerequisites. We write $r_{ij}(\lambda)$ for the Laplace transform of $p_{ij}(t)$:

$$r_{ij}(\lambda) = \int_0^\infty \exp(-\lambda t)\, p_{ij}(t)\, dt \qquad (\lambda > 0) \tag{18}$$

and $R(\lambda)$ for the $E \times E$ matrix matrix with $(i,j)^{\text{th}}$ component $r_{ij}(\lambda)$. Then $R(\lambda)$ satisfies:

$$\textit{(the resolvent equation)} \qquad R(\lambda) - R(\mu) = (\mu - \lambda)\, R(\lambda)\, R(\mu); \tag{19}$$

$$\textit{(the positivity condition)} \qquad R(\lambda) \geqq 0; \tag{20}$$

$$\textit{(the norm condition)} \qquad \lambda R(\lambda) 1 \leqq 1; \tag{21}$$

$$\textit{(the continuity condition)} \quad \lambda R(\lambda) \to I \quad \text{as} \quad \lambda \to \infty \tag{22}$$

Conversely, the relations (18)—(22) imply that $\{P(t) : t > 0\}$ is a substochastic transition function. $\{P(t)\}$ is stochastic if and only if $\lambda R(\lambda) 1 = 1$. It is clear that, in our situation,

$$R(\lambda) \geqq \Phi(\lambda) \tag{23}$$

and it is known that

$$(\lambda - Q)\, R(\lambda) = I. \tag{24}$$

* See Note 2 in the Introduction.

Equation (16) is simply the Laplace transformed version of the Kolmogorov backward equation:

$$P'(t) = Q\,P(t) \qquad (t \geqq 0)$$

which holds under Assumption A.

In view of the above results, *the problem of constructing* $\mathscr{I}_Q$ *is equivalent to that of constructing all matrix functions* $R(\lambda)$ *satisfying* (19), (21), (23) *and* (24).

On comparing equations (15) and (24), we observe that each column of the matrix $R(\lambda) - \Phi(\lambda)$ is a solution of the equation $(\lambda - Q)y = 0$ and hence, by Lemma 1, is a linear combination of the vectors $x^a(\lambda)$. This proves the first part of the following theorem.

Theorem 1. $R(\lambda)$ *has the decomposition:*

$$r_{ij}(\lambda) = \varphi_{ij}(\lambda) + \sum_{a\in A} x_i^a(\lambda)\, y_j^a(\lambda)\,. \tag{25}$$

The (row) vectors $y^a(\lambda)$ $(a \in A,\ \lambda > 0)$ *are non-negative and satisfy:*

$$\textit{(resolvent condition)} \qquad y^a(\lambda) - y^a(\mu) = (\mu - \lambda)\, y^a(\lambda)\, R(\mu)\,; \tag{26}$$

$$\textit{(norm condition)} \qquad \lambda \sum_{j\in E} y_j^a(\lambda) \leqq 1\,. \tag{27}$$

$\{P(t) : t \geqq 0\}$ *is stochastic if and only if*

$$\lambda \sum_{j\in E} y_j^a(\lambda) = 1 \qquad (a \in A,\ \lambda > 0)\,.$$

Analytically, the non-negativity of the y's and the equations (26) and (27) follow from Lemma 2 and (respectively) (23), (19) and (21). However, equation (25) has a simple probabilistic interpretation which implies the rest of Theorem 1.

Interpretation of Theorem 1. Suppose that the Markov chain $\{X(\cdot)\}$ started at i. If it is at j at time t, then either it has reached j without hitting the exit boundary or else it first hit the exit boundary at some point a during some time neighbourhood $(s, s + ds)$ (with $0 < s < t$) and then moved from a to j in the remaining time $t - s$. We therefore expect a decomposition:

$$p_{ij}(t) \equiv f_{ij}(t) + \sum_{a\in A} \int_0^t dL_i^a(s)\, \xi_j^a(t - s) \tag{28}$$

where

$$\xi_j^a(u) \equiv \mathrm{P}\{X(T_\infty + u) = j \mid X(T_\infty) = a\}\,.$$

A rigorous proof of (28) using these ideas and based on the Strong Markov Theorem is given in CHUNG [*2*; Theorem 5.1]. A completely different proof will be given in Section 4. Equation (25) is simply the Laplace transform of equation (28) and $y_j^a(\lambda)$ is therefore identified with the Laplace transform of $\xi_j^a(t)$.

3.2. Last exit decomposition of the $y^a(\lambda)$. Our purpose in this section is to decompose the $y^a(\lambda)$ into simpler elements. We shall see eventually that what is obtained is a decomposition modulo the last visit to the boundary though why the analytic method described below leads to probabilistically meaningful results, I do not know. What is clear however is that the behaviour of $\{X(\cdot)\}$ after its last visit to the boundary will be governed by "entrance solutions" of the type we now introduce.

Entrance solutions. By an *entrance solution* we shall mean a set of non-negative functions $v_i(\cdot)$ $(i \in E)$ on $(0, \infty)$ satisfying the equation:

$$\sum_{i \in E} v_i(s) f_{ij}(t) = v_j(s+t) \quad (v(s) F(t) = v(s+t)).$$

An entrance solution for $\{v_i(\cdot)\}$ is said to be *bounded* if for some $T > 0$, equivalently, for all $T > 0$,

$$\int_0^T \sum_{i \in E} v_i(t)\, dt < \infty.$$

The Laplace transform:

$$\eta(\lambda) = \int_0^\infty \exp(-\lambda t)\, v(t)\, dt$$

sets up a one-to-one correspondence between the set of bounded entrance solutions $v(\cdot)$ and the set of (row) vector functions $\eta(\cdot)$ such that, for $\lambda, \mu > 0$,

$$\eta(\lambda) - \eta(\mu) = (\mu - \lambda)\, \eta(\lambda)\, \Phi(\mu); \quad \sum_{j \in E} \eta_j(\lambda) < \infty. \tag{29}$$

See, for example, Neveu [*6*; Theorem 2.1.4]. Reuter has shown (Lemma 2.2 of [*8*]) how the solutions of (29) may be constructed. (See Lemma 7.)

Notation. We shall write l for the Banach space of row vectors y on E such that $\|y\|_1 = \sum_{i \in E} |y_i| < \infty$. The scalar product $\sum_{i \in E} y_i x_i$ of a row vector y on E and a column vector x on E will be denoted by $\langle y, x \rangle$. Lastly we introduce the matrix function

$$A(\lambda, \mu) \equiv I + (\lambda - \mu)\, \Phi(\mu) \qquad (\lambda, \mu > 0)$$

which satisfies the equation

$$A(\lambda, \nu) = A(\lambda, \mu)\, A(\mu, \nu) \qquad (\lambda, \mu, \nu > 0).$$

With this notation, (29) becomes:

$$\eta(\mu) = \eta(\lambda)\, A(\lambda, \mu); \quad \eta(\lambda) \in l. \tag{30}$$

Lemma 1. *There exist a non-negative matrix function* $\{M^{ab}(\lambda) : a, b \in A\}; \lambda > 0\}$ *and non-negative vector functions* $\eta^a(\cdot)$ $(a \in A)$ *satisfying conditions* (30) *such that*

$$y_{\cdot}(\lambda) = M(\lambda)\, \eta^{\cdot}(\lambda).$$

Proof. On expanding (26), we obtain

$$(\mu - \lambda)\, [y^a(\lambda)\, \Phi(\mu) + \sum_{c \in A} \langle y^a(\lambda), x^c(\mu) \rangle\, y^c(\mu)] = y^a(\lambda) - y^a(\mu),$$

which may be arranged as:

$$\sum_{c \in A} [\delta^{ac} + (\mu - \lambda) \langle y^a(\lambda), x^c(\mu) \rangle]\, y^c(\mu) = y^a(\lambda)\, A(\lambda, \mu). \tag{31}$$

Hence,

$$y^a(\lambda) = \sum_{c \in A} [\delta^{ac} + (\mu - \lambda) \langle y^a(\lambda), x^c(\mu) \rangle]\, y^c(\mu)\, A(\mu, \lambda). \tag{32}$$

Now $y^a(\mu)A(\mu,\lambda)$ is always non-negative. In fact,

$$y^a(\mu)A(\mu,\lambda) \geqq y^a(\mu) \quad \text{for} \quad \lambda \leqq \mu,$$
$$y^a(\mu)A(\mu,\lambda) = y^a(\mu) - (\lambda-\mu)y^a(\mu)\Phi(\lambda) \geqq$$
$$\geqq y^a(\mu) - (\lambda-\mu)y^a(\mu)R(\lambda) = y^a(\lambda) \quad \text{for} \quad \lambda > \mu.$$

Define

$$\eta^a(\lambda;\mu) = \frac{y^a(\mu)A(\mu,\lambda)}{\|y^a(\mu)A(\mu,1)\|_1} \tag{33}$$

and observe that, by virtue of equation (32), we may write

$$y^\cdot(\lambda) = M(\lambda;\mu)\,\eta^\cdot(\lambda;\mu)$$

where $M(\lambda;\mu)$ is an $A\times A$ matrix whose elements are non-negative if $\lambda<\mu$.

From (27) and (33), it follows that

$$1 \geqq \sum_{j\in E} y_j^a(1) = \sum_{b\in A} M^{ab}(1;\mu), \qquad (\mu>0).$$

By a diagonalization procedure, we may choose a sequence $\{\mu_n\}$ with $\mu_n\to\infty$ such that the limits:

$$M^{ab}(1) = \lim_n M^{ab}(1;\mu_n) \qquad (a,b\in A); \tag{34}$$

$$\eta_j^a(1) = \lim_n \eta_j^a(1;\mu_n) \qquad (a\in A,\ j\in E); \tag{34}$$

exist. Clearly,

$$\sum_{b\in A} M^{ab}(1) = \sum_{j\in E} y_j^a(1), \tag{35}$$

while, by FATOU's Lemma, each $\eta^a(1)$ is a non-negative l-vector whose norm does not exceed unity. But

$$y_j^a(1) = \sum_{b\in A} M^{ab}(1)\,\eta_j^b(1),$$

and, on summing this equation over j and comparing the result with equation (35), we see that

$$\sum_{j\in E}\eta_j^b(1) = 1 \qquad (b\in A). \tag{36}$$

(Note. Strictly speaking, equation (36) follows only for those b with the property that for some a,

$$\limsup_{\mu\to\infty} M^{ab}(1;\mu) \neq 0.$$

This difficulty is easily avoided in the following manner.

For fixed a and c in A, the function

$$\limsup_{\mu\to\infty} M^{ac}(\lambda;\mu)$$
$$= \limsup_{\mu\to\infty}\,(\delta^{ac} + (\mu-\lambda)\langle y^a(\lambda), x^c(\mu)\rangle]\,\|y^c(\mu)A(\mu,1)\|_1$$

is a non-negative, non-increasing function of λ (because $y^a(\lambda)$ is). Hence, we may find a number ν such that, for every pair a and c in A, either

$$\left.\begin{array}{ll}\text{(i)} & \limsup_{\mu\to\infty} M^{ac}(\cdot;\mu) \equiv 0 \quad \text{or} \\ \text{(ii)} & \limsup_{\mu\to\infty} M^{ac}(\nu;\mu) \neq 0 .\end{array}\right\} \tag{37}$$

We now redefine the η's, using ν instead of 1 as the value at which to normalize and adjust $M(\lambda;\mu)$ accordingly. It is easily checked that the *new* $M(\lambda;\mu)$ has the properties described at (37). It does no harm to suppose that 1 is a suitable ν.

If, for some b, $\limsup_{\mu\to\infty} M^{ab}(\cdot;\mu) \equiv 0$ for *every* a, then η^b plays a completely superfluous rôle and we may for the present choose for η^b any non-negative solution of (30) which also satisfies (36). But see the discussion following equation (39).)

From (34) and (36), it follows, by a well-known theorem on l-spaces, that

$$\|\eta^a(1) - \eta^a(1;\mu_n)\|_1 \to 0 \quad \text{as} \quad n\to\infty .$$

Define $\eta^a(\lambda) = \eta^a(1)A(1,\lambda)$. Then, in the strong topology of l,

$$\eta^a(\lambda;\mu_n) = \eta^a(1;\mu_n)A(1,\lambda) \to \eta^a(1)A(1,\lambda) = \eta^a(\lambda)$$

and so $\eta^a(\lambda)$ is non-negative. For fixed λ,

$$y^a(\lambda) = \sum_{b\in A} M^{ab}(\lambda;\mu_n)\,\eta^b(\lambda;\mu_n)$$

and we know that $\|\eta^b(\lambda;\mu_n)\|_1 \to \|\eta^b(\lambda)\| \neq 0$. Hence, in a suitable subsequence of $\{\mu_n\}$, $M^{ab}(\lambda;\mu_n) \to M^{ab}(\lambda)$ and

$$y^a(\lambda) = \sum_{b\in A} M^{ab}(\lambda)\,\eta^b(\lambda)$$

as was to be shown.

3.3. Some identities. Returning to the definition (33) of $\eta^a(\lambda;\mu)$, we see now that

$$\eta^a(\lambda;\mu) = \frac{\sum_{b\in A} M^{ab}(\mu)\,\eta^b(\mu)\,A(\mu,\lambda)}{\|\sum_{c\in A} M^{ac}(\mu)\,\eta^c(\mu)\,A(\mu,1)\|_1} .$$

Since $\eta^a(\mu)A(\mu,\lambda) = \eta^a(\lambda)$ and $\|\eta^a(1)\|_1 = 1$, $(a\in A)$, we may write the above equation as

$$\eta^a(\lambda;\mu) = \sum_{b\in A} H^{ab}(\mu)\,\eta^b(\lambda), \tag{38}$$

where $H(\mu)$ is the stochastic matrix with elements

$$H^{ab}(\mu) = M^{ab}(\mu)/\sum_{c\in A} M^{ac}(\mu). \tag{39}$$

In a suitable subsequence of $\{\mu_n\}$, $H^{ab}(\mu) \to H^{ab}$ where H is a stochastic matrix. It follows from (38) that for this subsequence, which we may as well take to be $\{\mu_n\}$ itself, $\lim_n \eta^a(\lambda;\mu_n)$ exists in the strong topology of l for *every* a in A and every $\lambda > 0$. *We may therefore define* $\eta^a(\lambda) = \lim_n \eta^a(\lambda;\mu_n)$ *for every a and every λ without affecting any of the foregoing results.* Then, from (38),

$$H\eta^{\cdot}(\lambda) = \eta^{\cdot}(\lambda) \quad \text{for every} \quad \lambda > 0 . \tag{40}$$

Many identities in the sequel are most conveniently expressed in terms of the matrix function $\{U^{bc}(\lambda); b, c \in A; \lambda > 0\}$ defined by

$$U^{bc}(\lambda) = \lambda \langle \eta^b(\lambda), x^c \rangle .$$

A similar function was used by FELLER. The importance of $U(\lambda)$ rests on the identity

$$(\lambda - \mu) \langle \eta^b(\lambda), \quad x^c(\mu) \rangle = U^{bc}(\lambda) - U^{bc}(\mu) \tag{41}$$

which is a consequence of equations (11), (12) and (30).

Other properties of $U(\lambda)$ which will be needed are:

$$U(\lambda) \geqq U(\mu) \quad \text{for} \quad \lambda \geqq \mu > 0; \tag{42}$$

$$H\,U(\lambda) = U(\lambda) \quad \text{for} \quad \lambda > 0; \tag{43}$$

$$\sum_{b \in A} \sum_{c \in A} M^{ab}(\mu)\, U^{bc}(\mu) \leqq 1 . \tag{44}$$

The relations (42) and (43) are implied by equations (41) and (40) (respectively). The inequality (44) follows from the norm condition (27) and the fact that $\sum_{c \in A} x^c(\mu) \leqq 1$.

3.4. Redundancy. We may have defined more y's and η's than is good for us. Notice that if the boundary points a_1 and a_2 are *indistinguishable*, i. e. if $y^{a_1}(\lambda) = y^{a_2}(\lambda)$ for some (and then all) $\lambda > 0$, then we may combine a_1 and a_2 into a single exit boundary point a, writing

$$x^a(\lambda) = x^{a_1}(\lambda) + x^{a_2}(\lambda); \quad y^a(\lambda) = y^{a_1}(\lambda) = y^{a_2}(\lambda) .$$

(For a detailed discussion, see CHUNG's paper.) *It therefore does not restrict the generality to assume that all the $y^a(\cdot)$ are distinct.*

Even after indistinguishable boundary points have been merged, there may still be "too many" y's. A method of making a suitable selection from among them will now be described and then the probabilistic significance of our choice will be discussed.

For a moment, let $\lambda > 0$ be fixed. The convex hull of the set $\{0\} \cup \{y^a(\lambda) : a \in A\}$ is a compact convex polyhedron in the space l. The vertices (extremal points) of this polyhedron may be written $\{0\} \cup \{y^{\bar a}(\lambda) : \bar a \in \bar A\}$ where $\bar A$ is a subset of A. It follows from the Krein-Milman Theorem that every $y^a(\lambda)$ $(a \in A)$ may be written:

$$y^a(\lambda) = \sum_{\bar a \in \bar A} G^{a \bar a}\, y^{\bar a}(\lambda) \tag{45}$$

where $G^{a\bar a} \geqq 0$, $\sum_{\bar a \in \bar A} G^{a \bar a} \leqq 1$ $(a \in A)$ and $G^{\bar a \bar a} = 1$ $(\bar a \in \bar A)$.

The matrix G may appear to depend on λ but, on post-multiplying equation (45) by $I - (\mu - \lambda) R(\mu)$, we obtain:

$$y^a(\mu) = \sum_{\bar a \in \bar A} G^{a \bar a}\, y^{\bar a}(\mu) \quad \text{for every} \quad \mu > 0 . \tag{46}$$

(There may however exist other substochastic matrices G for which (46) holds. Throughout the remainder of the paper, G will denote one fixed such matrix.) A similar argument shows that, for each $\mu > 0$, the extremal points of the set $\{0\} \cup \{y^a(\mu) : a \in A\}$ are precisely $\{0\} \cup \{y^{\bar a}(\mu) : \bar a \in \bar A\}$ i. e. $\bar A$ is independent of λ.

Probabilistically, equation (46) may be interpreted as stating that, on reaching $a \in A - \bar{A}$, $\{X(\cdot)\}$ decides to jump immediately to a point of $\bar{A}$, choosing $\bar{a}$ with probability $G^{a\bar{a}}$.

3.5. Interchanging λ and μ in equation (31), we obtain

$$y^a(\mu)\,A(\mu,\lambda) = \sum_{c\in A}[\delta^{ac} + (\lambda-\mu)\langle y^a(\lambda), x^c(\lambda)\rangle]\,y^c(\lambda). \tag{47}$$

Equations (47), (33) and (41) and Lemma 1 lead to the following formula for $\eta^a(\lambda;\mu)$:

$$\eta^a(\lambda;\mu) = \frac{\sum\limits_c \Big[\delta^{ac} + \sum\limits_b M^{ab}(\mu)\,[U^{bc}(\lambda) - U^{bc}(\mu)]\Big] \sum\limits_{\bar{e}} G^{c\bar{e}} y^{\bar{e}}(\lambda)}{\sum\limits_t M^{at}(\mu)}$$

(*Convention: symbols a, b, c, d, t etc. range over A; "barred" symbols such as $\bar{a}$ and $\bar{e}$ range over $\bar{A}$.*) Hence, by definition of $H(\mu)$,

$$\eta^{\bar{a}}(\lambda;\mu) = \sum_b \sum_c H^{\bar{a}b}(\mu)\,U^{bc}(\lambda) \sum_{\bar{e}} G^{c\bar{e}} y^{\bar{e}}(\lambda) + \sum_{\bar{e}} T^{\bar{a}\bar{e}}(\mu)\,y^{\bar{e}}(\lambda), \tag{48}$$

where

$$T^{\bar{a}\bar{e}}(\mu) = \frac{\delta^{\bar{a}\bar{e}} - \sum\limits_b \sum\limits_c M^{\bar{a}b}(\mu)\,U^{bc}(\mu)\,G^{c\bar{e}}}{\sum\limits_t M^{\bar{a}t}(\mu)}.$$

The inequality (44) implies that

$$\sum_b \sum_c \sum_{\bar{e}} M^{\bar{a}b}(\mu)\,U^{bc}(\mu)\,G^{c\bar{e}} \leqq 1. \tag{49}$$

It will now be shown that

(Lemma 2) *if T denotes any subsequential limit of $T(\mu_n)$, then T is finite.*

Proof. We may as well suppose that $T = \lim T(\mu_n)$. Notice that if, for some $\bar{a}$, $T^{\bar{a}\bar{a}} = 0$, then, by (49), $T^{\bar{a}\bar{c}} = 0$ for all $\bar{c}$. If $T^{\bar{a}\bar{a}} \neq 0$, define

$$S^{\bar{a}\bar{c}}(\mu) = -\,T^{\bar{a}\bar{c}}(\mu)/T^{\bar{a}\bar{a}}(\mu) \qquad (\bar{c} \neq \bar{a},\ \mu > 0).$$

Then $\sum\limits_{\bar{c}\neq\bar{a}} S^{\bar{a}\bar{c}}(\mu) \leqq 1$ and hence, in a suitable subsequence of $\{\mu_n\}$, $S^{\bar{a}\bar{c}}(\mu) \to S^{\bar{a}\bar{c}}$ with $\sum\limits_{\bar{c}\neq\bar{a}} S^{\bar{a}\bar{c}} \leqq 1$. If $T^{\bar{a}\bar{a}} = \infty$, we obtain, on dividing (48) through by $T^{\bar{a}\bar{a}}(\mu)$ and letting $\mu \to \infty$ suitably,

$$y^{\bar{a}}(\lambda) = \sum_{\bar{c}\neq\bar{a}} S^{\bar{a}\bar{c}} y^{\bar{c}}(\lambda),$$

contradicting the choice of $y^{\bar{c}}(\cdot)$ as an extremal point. Lemma 2 is therefore proved.

In the limit as $\mu \to \infty$ through the sequence $\{\mu_n\}$, equation (48) becomes

$$\bar{\eta}(\lambda) = [\bar{H}\,U(\lambda)\,G + T]\,\bar{y}(\lambda) = [\bar{U}(\lambda)\,G + T]\,\bar{y}(\lambda) \tag{50}$$

where $\eta(\lambda)$ and $\bar{y}(\lambda)$ denote the restrictions of $\eta^{\cdot}(\lambda)$ and $y^{\cdot}(\lambda)$ to $\bar{A}$ and where $\bar{H}$ and $\bar{U}(\lambda)$ denote the restrictions of H and $U(\lambda)$ to $\bar{A}\times A$.

It it easily checked that each off-diagonal term of the matrix $\bar{U}(\lambda)\,G + T$ is non-positive. Also, it follows from equation (50) that for $\bar{a}$ in $\bar{A}$ and $\lambda > 0$,

$$0 > \lambda\langle\eta^{\bar{a}}(\lambda), 1\rangle = \sum_{\bar{c}} [\bar{U}(\lambda)\,G + T]^{\bar{a}\bar{c}}\,\lambda\langle y^{\bar{c}}(\lambda), 1\rangle. \tag{51}$$

The matrix with $(\bar{a}, \bar{c})^{\text{th}}$ component

$$\lambda[\bar{U}(\lambda)\,G + T]^{\bar{a}\bar{c}}\langle y^{\bar{c}}(\lambda), 1\rangle$$

therefore has a non-negative inverse and so the same is true of the matrix $[\bar{U}(\lambda)\,G + T]$. Hence

(Lemma 3) $$y(\lambda) = G[\bar{U}(\lambda)\,G + T]^{-1}\,\bar{\eta}(\lambda)\,.$$

3.6. The norm condition and positivity. Define

$$\tau^{\bar{a}} = \lambda\langle \eta^{\bar{a}}(\lambda), x^0\rangle \qquad (\bar{a} \in \bar{A})\,,$$

the definition being independent of λ because of equation (13). Substituting $1 = x^0 + \sum_{b\in A} x^b$ in the left hand side of equation (51) and using the norm condition $\lambda\langle y^c(\lambda), 1\rangle \leqq 1$, we obtain

$$\sum_b U^{\bar{a}b}(\lambda)\,[1 - \sum_{\bar{c}} G^{b\bar{c}}] + \tau^{\bar{a}} \leqq \sum_{\bar{c}} T^{\bar{a}\bar{c}}, \qquad (\lambda > 0)\,. \tag{52}$$

By (42), $U(\lambda) \uparrow U(\infty)$ as $\lambda \to \infty$, where the matrix $U(\infty)$ may contain infinite elements. It is clear that (52) is equivalent to the statement:

$$\sum_b U^{\bar{a}b}(\infty)\,[1 - \sum_{\bar{c}} G^{b\bar{c}}] + \tau^{\bar{a}} \leqq \sum_{\bar{c}} T^{\bar{a}\bar{c}}\,. \tag{53}$$

Since, for every $\lambda > 0$, $\bar{U}(\lambda)\,G + T$ is non-positive off diagonal, we also obtain the condition:

$$\infty > -\,T^{\bar{a}\bar{c}} \geqq \sum_b U^{\bar{a}b}(\infty)\,G^{b\bar{c}} \qquad (\bar{a} \neq \bar{c})\,. \tag{54}$$

Conditions (53) and (54) imply that

(Lemma 4) *$U^{\bar{a}b}(\infty)$ is finite if $b \neq \bar{a}$ ($\bar{a} \in \bar{A}$, $b \in A$). Equivalently:*

$$\lim_{\lambda\to\infty} \langle \lambda\,\eta^{\bar{a}}(\lambda), 1 - x^{\bar{a}}\rangle < \infty \qquad (\bar{a} \in \bar{A})\,.$$

Proof. If $U^{\bar{a}b}(\infty) = \infty$ for some $\bar{a}$ in $\bar{A}$ and b in A with $b \neq \bar{a}$, then, from (53), $\sum_{\bar{c}} G^{b\bar{c}} = 1$, while, from (54), $G^{b\bar{c}} = 0$ for $\bar{c} \neq \bar{a}$. Hence

$$y^b(\mu) = y^{\bar{a}}(\mu) \quad \text{for every } \mu\,,$$

a possibility we have already ruled out.

3.7. The general solution. The above relations among the vectors $x^a(\cdot)$, $\eta^a(\cdot)$ and the matrices G and T provide *sufficient* (as well as necessary) criteria that the function $\{P(t)\}$, defined via Lemma 3 and Theorem 1, belongs to $\mathscr{I}_Q$. This is the content of Theorem 2, which may be regarded as a complete, though clumsy, solution to our problem. We shall see in the next subsection that the construction of the strictly stochastic members of $\mathscr{I}_Q$ may be described in somewhat simpler terms.

The introduction of the set C in the statement of Theorem 2 is merely a device to cover the case of indistinguishable boundary points.

Theorem 2. *Let E be a countable set and let Q be an $E \times E$ matrix such that*

$$q_{ij} \geqq 0; \quad -\infty < q_{ii} < 0; \quad \sum_{k\in E} q_{ik} = 0 \qquad (i, j \in E,\; i \neq j)\,.$$

Let $\Phi(\cdot) = \{\varphi_{ij}(\cdot)\}$ denote the resolvent matrix of the minimal transition function F associated with Q. Assume that for some (equivalently, for every) $\lambda > 0$, the space of bounded vectors x on E such that

$$\lambda x = Qx$$

is of finite dimension. The exit boundary A induced by Q then consists of a finite number of points.

Choose any disjoint partition:

$$A = \bigcup_{a \in C} A^a, \quad A^a \bigcap A^c = \emptyset \qquad (a \neq c;\ a, c \in C)$$

of A and, for $a \in C$, let x^a, $x^a(\lambda)$ denote respectively the sojourn solutions of $Qx = 0$, $(\lambda - Q)x = 0$ corresponding to the boundary set A^a. Let $x^0 = 1 - \sum_{a \in C} x^a$ denote the maximal passive element.

Next, choose any subset $\bar{C}$ of C and, for each $\bar{a}$ in $\bar{C}$, choose any non-negative l-valued function $\eta^{\bar{a}}(\cdot)$ which satisfies both

$$\eta^{\bar{a}}(\lambda) - \eta^{\bar{a}}(\mu) = (\mu - \lambda)\,\eta^{\bar{a}}(\lambda)\,\Phi(\mu) \qquad (\lambda, \mu > 0)$$

and

$$\lim_{\lambda \to \infty} \langle \lambda\, \eta^{\bar{a}}(\lambda),\ 1 - x^{\bar{a}} \rangle < \infty .$$

For $\bar{a}$ in $\bar{C}$ and b in C, define

$$U^{\bar{a}b}(\lambda) = \langle \lambda\, \eta^{\bar{a}}(\lambda), x^b \rangle \qquad (\lambda > 0)$$

and

$$\tau^{\bar{a}} = \langle \lambda\, \eta^{\bar{a}}(\lambda), x^0 \rangle .$$

$\tau^{\bar{a}}$ is independent of λ. Let $U^{\bar{a}b}(\infty) = \lim_{\lambda \to \infty} U^{\bar{a}b}(\lambda)$.

Choose a non-negative $C \times \bar{C}$ matrix G such that

$$\sum_{\bar{a} \in \bar{C}} G^{a\bar{a}} \leqq 1 \qquad (a \in C);\ G^{\bar{a}\bar{a}} = 1\ (\bar{a} \in \bar{C}) .$$

Now choose any (finite) matrix T on $\bar{C} \times \bar{C}$ such that

$$-\,T^{\bar{a}\bar{c}} \geqq \sum_{b \in C} U^{\bar{a}b}(\infty)\, G^{b\bar{c}} \qquad (\bar{a} \neq \bar{c})$$

and

$$\sum_{\bar{c} \in \bar{C}} T^{\bar{a}\bar{c}} \geqq \tau^{\bar{a}} + \sum_{b \in C} U^{\bar{a}b}(\infty)\,[1 - \sum_{\bar{c} \in \bar{C}} G^{b\bar{c}}] .$$

Then, for every $\lambda > 0$, the matrix

$$K(\lambda) = G[\bar{U}(\lambda)\, G + T]^{-1}$$

(on $C \times \bar{C}$) exists and is non-negative. Lastly, define

$$r_{ij}(\lambda) = \varphi_{ij}(\lambda) + \sum_{a \in C} \sum_{\bar{c} \in \bar{C}} x_i^a(\lambda)\, K^{a\bar{c}}(\lambda)\, \eta_j^{\bar{c}}(\lambda) .$$

Then $R(\cdot) = \{r_{ij}(\cdot)\}$ is the resolvent of a substochastic transition function $P(\cdot)$ on $E \times E$ with initial derivative matrix Q. Conversely, every substochastic transition function with initial derivative matrix Q may be constructed in the above manner.

$R(\cdot)$ is the resolvent of a strictly stochastic transition function if and only if

$$\sum_{\bar{c} \in \bar{C}} G^{a\bar{c}} = 1 \ \ (a \in C) \quad \textit{and} \quad \sum_{\bar{e} \in \bar{C}} T^{\bar{a}\bar{e}} = \tau^{\bar{a}} \ \ (\bar{a} \in \bar{C}) .$$

Proof of the sufficiency of the conditions. Suppose that the vectors x^a, $x^a(\lambda)$, $\eta^{\bar{a}}(\lambda)$ ($a \in C$, $\bar{a} \in \bar{C}$) and the matrices G and T have been defined in accordance with the conditions of the theorem. Then equation (52) holds for all λ, and hence

$$0 < \langle \lambda \eta^a(\lambda), 1 \rangle \leqq [\bar{U}(\lambda) G + T] 1_{\bar{C}}, \tag{55}$$

$1_{\bar{C}}$ denoting the vector $(1, 1, 1, \ldots)$ on $\bar{C}$. The relations (54) and (55) imply that $\bar{U}(\lambda) G + T$ has a non-negative inverse. On multiplying (55) by $K(\lambda)$ we obtain

$$\langle \lambda y^a(\lambda), 1 \rangle \leqq 1, \qquad (a \in C) \tag{56}$$

the norm condition for $R(\lambda)$. The resolvent equation for $R(\lambda)$ follows from the equation

$$K(\lambda) - K(\mu) = K(\lambda) [\bar{U}(\mu) - \bar{U}(\lambda)] K(\mu).$$

This equation, which is equivalent to the matrix Riccati equation

$$K'(\mu) = - K(\mu) \bar{U}'(\mu) K(\mu),$$

plays a central rôle in all construction problems.

Notes. 1. We have deduced equation (56) from our explicit formula for $K(\lambda)$. In the cases studied by FELLER and CHUNG, one may proceed in the opposite direction, first proving equation (56) from the resolvent equation and then deducing from it the form of $K(\lambda)$. Incidentally, this method yields the shortest proof of FELLER's construction theorem. The reason that the same method does not work here is that our η's may be linearly dependent.

2. Theorem 2 simply rests on the fact that the resolvent $R(\cdot)$ to be constructed is related to a known resolvent $\Phi(\cdot)$ by an equation of the type

$$r_{ij}(\lambda) = \varphi_{ij}(\lambda) + \sum_a x_i^a(\lambda) y_j^a(\lambda)$$

where the non-negative vectors x^a and y^a satisfy respectively the relations (9)—(13) and (26)—(27). There are many problems in Markov chain theory where a similar situation obtains and to which therefore Theorem 2 provides the answer. Suppose, for example, that $\{X(t)\}$ is any Markov chain (with or without instantaneous states) on a countable set E. Suppose that we know the "taboo" transition function $\{{}_D P(t)\}$ which determines the behaviour of $\{X(t)\}$ prior to its first entry into a certain finite subset D of E. To what extent can we reconstruct the transition function $\{P(t)\}$ of $\{X(t)\}$? This problem has already been solved by NEVEU ([*5*], [*6*]). Theorem 2 confirms his solution.

3.8. The strictly stochastic case; detailed construction. Q will continue to denote an $E \times E$ matrix satisfying Assumptions A and B. The purpose of this subsection is to describe, in as direct a manner as possible, the complete procedure for constructing the most general strictly stochastic element of $\mathscr{I}_Q$ from the matrix Q. *No topological considerations will be used in the following formulation and the concept of an "exit boundary" will be replaced by that of a perfectly arbitrary parametrizing set.*

First, let us recall some known results.

Lemma 5 (FELLER [*3*; *Theorem* 4.1]). *For $\lambda > 0$, define the matrix $\Delta(\lambda)$ by the relation*

$$\Delta_{ij}(\lambda) = q_{ij}/(\lambda + q_i)$$

and let

$$S(\lambda) = \sum_{n=0}^{\infty} [\Delta(\lambda)]^n .$$

Then

$$\varphi_{ij}(\lambda) = S_{ij}(\lambda)/(\lambda + q_j) .$$

Lemma 6 (Feller).

$$x^0 = \lim_{\lambda \downarrow 0} \lambda \Phi(\lambda) 1 .$$

This follows from equations (9) and (10).

Lemma 7 (Reuter [*8*; *Lemma* 2.2]). *To construct the most general l-valued function $\eta(\cdot)$ on* $(0, \infty)$ *satisfying*

$$\eta(\mu) = \eta(\lambda) A(\lambda, \mu) \qquad (\lambda, \mu > 0),$$

where

$$A(\lambda, \mu) = I + (\lambda - \mu) \Phi(\mu) \qquad (\lambda, \mu > 0),$$

choose (i) *a non-negative vector w on E such that $w\Phi(\lambda) \in l$ for every $\lambda > 0$;*

(ii) *a positive number ν and a non-negative l-vector $\bar{\eta}$ satisfying $\bar{\eta}(\nu I - Q) = 0$ and define*

$$\eta(\lambda) = w\Phi(\lambda) + \bar{\eta} A(\nu, \lambda) .$$

The set B in the following theorem corresponds to the set $\bar{C}$ in Theorem 2 and the vector z^b to $\sum_a x^a G^{ab}$. When these substitutions have been made, Theorem 3 becomes an immediate corollary of Theorem 2.

Theorem 3. *Let E be a countable set and let Q be an $E \times E$ matrix satisfying Assumptions A and B. Define the minimal resolvent $\Phi(\lambda)$ as in Lemma* 5 *and the vector x^0 as in Lemma* 6.

Let B denote any finite set disjoint from E. Choose non-negative vectors z^b $(b \in B)$ on E such that

$$Q z^b = 0 \;\; (b \in B), \quad \sum_{b \in B} z^b = 1 - x^0 ,$$

and, for $\lambda > 0$, define

$$z^b(\lambda) = z^b - \lambda \Phi(\lambda) z^b \geqq 0 \qquad (b \in B) .$$

For each b in B, choose an l-valued function $\eta^b(\cdot)$ on $(0, \infty)$ *satisfying*

$$\eta^b(\mu) = \eta^b(\lambda) A(\lambda, \mu) \qquad (\lambda, \mu > 0)$$

(see Lemma 7) and also

$$\lim_{\lambda \to \infty} \langle \lambda \eta^b(\lambda), 1 - z^b \rangle < \infty \qquad (b \in B) .$$

Define

$$V^{ab}(\lambda) = \langle \lambda \eta^a(\lambda), z^b \rangle \qquad (a, b \in B),$$

$$V^{ab}(\infty) = \lim_{\lambda \to \infty} V^{ab}(\lambda) \qquad (a, b \in B)$$

and (independently of λ)

$$\tau^a = \langle \lambda \eta^a(\lambda), x^0 \rangle \qquad (a \in B) .$$

Next, choose a matrix T on $B \times B$ such that

$$-T^{ab} \geqq V^{ab} \qquad (a \neq b;\; a, b \in B),$$

$$\sum_{c \in B} T^{ac} = \tau^a \qquad (a \in B) .$$

Then the matrix

$$K(\lambda) = [V(\lambda) + T]^{-1}$$

exists and is non-negative. Now set

$$r_{ij}(\lambda) = \varphi_{ij}(\lambda) + \sum_{a \in B} \sum_{a \in B} z_i^a(\lambda) K^{ab}(\lambda) \eta_j^b(\lambda).$$

Then $R(\cdot) = \{r_{ij}(\cdot)\}$ *is the resolvent of a strictly stochastic transition function* $\{P(t)\}$ *with* $P'(0) = Q$. *Conversely, every such resolvent may be constructed in the above manner.*

Theorem 3 is now in the most convenient form for the type of probabilistic interpretation described in Section 4. I am convinced that the general case of Theorem 2 may be reduced in a similar manner and think it better that the interpretation of Theorem 2 be postponed until this has been done. An analysis of Theorem 2 in its present form would involve complicated calculations very like those needed to prove Theorem 2 in [*10*].

4. Probabilistic interpretation

4.1. A theorem on time substitution. This subsection, which is a slightly modified form of § 1 of the author's paper [*10*], is independent of the preceding results. Various technical conditions (e. g. that $Z(t, \omega)$ be "Borel measurable, well-separable, ...") essential for a rigorous proof of Theorem 4, while of mere academic interest in the present context, are included for the sake of completeness. What is important is that these conditions do not restrict the generality in any real sense because *every* Markov chain with standard transition matrix has a version with the properties assumed of $Z(\cdot, \cdot)$. (For the relevant terminology and for a proof of this result, see CHUNG's book [*1*].)

Suppose that $(\Omega, \mathscr{F}, \mathrm{P})$ is a complete probability triple. Let E be a countable set and let B be a finite set disjoint from E. Suppose that $\{Z(t, \omega) : t > 0, \omega \in \Omega\}$ is a Borel measurable, well-separable, right lower semicontinuous Markov chain with minimal state-space $E + B$ and with standard transition function

$$\{N(t; i, j) : t \geqq 0; i, j \in E + B\}.$$

(The set $E + B$ with the *discrete* topology is assumed compactified by *one-point compactification*, the adjoined point being denoted by "∞". In general, ∞ will belong to the range of $Z(\cdot, \omega)$ for almost every ω.)

The *process* $\{Z^E(\cdot)\}$ *on E induced by* $\{Z(\cdot)\}$ is obtained from the process $\{Z(\cdot)\}$ by ignoring the time spent by the latter in B. Thus $Z^E(\cdot)$ *represents the position of* $Z(\cdot)$ *when the time spent by* $Z(\cdot)$ *in E is exactly t.* (Similarly for $\{Z^B(\cdot)\}$, the *process on B*.) We assume that B is not an absorbing set so that $Z(\cdot)$ will spend an infinite time in E, i. e., $Z^E(\cdot)$ will be everywhere defined. On the other hand, $Z^B(t, \omega)$ will be defined only for $t < \sum_{i \in B} \beta_i(\infty, \omega) \leqq \infty$, so that $\{Z^B(\cdot)\}$ will be a stopped process in the case when the states in B are transient. These ideas will now be made precise.

Let μ denote Lebesgue measure on the real line and, for $t_Q \geqq 0$, $\omega \in \Omega$ and $i \in E + B$, let

$$\beta_i(t, \omega) = \mu\{s : s \leqq t, Z(s, \omega) = i\};$$

in other words, $\beta_i(t)$ *is the time spent by* $\{Z(\cdot)\}$ *in* i *before time* t. *Define*:

$$\tau^E(t,\omega) = \inf\{s : \sum_{i\in E} \beta_i(s,\omega) > t\};$$

$$\begin{aligned} Z^E(t,\omega) &= Z(\tau^E(t,\omega),\omega) \quad \text{if} \quad Z(\tau^E(t,\omega),\omega) \notin B, \\ &= \infty \qquad\qquad\qquad\; \text{if} \quad Z(\tau^E(t,\omega),\omega) \in B; \end{aligned}$$

$$\tau^B(t,\omega) = \inf\{s : \sum_{i\in B} \beta_i(s,\omega) > t\};$$

$$Z^B(t,\omega) = Z(\tau^B(t,\omega),\omega).$$

It is known (see Williams [9; Theorem 1.1]) that both $\{Z^E(\cdot)\}$ and $\{Z^B(\cdot)\}$ are (Borel measurable, well-separable, right lower semicontinuous) Markov chains. Let $\{N^E(t;i,j) : t \geqq 0;\ i,j\in E\}$ and $\{N^B(t;i,j) : t \geqq 0;\ i,j\in B\}$ denote their respective standard transition functions. For $\lambda > 0$, we write $N_\lambda(i,j)$, $N_\lambda^E(i,j)$ for the Laplace transforms:

$$N_\lambda(i,j) = \int_0^\infty \exp(-\lambda t)\, N(t;i,j)\, dt; \quad N_\lambda^E(i,j) = \int_0^\infty \exp(-\lambda t)\, N^E(t;i,j)\, dt$$

of $N(\cdot\,;i,j)$ and $N^E(\cdot\,;i,j)$.

With the transition function $\{N(t;i,j) : t \geqq 0\}$ is associated the strongly continuous semigroup $\{N(t) : t \geqq 0\}$ of operators on the Banach space l_{E+B} of vectors y such that

$$\|y\| = \sum_{i\in E+B} |y(i)| < \infty.$$

The operator $N(t)$ $(t > 0)$ is defined by

$$\{y N(t)\}(j) = \sum_{i\in E+B} y(i)\, N(t;i,j) \qquad (y\in l_{E+B},\ j\in E+B).$$

In a similar fashion, we define the *resolvent* operator $N_\lambda (\lambda > 0)$ by the equation

$$(y N_\lambda)(j) = \sum_{i\in E+B} y(i)\, N_\lambda(i,j) \qquad (y\in l_{E+B},\ j\in E+B).$$

It will be recalled that the *infinitesimal generator* $\mathscr{A}$ of the semigroup $\{N(t) : t \geqq 0\}$ is defined as follows: a vector y in l_{E+B} belongs to the domain $\mathscr{D}(\mathscr{A})$ of $\mathscr{A}$ if and only if there is a vector z in l_{E+B} such that

$$\|\{y N(t) - y\}/t - z\| \to 0 \quad \text{as} \quad t \downarrow 0;$$

and then $y\mathscr{A} = z$. $\mathscr{A}$ is therefore the strong derivative of $N(\cdot)$ at zero. In exactly analogous fashion, we introduce the generator $\mathscr{A}^E$ of the semigroup $\{N^E(t) : t \geqq 0\}$ (on the space l_E) associated with the transition function $\{N^E(t;i,j)\}$. One of the important properties of the operator $\mathscr{A}$ is that *it uniquely determines the transition function* $\{N(t;i,j)\}$. (*Reason*. Each $N_\lambda (\lambda > 0)$ is precisely the operator inverse to $(\lambda - \mathscr{A})$ and so is certainly determined by $\mathscr{A}$. Lerch's theorem on the uniqueness of the determining functions of Laplace transforms completes the proof.)

Our sequence of definitions is completed with the introduction of the "taboo" transition function $\{{}_BN(t;i,j) : t \geqq 0;\ i,j\in E\}$:

$${}_BN(t;i,j) = \mathrm{P}\{Z(t) = j;\ Z(s) \notin B,\ 0 < x < t \mid Z(0) = i\}.$$

$\{{}_BN(t;i,j)\}$ therefore determines the behaviour of $Z(\cdot)$ up to its first entry into B. The Laplace transforms and generator associated with $\{{}_BN(t;i,j)\}$ will be denoted by $\{{}_BN_\lambda(i,j)\}$ and ${}_B\mathscr{A}$ respectively.

The interrelations among the transition matrices described above are summarized in the following theorem.

Theorem 4.

(I). $$N^E_\lambda(\cdot,\cdot) = N_\lambda(\cdot,\cdot) + \sum_{a,\, b \in B}\sum N_\lambda(\cdot, a)\, \Gamma_\lambda(a,b)\, N_\lambda(b,\cdot)$$

on $E \times E$, *where*

$$\Gamma_\lambda^{-1} = \lambda^{-1} - N_\lambda \quad \text{on} \quad B \times B.$$

(II). *For each* $\lambda > 0$ *there exist uniquely defined matrices* L_λ *on* $E \times B$ *and* G_λ *on* $B \times E$ *(first-entrance and last-exit functions for the set* B*) and a matrix* Π_λ *on* $B \times B$ *such that:*

$$N_\lambda(\cdot,\cdot) = \Pi_\lambda(\cdot,\cdot) \quad \text{on} \quad B \times B;$$

$$N_\lambda(\cdot,\cdot) = \sum_{b\in B} \Pi_\lambda(\cdot, b)\, G_\lambda(b,\cdot) \quad \text{on} \quad B \times E;$$

$$N_\lambda(\cdot,\cdot) = \sum_{b\in B} L_\lambda(\cdot, a)\, \Pi_\lambda(a,\cdot) \quad \text{on} \quad E \times B;$$

$$N_\lambda(\cdot,\cdot) = {}_BN_\lambda(\cdot,\cdot) + \sum_{a,\, b\in B}\sum L_\lambda(\cdot, a)\, \Pi_\lambda(a,b)\, G_\lambda(b,\cdot)$$

on $E \times E$.

(III). *In terms of* G_λ *and* L_λ, $N^E_\lambda(\cdot,\cdot)$ *may be expressed:*

$$N^E_\lambda(\cdot,\cdot) = {}_BN_\lambda(\cdot,\cdot) + \sum_{a,\, b\in B}\sum L_\lambda(\cdot, a)\, \Lambda_\lambda(a,b)\, G_\lambda(b,\cdot)$$

on $E \times E$, *where*

$$\Lambda_\lambda = \Pi_\lambda[I - \lambda \Pi_\lambda]^{-1} \quad \text{on} \quad B \times B.$$

(IV). $$\int_0^\infty \exp(-\lambda s)\, \mathrm{P.}\{\tau^B(t) \in ds;\ Z^B(t) = \cdot\} = \exp[t\Psi_\lambda]$$

on $B \times B$, *where*

$$\Psi_\lambda = -\Pi_\lambda^{-1}.$$

Hence $N^B(t) = \exp(Q^B t)$ *where*

$$Q^B = -\lim_{\lambda \downarrow 0} \Pi_\lambda^{-1}.$$

(V). $$\Lambda_\lambda(a,b) = \int_0^\infty \exp(-\lambda t)\, d\mathrm{E}_a \beta_b(\tau^E(t)),$$

$\mathrm{E}_a\{\cdot\}$ *denoting the conditional expectation* $\mathrm{E}\{\cdot \mid Z(0) = a\}$.

(VI). *The generators* $\mathscr{A}$, $\mathscr{A}^E$ *and* ${}_B\mathscr{A}$ *are related as follows:*

(i) $y \in \mathscr{D}(\mathscr{A}^E)$ *if and only if there exists a vector* y^+ *on* $E + B$ *such that*

$$y = y^+ \quad \text{on} \quad E,\ y^+ \in \mathscr{D}(\mathscr{A}) \quad \text{and} \quad (y^+\mathscr{A}) = 0 \quad \text{on} \quad B;$$

then

$$y\mathscr{A}^E = y^+\mathscr{A} \quad \text{on} \quad E;$$

(ii) $y \in \mathscr{D}({}_B\mathscr{A})$ *if and only if the vector* y^0, *with*

$$y^0 = y \text{ on } E \quad \text{and} \quad y^0 = 0 \text{ on } B,$$

belongs to $\mathscr{D}(\mathscr{A})$ *and then*

$$y \cdot {}_B\mathscr{A} = y^0 \cdot \mathscr{A} \quad \text{on} \quad E.$$

Comments and References. For a proof of Part I of Theorem 4 see "An extension to Theorem 3.2" in the author's paper [*9*].

An analytic proof of Part II is given in NEVEU [5; Theorem 4]. As stated by NEVEU, the intuitive content of the result is clear, though, as usual, the heuristic arguments are rather difficult to rigorize. Let us write

$$\alpha_B = \inf\{u : Z(u) \in B\}$$

and $L(s; i, a)$ ($s \geqq 0$, $i \in E$, $a \in B$) for the probability

$$L(s; i, a) = \mathrm{P}_i\{\alpha_B \leqq s;\ Z(\alpha_B) = a\}.$$

Decomposition of the event $\{Z(t) = b\}$ ($b \in B$) according to the time (α_B) and place of first entry into B leads to the formula:

$$N(t; i, b) = \sum_{a \in B} \int_0^t dt(s; i, a)\, N(t-s; a, b) \qquad (i \in E,\ b \in B).$$

The second and fourth equations of Part II involve a decomposition of $N(t; a, j)$ ($a \in B$, $j \in E$) according to the time and place of the *last exit* of $Z(\cdot)$ from B. It may be shown that $G_\lambda(b, j)$ ($b \in B$, $j \in E$) is the Laplace transform of the continuous function $G(t; b, j)$ defined by the equation:

$$G(t; i, j) = \lim_{u \downarrow 0} (1/u)\, \mathrm{P}_b\{Z(t) = j;\ Z(s) \notin B,\ u < s < t\}.$$

The discussion of last-exit decompositions relative to a single state in the Appendix to CHUNG [1] extends with obvious modifications to the case we are now considering and the reader will find there (see especially Theorem 2) a basis for further interpretations of $G(t; b, j)$.

Parts I and II of Theorem 4 imply Part III. The formula for the transition function

$$N^B(t; \cdot, \cdot; \lambda) = \int_0^\infty \exp(-\lambda s)\, \mathrm{P}.\{\tau^B(t) \in ds;\ Z^B(t) = \cdot\}$$

on $B \times B$ is due to NEVEU (Theorem 5 in [5]). It is also possible to calculate the transition function

$$N^E(t; \cdot, \cdot; \lambda) = \int_0^\infty \exp(-\lambda s)\, \mathrm{P}.\{\tau^E(t) \in ds;\ Z^E(t) = \cdot\}$$

on $E \times E$ (see WILLIAMS [9]). From either result, one may deduce Part V.

The formulae for the generators $\mathscr{A}^E$ and ${}_B\mathscr{A}$ (see Lemmas 3.3 and 3.4 in [9]) were included because they provide the most concise and natural expression ot the relations among the various transition functions.

4.2. Application to preceding results.

Theorem 5. *Let the sets E and B be as in Theorem* 3 *and let* $R(\lambda)$ *be the resolvent there constructed. Then there exists a strictly stochastic transition function* $\{N(t; i, j)\}$ *on the set* $E + B$ *such that Theorem* 4 *holds with*

$$N_\lambda^E(i, j) = r_{ij}(\lambda) \quad (i, j \in E), \qquad {}_BN_\lambda(i, j) = \varphi_{ij}(\lambda) \quad (i, j \in E)$$

$$G_\lambda(b, j) = \eta_j^b(\lambda) \quad (b \in B,\ j \in E), \qquad L_\lambda(i, a) = z_i^a(\lambda) \quad (i \in E,\ a \in B)$$

$$\Lambda_\lambda = K(\lambda) \quad , \qquad Q^B = -T.$$

The matrix

$$\Pi_\lambda = [\lambda + V(\lambda) + T]^{-1}$$

on $B \times B$ exists, is non-negative and defines N_λ (and hence $N(t)$) via the equations of Part II *of Theorem* 4.

Particular attention is drawn to the representation of $K(\lambda)$ given by Theorem 4, *Part* V.

The verification of Theorem 5 is straightforward and will be left to the reader.

Comments. The solution to the construction problem afforded by Theorems 3 and 5 certainly has the merit of simplicity and yet it is unsatisfactory in some respects. The chief drawback is that our solution assigns the central probabilistic rôle to an arbitrary well-behaved Markov chain $\{Z(t)\}$ on $E + B$ with transition matrix $\{N(t; i, j)\}$. (In the present context, "well-behaved" means "Borel measurable, wellseparable and right lower semicontinuous".) In a systematic probabilistic treatment of the subject, the main rôle should be taken by an arbitrary well-behaved Markov chain $\{X(t)\}$ on E with transition matrix $\{P(t; i, j)\}$. The distinction is of little importance from a "practical" standpoint because *in calculating any function of $\{X(t)\}$ which depends only on the transition function of $\{X(t)\}$, we may replace $\{X(t)\}$ by a process $\{Z^E(t)\}$*. However, a rigorous proof that *every (well-behaved) $\{X(t)\}$ is of the form $\{Z^E(t)\}$ for a suitable $\{Z(\lambda)\}$* would add considerable insight into the structure of Markov chains. Progress towards such a proof has been made by LÉVY ([*4*]) and NEVEU ([*7*]) but it is not yet possible to define $Z(\cdot, \omega)$ for every ω.

Another possible objection is that the use of non-extremal exit solutions obscures the rôle of the exit boundary as the set of limit points of traversible jump chains. Comparison of Theorem 5 with Theorem 2 of [*10*] inclines the author to the view that what the present approach gains in simplicity more than justifies the sacrifice of some fine probabilistic detail.

References

[*1*] CHUNG, K. L.: Markov chains with stationary transition probabilities. Berlin-Göttingen-Heidelberg: Springer 1960.

[*2*] — On the boundary theory for Markov chains. Acta math. **110**, 19—77 (1963).

[*3*] FELLER, W.: On boundaries and lateral conditions for the Kolmogorov equations. Ann. of Math., II. Ser. **65**, 527—570 (1957).

[*4*] LÉVY, P.: Systèmes markoviens et stationnaires; cas dénombrable. Ann. sci. École norm. sup., III. Sér. **68**, 327—381 (1951).

[*5*] NEVEU, J.: Une généralisation des processus à accroisements positifs indépendants. Abh. math. Sem. Univ. Hamburg **25**, 36—61 (1961).

[*6*] — Lattice methods and submarkovian processes. Proc. Fourth Berkeley Sympos math. Statist Probability **2**, 347—391 (1960).

[*7*] — Sur les états d'entrée et les états fictits d'un processus de Markov. Ann. Inst. Henri Poincaré **17**, 323—337 (1962).

[*8*] REUTER, G. E. H.: Denumerable Markov processes (II). J. London math. Soc. **34**, 81—91 (1959).

[*9*] WILLIAMS, D.: A new method of approximation in Markov chain theory and its application to some problems in the theory of random time substitution. (To appear.)

[*10*] — The process extended to the boundary. Z. Wahrscheinlichkeitstheorie **2**, 332—339 (1963).

Department of Mathematics,
University of Durham,
Great Britain

(Received April 14, 1964)