DAVID WILLIAMS

The Q-matrix problem 3 : the Lévy-kernel problem for chains

Séminaire de probabilités (Strasbourg), tome 12 (1978), p. 310-331

<http://www.numdam.org/item?id=SPS_1978__12__310_0>

THE Q-MATRIX PROBLEM 3: THE LEVY-KERNEL PROBLEM FOR CHAINS

by

David Williams

Part 1 Introduction (including a 'correction' to [QMP2])

(1a) Experts on Markov process theory will probably find this paper, [QMP3], more interesting than either of its predecessors, [QMP1] = [6] and [QMP2] = [7]. Indeed, since chain theory will soon become far too difficult for me, this is a deliberate attempt to persuade process-theory experts to take a more active interest in chains.

This paper is largely independent of its predecessors. Martingale-problem techniques will underlie much of our work, so that we begin a serious effort to tie in the probabilistic side of the theory with the analysis of infinitesimal generators. At the very least, this paper looks more like (say) spin-flip theory than chain theory has tended to do in the past. In particular, we strive to work with strong FELLER, stochastically continuous transition functions on compact metric spaces.

That said, it is perhaps as well to emphasise right away (even before discussing the LEVY kernel problem) some of the rather peculiar features of chains. Let I be a countable set, and let $\{P(t)\}$ be a transition function on I. That $\{P(t)\}$ is "standard" in CHUNG's sense:

$$\lim_{t \downarrow 0} p_{ii}(t) = 1 \qquad (\forall i),$$

is taken to be part of the definition of transition function. Further, we shall deal only with transition functions $\{P(t)\}$ which are honest in the sense that

$$P(t)1 = 1, \ \forall t > 0.$$

(1b) VIA RAY TO FELLER

When it comes to proving that a certain special $\{P(t)\}$ has a strong FELLER, stochastically continuous extension to a given compactification E of I, I find myself unable to do this directly. I first show that the resolvent $\{R(\lambda)\} = \{R_\lambda\}$

of $\{P(t)\}$ acts as a RAY resolvent on E and then show that E has no branch-points. (Because we use the RAY property only as a stepping-stone, we need very little RAY theory here. The very brief Chapters 1 - 3 of GETOOR [2] provide all that we need.)

(1c) 'NATURAL' INFINITESIMAL GENERATOR

A closely related 'difficulty' must also be clarified now. The space $Z = R(\lambda)B(I)$, where $B(I)$ denotes the space of bounded functions on I, is independent of λ, and the closure $\bar{Z}$ of Z is exactly the domain of strong continuity (at $t = 0$) of $\{P(t)\}$ acting on $B(I)$. The <u>strong infinitesimal generator</u> A of $\{P(t)\}$ has domain $R(\lambda)\bar{Z}$ (independently of λ) and, of course

$$(\lambda - A)R(\lambda)f = f, \quad \forall f \in \bar{Z}.$$

If we are to use the martingale-problem method effectively, we must abandon A and instead work with the '<u>natural' infinitesimal generator</u> N defined unambiguously on $\mathfrak{D}(N) = Z$ via the equation

$$(\lambda - N)R(\lambda)f = f, \quad \forall f \in B(I).$$

(The reader can easily check from the resolvent equation and the fact that

$$\lim_{\mu \to \infty} \mu R(\mu)f(i) = f(i) \qquad (\forall i),$$

that the map $R(\lambda) : B(I) \to Z$ is injective for every λ; etc. .)

It is important that <u>if</u> $f \in \mathfrak{D}(N)$ <u>and</u> X <u>is a (nice) chain with transition function</u> $\{P(t)\}$, <u>then</u>

$$f \circ X(t) - \int_0^t Nf \circ X(s)ds$$

<u>is a martingale</u> (for every initial law). See Exposé II of MEYER [4].

(1d) NEVEU's LEMMA

Let us recall one particularly good feature of honest chains. Suppose that E is some given metrizable compactification of I. Then the two properties:

$$P(t) : B(I) \to C(E) \quad (\forall t > 0); \quad R(\lambda) : B(I) \to C(E) \quad (\forall \lambda > 0)$$

are <u>equivalent</u>. This is an immediate consequence of Proposition 2 of NEVEU [5]. It will henceforth be used without further comment.

(1e) KOLMOGOROV BACKWARD EQUATIONS

Let Q be an $I \times I$ matrix satisfying

(DK): $$0 \le q_{ij} < \infty \qquad (\forall i,j : i \neq j)$$

(TIΣ): $$\sum_{k \neq i} q_{ik} = \infty = -q_{ii} \qquad (\forall i).$$

Set

$$(\mathfrak{Q}f)_i \equiv \sum_{k \neq i} q_{ik}(f_k - f_i),$$

with $\mathfrak{D}(\mathfrak{Q})$ consisting of those f in $B(I)$ such that

(i) for each i, the series defining $(\mathfrak{Q}f)_i$ converges absolutely,

(ii) $\mathfrak{Q}f \in B(I)$.

We know from [QMP1,2] that $Q = P'(0)$ for some $\{P(t)\}$ if and only if the condition

(N): $$\sum_{j \notin \{a,b\}} q_{aj} \wedge q_{bj} < \infty$$

holds. The main result of [QMP2] is the following:

THEOREM 1 ([QMP2]) <u>When</u> (DK), (TIΣ) <u>and</u> (N) <u>hold, we can choose</u> $\{P(t)\}$ <u>with natural infinitesimal generator</u> $\mathcal{N}$ <u>satisfying</u> $\mathcal{N} \subset \mathfrak{Q}$.

Theorem 1 is the result actually <u>proved</u> in [QMP2] though the result <u>stated</u> in [QMP2] is the <u>apparently</u> weaker result with $\mathcal{A}$ replacing $\mathcal{N}$.

Let us put [QMP2] right. Introduce the three conditions:

$(KBE)_1$: $\mathcal{A} \subset \mathfrak{Q}$;

$(KBE)_2$: $(\lambda - \mathfrak{Q})R(\lambda)f = f, \ \forall f \in B(I)$;

$(KBE)_3$: $\mathcal{N} \subset \mathfrak{Q}$.

The fact is that <u>these three conditions are equivalent</u>. What is obvious is that

$$(KBE)_2 \iff (KBE)_3 \Rightarrow (KBE)_1 .$$

The proof that $(KBE)_1 \Rightarrow (I \overset{Q}{\rightarrow})$ beginning at the bottom of page 508 in [QMP2] is fallacious but easily corrected. ⟦It is <u>always</u> true that

$$\sum_{i \neq b} q_{bi}[1 - \hat{f}_{ib}(\mu)] \le \mu \hat{g}_b(\mu).$$

Thus

$$\sum_{i \neq b} q_{bi}[\hat{p}_{bb}(\mu) - \hat{p}_{ib}(\mu)] \leq \mu\hat{g}_b(\mu)\hat{p}_{bb}(\mu) \leq 1 - \mu\hat{p}_{bb}(\mu) .$$

This estimate allows us to replace $u = \chi_{\{b\}}$ (which is not necessarily in $\overline{Z}$ - and certainly not in $\overline{Z}$ under hypothesis $(TI\Sigma)!$) by $\mu R(\mu)u$ (which is in Z) and then let $\mu \to \infty$.]] The proof that $(I \overset{Q}{\to}) \Rightarrow (KBE)_2$ in [QMP2] is correct.

The equivalence of $(KBE)_1$, $(KBE)_2$ and $(KBE)_3$ is now established. We simply write (KBE) and say that X (or $\{P(t)\}$) satisfies the KOLMOGOROV backward equation if (KBE) holds.

The probabilistic significance of (KBE) is that when (KBE) holds, each excursion from a point i of I will begin at a point of $I\setminus\{i\}$. Thus (KBE) rules out both continuous exiting from i and jumping from i to a fictitious state.

(1f) THE LEVY-KERNEL PROBLEM

We assume from now on that Q satisfies (DK), $(TI\Sigma)$ and (N). In [QMP2], we proved Theorem 1 by explicit construction of a chain X satisfying (KBE). What makes Theorem 1 and its proof in [QMP2] unsatisfactory is that, however we choose the parameters, the chain X of [QMP2] will make jumps from fictitious states; such jumps are not controlled by Q. For example, topological considerations show that that chain can jump from the bottom of the tree to the top.

What we would like to be able to say is that a chain X can be chosen to satisfy (KBE) and also have the property that

(FLK): Q is the full LEVY kernel of X

in the sense that, almost surely,

$$\forall t, (X_{t-}(\omega) \neq X_t(\omega)) \Rightarrow (X_{t-}(\omega) \in I, X_t(\omega) \in I).$$

The jumps of such a chain will be entirely controlled by Q in the sense of LEVY kernel theory. (Of course, we are light-years away from a situation where the transition function of X is uniquely specified by Q.)

(1g) THE MAIN RESULT

The purpose of this paper is to construct such a chain X and further arrange that its transition function has certain very desirable smoothness properties.

THEOREM 2 We can choose $\{P(t)\}$ satisfying (KBE) such that for some metrizable compactification E of I,

(i) $P(t) : B(I) \to C(E)$, $\forall t > 0$;

(ii) $\{P(t)\}$ is strongly continuous on $C(E)$;

(iii) any (E-valued, HUNT) chain X with transition function $\{P(t)\}$ satisfies (FLK).

Conditions (i) and (ii) imply that the domain $\bar{Z}$ of strong continuity of $\{P(t)\}$ on $B(I)$ is exactly $C(E)$. More significantly, they imply that the strong generator A of $\{P(t)\}$ is exactly DYNKIN's characteristic operator on $C(E)$. (See Theorem 5.5 of DYNKIN [1]). Property (FLK) therefore corresponds to the fact that A is local at points of $E \setminus I$. (See (4ℓ).) The problem of contracting $\mathcal{A}$ to a generator by imposing appropriate boundary and/or lateral conditions is of course the exact analogue of the problem considered by FELLER under the assumption that all states are stable. Provided that we always remember that "the process is the thing", FELLER's ideas act as valuable guides in the present situation.

(1h) The experts on process theory to whom this paper is chiefly addressed will, in the first instance, wish to see general principles rather than merely another complicated construction, even though the construction requires much more cunning this time.

Parts 2 and 3 collect together some 'theoretical' lemmas which outline our basic strategy. On the analytic side, we need the 'perturbation' Lemma 3 which guarantees that, under certain conditions, we can extend the LEVY kernel without destroying the strong FELLER property.

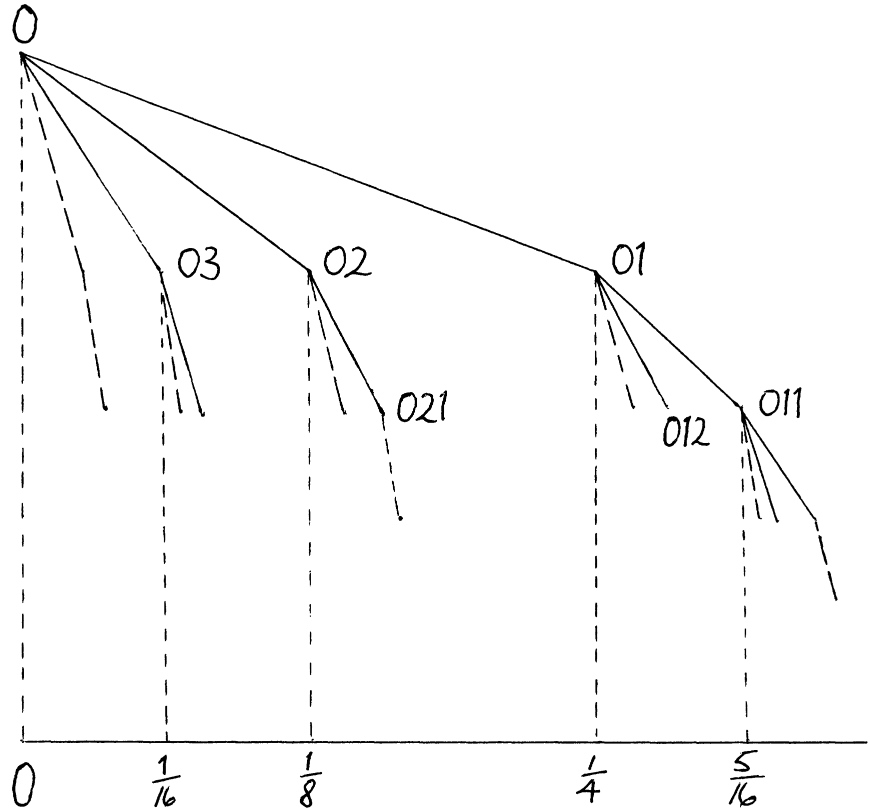

Figure 0

I had anticipated that it would be technically difficult to prove rigorously that a chain constructed by a limiting operation has property (FLK). It therefore surprised me that Lemma 1 provides all that is required.

This paper adds weight to the idea expressed in [8] that chains are essentially one-dimensional because I provides a dense set of points each of which is regular (for itself).

Part 2 A crucial lemma

(2a) TREE-LABELLING

I is said to be tree-labelled if I is labelled as the set of vertices

$I = \{0\} \cup I_1 \cup I_2 \cup \ldots$ (where $I_n = \{0i_1i_2 \ldots i_n : i_1, i_2, \ldots, i_n \in \mathbb{N}\}$)

of the tree shown in Figure 0. Figure 0 also illustrates a certain identification (described in a moment) of I with a certain subset of $[0,1]$.

In Theorem 2, it is always possible to label I as this tree and to take for E the 'compactified' tree $E = I \cup I_\infty$, where

$$I_\infty = \{0i_1i_2 \ldots : i_1, i_2, \ldots \in \mathbb{N}\}.$$

Topology of E. For $i \in I$, take $K(i)$ to be the 'compactified' sub-tree ($K(i) \subset E$) with vertex i in the obvious way. We obtain the topology on E (as in KENDALL [3]) by making the sets $K(i)$ and their complements a sub-base for open sets. Then each $K(i)$ is truly compact (and open).

For us, it is best to think of this compactification in a more down-to-earth way. We shall regard the vertices of the tree as identified with a subset of $[0,1]$ via the identification

(1) $$0i_1i_2 \ldots i_n \longleftrightarrow \sum_{1 \le k \le n} 2^{-(i_1+i_2+\ldots+i_k+k)}$$

which preserves (reverses?!) lexicographic order. Then E is the compact completion of I under the Euclidean metric ρ. The identification (1) transfers ρ to the tree.

Throughout Parts 2 and 3, we shall assume that I is already tree-labelled as described. The symbols E, ρ, $K(i)$ have the significance described above and we write $D(i)$ for the 'finite' subtree with vertex $i : D(i) \equiv I \cap K(i)$. Note that $i \in D(i)$.

Note. The problem of how to label I is difficult. SEYMOUR's labelling is no longer suitable, though SEYMOUR's lemma (Lemma 4 below) is essential as a first step. The required perturbation of SEYMOUR's labelling is very complicated, and it is just as well that, for the time being, the manner in which I is labelled does not concern us.

(2b) A USE OF RAY RESOLVENTS

We now come to the "crucial" lemma. We write χ_F for the characteristic (indicator) function of a set F.

LEMMA 1 <u>Suppose that</u> $\{P(t)\}$ <u>is an honest transition function on</u> I <u>with resolvent</u> $\{R(\lambda)\}$ <u>satisfying</u>

(2) $$R(\lambda) : B(I) \to C(E)$$

<u>Suppose also that for each</u> i, <u>we can find a function</u> f^i <u>in</u> $B(I)$ <u>such that</u>

(3.i) $$R(1)f^i = \chi_{K(i)},$$

(3.ii) $$f^i = 1 \text{ on } D(i)\setminus\{i\}.$$

<u>Then</u>

(I) $\{R(\lambda)\}$ <u>is a RAY resolvent on</u> E;

(II) E <u>has no branch points, so that</u> $\{P(t)\}$ <u>is strongly continuous on</u> $C(E)$;

(III) <u>if</u> X <u>is any</u> (E-<u>valued, HUNT) process with transition function</u> $\{P(t)\}$, <u>then, almost surely</u>,

$$\forall i, \forall t, (X_{t-}(\omega) \in K(i), X_t(\omega) \in E\setminus K(i)) \Rightarrow (X_{t-}(\omega) = i).$$

<u>Proof</u> (I) Let $f^i = f^i_+ - f^i_-$ be the decomposition of f^i into its positive and negative parts. From (3.i) we see that the family

$$\{R(1)f^i_+, R(1)f^i_- : i \in I\}$$

of <u>continuous</u> 1-super-median functions separates points of E. This fact (together with (2)) implies that $\{R(\lambda)\}$ is a RAY resolvent on E.

(II) Let $\{P^*_t\}$ be the RAY transition function on E associated with $\{R(\lambda)\}$ as described in Theorem (3.6) of GETOOR [2]. Since $P^*_0R(1) = R(1)$, it follows from (3.i) that

(4) $$P^*_0\chi_{K(i)} = \chi_{K(i)} \quad (\forall i).$$

It is an immediate consequence of (4) that P^*_0 is the identity operator. Hence E is free from branch points, and for $f \in C(E)$,

$$\lambda R(\lambda)f \to f \qquad (\lambda \uparrow \infty)$$

both in the pointwise sense and, by a well-known argument, in the norm of $C(E)$. The classical HILLE-YOSIDA theorem now implies that

$$P^*_t : C(E) \to C(E)$$

and that $\{P^*_t\}$ is strongly continuous on $C(E)$. Since $P(t) : B(I) \to C(E)$ for

$t > 0$, we can now conclude that $\{P_t^*\} = \{P(t)\}$ in the obvious sense and that (therefore) $\{P(t)\}$ is strongly continuous on $C(E)$.

(III) Let X be an (E-valued, HUNT) chain with transition function $\{P(t)\}$. Let $\xi \in K(i) \setminus \{i\}$, and set

$$U \equiv \inf\{t > 0 : X(t) \in i \cup (E \setminus K(i))\}.$$

By DYNKIN's formula,

$$\begin{aligned} 1 &= \chi_{K(i)}(\xi) = R(1)f^i(\xi) \\ &= E^{\xi} \int_0^U e^{-t} f^i \circ X(t)dt + E^{\xi}[e^{-U}; R(1)f^i \circ X(U)] \\ &= E^{\xi}[1 - e^{-U}] + E^{\xi}[e^{-U}; \chi_{K(i)} \circ X(U)], \end{aligned}$$

from (3). It is therefore obvious that

$$P^{\xi}[U < \infty, X(U) \neq i] = 0.$$

It is now straightforward to finish the proof of (III).

<u>Part 3</u> <u>Establishing the strong FELLER property</u>.

(3a) A 'PRACTICABLE' SUFFICIENT CONDITION

In Part 3, we concentrate on I and E as subsets of $[0,1]$. Statements like "$i > j$" refer to the natural order of $[0,1]$.

If $\{P(t)\}$ is a transition function on I, we write $\{R(\lambda)\}$ or $\{R_\lambda\}$ for the resolvent of $\{P(t)\}$. We write X for an associated chain and T_j $(j \in I)$ for the hitting time of j by X. (Technically, it is best to choose X to be a RAY chain, but as we are interested only in behaviour on I, we can, if we wish, choose X to be (say) a right-lower-semicontinuous chain of the type found in CHUNG's book.)

LEMMA 2 <u>Let</u> $\{P(t)\}$ <u>be a transition function on</u> I <u>such that</u>

(5) $$E^i T_j \leq \rho(i,j) \qquad (\forall i,j \in I : i > j).$$

<u>Then</u>

$$R_\lambda : B(I) \to C(E).$$

<u>Proof</u> For $f \in B(I)$, DYNKIN's formula gives

(6) $$R_\lambda f(i) - R_\lambda f(j) = E^i\int_0^{T_j} e^{-\lambda t} f \circ X(t)dt - [1 - E^i e^{-T_j}]R_\lambda f(j),$$

so that

$$|R_\lambda f(i) - R_\lambda f(j)| \le 2\lambda^{-1}[1 - E^i e^{-\lambda T_j}]\|f\| \le 2\|f\|E^i(T_j).$$

Hence, by (5), $R_\lambda f(\cdot)$ is <u>uniformly</u> continuous on I and so extends to a continuous function on E.

(3b) A PERTURBATION RESULT

LEMMA 3 <u>Let</u> Q <u>be an</u> $I \times I$ <u>matrix satisfying</u> (DK), (TIΣ) <u>and</u> (N). <u>Assume that the conclusion of Theorem 2 is valid for</u> Q <u>with</u> E <u>the given compactification of</u> I, <u>and let</u> $\{P(t)\}$ <u>and</u> X <u>be appropriate entities satisfying this conclusion</u>.

<u>Let</u> $V = (v_{ij})$ <u>be an</u> $I \times I$ <u>matrix satisfying the conditions</u>:

$$0 \le v_{ij} < \infty \qquad (\forall i,j : i \neq j),$$

$$v(i) \equiv -v_{ii} = \sum_{j \neq i} v_{ij} < \infty \qquad (\forall i).$$

<u>Assume that</u>

(7) $$R_\lambda v(0) < \infty \qquad (\forall \lambda > 0)$$

(<u>recall that</u> $0 \in I$) <u>and that</u>

(8) $$E^i\int_0^{T_j} v(X_s)ds \le \rho(i,j) \qquad (\forall i,j : i > j).$$

<u>Then the conclusion of Theorem 2 (with the same</u> E <u>and with appropriate</u> $\{\tilde{P}(t)\}$ <u>and</u> $\tilde{X}$) <u>is valid for</u> $\tilde{Q}$.

<u>Proof</u> We construct $\tilde{X}$ by extending the LEVY kernel of X by V in the usual way. Expand the sample space so that in particular it carries a variable σ_1 (the first 'new' jump time) such that

(9.i) $$P[\sigma_1 > t \mid X] = \exp\left[-\int_0^t v \circ X(s)ds\right].$$

Let $\tilde{X}$ agree with X up to time σ_1, and arrange that

(9.ii) $$P[\tilde{X}(\sigma_1) = j \mid \tilde{X}(\sigma_1 -) = i] = v(i)^{-1}v_{ij} \qquad (j \neq i);$$

and so on. That the right-hand-side of (9.i) makes proper sense is easily checked from (7) and (8), as is implicit in the analysis below. The analysis also implies that the 'new' jump times $\sigma_1, \sigma_2, \ldots$ of $\tilde{X}$ satisfy $\sigma_\infty \equiv \lim \sigma_n = \infty$ (almost surely).

It is clear that we must make precise sense of the formal equation

$$\tilde{A} = A + V. \tag{10}$$

The operator V on $B(I)$ induced by the matrix V is generally unbounded. However, hypotheses (7) and (8) imply that $R_\lambda V$ extends to a <u>bounded</u> operator on $B(I)$ and that

$$\|R_\lambda V\| \to 0 \quad \text{as} \quad \lambda \to \infty. \tag{11}$$

To begin the proof of (11), look at the matrix $R_\lambda V$ and estimate (the notation is self-explanatory)

$$|(R_\lambda V)_{i\ell}| = |\sum_k r_{ik}(\lambda) v_{k\ell}| \le \sum_k r_{ik}(\lambda)|v_{k\ell}|,$$

whence

$$\sum_\ell |(R_\lambda V)_{i\ell}| \le 2\sum_k r_{ik}(\lambda) v_k = 2R_\lambda v(i). \tag{12}$$

From (6) (with $f = v \ge 0$) and (8), we see that

$$R_\lambda v(i) \le R_\lambda v(j) + \rho(i,j) \qquad (\forall i,j : i > j). \tag{13}$$

It now follows from (7), (12), (13) and the well-known DUNFORD-PETTIS result

$$\|R_\lambda V\| = \sup_i \sum_\ell |(R_\lambda V)_{i\ell}|$$

that

$$\|R_\lambda V\| \le 2\sup_i R_\lambda v(i) \le 2R_\lambda v(0) + 2 < \infty \tag{14}$$

By monotone convergence, $R_\lambda v(i) \downarrow 0$ as $\lambda \uparrow \infty$, for each i. There is obviously enough 'equi-uniform-continuity' in property (13) to allow us to deduce that

$$\sup_i R_\lambda v(i) \downarrow 0 \qquad (\lambda \uparrow \infty).$$

Because of (14), <u>property</u> (11) <u>is therefore proved</u>.

Set

$$R_\lambda^{-v} f(i) \equiv \int_0^\infty E^i[\exp\{-\lambda t - \int_0^t v \circ X(s)ds\} f \circ X(t)]\,dt$$

and let $\{\tilde{R}_\lambda\}$ be the resolvent of $\tilde{X}$. If we regard v both as a function on I and as the operation of multiplication by that function, then the FEYNMAN-KAC formula gives (see, for example, Theorem 9.5 of DYNKIN [1])

$$(15) \qquad R_\lambda^{-v} f = R_\lambda f - R_\lambda v R_\lambda^{-v} f .$$

By (14) and FUBINI's theorem, interpretation of $R_\lambda v R_\lambda^{-v} f$ is unambiguous. Next, there is an obvious probabilistic interpretation of the formula

$$(16) \qquad \tilde{R}_\lambda f = R_\lambda^{-v} f + R_\lambda^{-v} V^+ \tilde{R}_\lambda f$$

where V^+ denotes the positive off-diagonal part of V. Again, (14) makes interpretation unambiguous. From (15) and (16),

$$[I + R_\lambda v]\tilde{R}_\lambda = R_\lambda + R_\lambda V^+ \tilde{R}_\lambda ,$$

or equivalently,

$$(17) \qquad \tilde{R}_\lambda - R_\lambda = R_\lambda V \tilde{R}_\lambda ,$$

which is the precise interpretation of the formal equation (10). If λ is so large that $\|R_\lambda V\| < 1$ (see (11)), then equation (17) has the unique solution

$$\tilde{R}_\lambda = R_\lambda + R_\lambda V R_\lambda + R_\lambda V R_\lambda V R_\lambda + \dots .$$

As a minor consequence, we see that (for large λ) $\lambda \tilde{R}_\lambda 1 = 1$, so that $\sigma_\infty = \infty$ almost surely.

A more significant consequence of (11) and (17) is that

$$\|\lambda \tilde{R}_\lambda - \lambda R_\lambda\| \le \|R_\lambda V\| \to 0 \qquad (\lambda \to \infty) .$$

Hence, for $f \in B(I)$, the statements

$$(18.\text{i}) \qquad \lambda R_\lambda f \to f \qquad (\lambda \to \infty)$$

$$(18.\text{ii}) \qquad \lambda \tilde{R}_\lambda f \to f \qquad (\lambda \to \infty)$$

are equivalent. However, since $\{P(t)\}$ is assumed to satisfy the conclusions (i) and (ii) of Theorem 2, property (18.i) is equivalent to the statement

$$(18.\text{iii}) \qquad f \in C(E) .$$

Hence $C(E)$ is the closure of $\tilde{R}_\lambda B(I)$ for each λ, and it is now clear that $\{\tilde{P}(t)\}$ satisfies conclusions (i) and (ii) of Theorem 2.

That $\tilde{X}$ inherits the FLK property from X is obvious.

Part 4 Proof of Theorem 2

(4a) SEYMOUR's LEMMA

We wish to reduce the problem of proving Theorem 2 for a general matrix $\tilde{Q}$ satisfying (DK), (TIΣ) and (N) to that of proving a somewhat stronger result for a special type of matrix Q 'near' to $\tilde{Q}$. The 'strengthening' of Theorem 2 for Q is of course exactly that needed to allow us to transfer Theorem 2 from Q to $\tilde{Q}$ via (a slight extension of) Lemma 3.

LEMMA 4 (P.D. SEYMOUR, [QMP1]). Suppose that $\tilde{Q}$ is an $I \times I$ matrix satisfying (DK), (TIΣ) and (N). Then I may be tree-labelled in such a way that

$$\tilde{Q} = Q + V,$$

where

$$0 \le v_{ij} < \infty \qquad (\forall i,j : i \neq j),$$

$$v(i) \equiv -v_{ii} = \sum_{j \neq i} v_{ij} < \infty \qquad (\forall i),$$

and Q satisfies (DK), (TIΣ), (N) and

($Q \in S\downarrow$): $$q_{ij} > 0 \Rightarrow j \in S(i)$$

where $S(i)$ denotes the set of immediate successors of i in the tree.

(4b) BASIC STRATEGY

We may now consider that our basic 'data' are represented by the following set-up:

(i) a fixed tree-labelling of I for which $S(i)$ denotes the set of immediate successors of i;

(ii) a matrix Q satisfying (DK), (TIΣ), (N) and

($Q \in S\downarrow$): $$q_{ij} > 0 \Rightarrow j \in S(i);$$

(iii) a function $v : I \to [0,\infty)$.

The point is that the perturbation $Q \to \tilde{Q} = Q + V$ will be justified by criteria which depend on V only via the function v. See Lemma 3.

We must construct a chain X which will settle Theorem 2 for Q and also allow the $Q \to \tilde{Q}$ perturbation. As in [QMP2] we shall obtain the desired chain X as a time-projective limit of chains $X_{[n]}$, $X_{[n]}$ being a chain on

$$I_{[n]} \equiv \{0\} \cup I_1 \cup I_2 \cup \ldots \cup I_n \cup I_{n+1}$$

for which states in $I_{[n-1]}$ are instantaneous and states in I_{n+1} are stable.

The time-projective property of the sequence $\{X_{[n]} : n = 0,1,2,\ldots\}$ is the following: $X_{[n]}$ <u>represents</u> "$X_{[n+1]}$ <u>observed only while</u> $X_{[n+1]}$ <u>is in</u> $I_{[n]}$". The martingale-problem method explains why this property translates neatly into the language of infinitesimal generators.

(4c) DIFFICULTIES

We have carefully prepared the topology of I and E, and also the 'combinatorics' of labelling via SEYMOUR's lemma. However, a moment's thought about the totally disconnected character of E will convince the reader that if I <u>is compactified to</u> E <u>via the given labelling for which</u> $(Q \in S\downarrow)$ <u>holds, then no right-continuous</u> E-<u>valued chain with</u> $A \subset \mathfrak{A}$ <u>can satisfy</u> (FLK).

We are forced to <u>relabel</u> I, and this will involve us in some heavy notation. The new labelling will be obtained by making a very slight perturbation of the old labelling. For the time being, we shall use 'stars' to differentiate between the labellings, but when the new labelling is firmly established the stars will be dropped.

From now on, you should from time to time glance at Figure 1. The labelling in Figure 1 is the new labelling. You can see that certain old successors of the new 01 have become new successors of the new 02. (It is obvious from the last sentence that we have to give some thought to problems of notation. I hope that what follows is reasonably clear.)

(4d) (TEMPORARY) NOTATION FOR THE RELABELLING

We shall use i^* to denote the new label for the point originally labelled as i. We shall write ${}^*i$ for the old label of the point which is labelled as i in the new labelling. (Thus ${}^*(i^*) = ({}^*i)^* = i$.) When a function is 'starred', it is to be understood that its argument is to be interpreted in the new labelling.

Let Q^* and v^* be the descriptions of Q and v in the new labelling so that

$$Q^*(i,j) = Q({}^*i,{}^*j)\,,\ v^*(i) = v({}^*i)\,.$$

We must be careful however because sensible usage leads to

$$\rho^*(i,j) = \rho(i,j)\,;$$

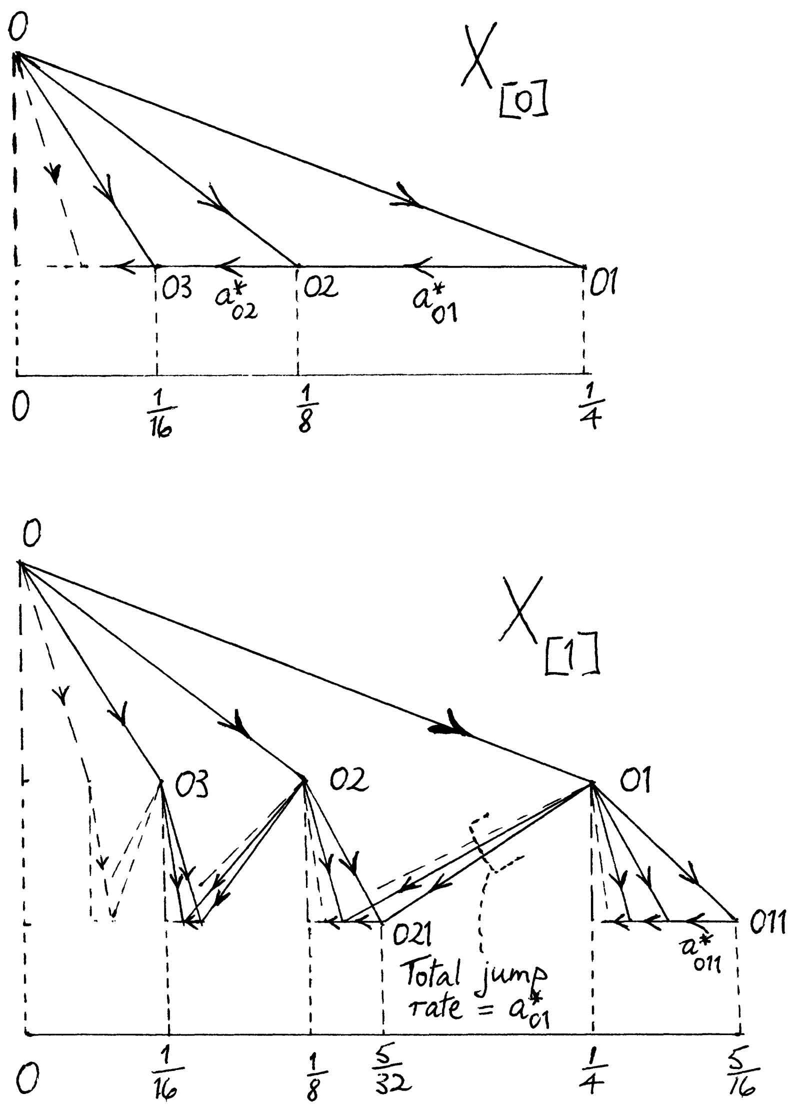

Figure 1.

an equation $\rho^*(i,j) = \rho(^*i, ^*j)$ would be absurd.

Let $S^*(i)$ denote the set of immediate successors in the new labelling of the point with new label i. Since we are going to change the successor relation, it will not be true in general that $S^*(i) = [S(^*i)]^*$.

The new labelling will preserve 0 and will also 'preserve but permute' each level I_n.

(4e) THE MECHANICS OF RELABELLING

The kindest thing to do is to present the reader with the complete new labelling as a fait accompli and explain the motivation afterwards. Figure 1 is a helpful guide to the $(H^*: 4)$ conditions.

It will be convenient (and it will cause no problems of 'circular' arguments) if we introduce the function γ^* on the newly-labelled I as follows:

$$\gamma^*(i) \equiv 1 + v^*(i) = 1 + v({}^*i) .$$

We now proceed via a rather involved inductive method.

Take $0^* = 0$.

$S(0)$ (= the old I_1) is already labelled as $\{01,02,03,\ldots\}$. Let I_1 become $S^*(0)$ and hence be relabelled as $\{01,02,03,\ldots\}$. The only constraint on the new labelling is that

$(H_0^*: 1)$ $$Q^*(0,01) > 0 .$$

Pick a function a^* on $S^*(0)$ such that

$(H_0^*: 2)$ $$(a^*_{0n})^{-1}\gamma^*_{0n} < \rho^*(0n,0\overline{n+1}) ,$$

$(H_0^*: 3)$ $$(a^*_{0n})^{-1}\gamma^*_{0n} \sum_{k\leq n} Q^*(0,0k) < 2^{-(n+2)}$$

and also such that there exists a finite subset $W({}^*0n)$ of $S({}^*0n)$ such that

$(H_0^*: 4)$ $$a^*_{0n} = \sum_{j\in W({}^*0n)} Q({}^*0n,j) .$$

Now define

$$S^*(01) \equiv S({}^*01)\setminus W({}^*01)$$

and for $m \geq 2$, define

$$S^*(0m) \equiv W^*(0\overline{m-1}) \cup [S({}^*0m)\setminus W({}^*0m)] .$$

For each $m \in \mathbb{N}$, perform the following operations: choose a labelling of $S^*(0m)$ as $\{0m1,0m2,\ldots\}$ subject to the sole restraint:

$(H^*_{0m}: 1)$ $$Q^*(0m,0m1) > 0 ;$$

pick a function a^* on $S^*(0m)$ so that

$(H^*_{Om} : 2)$ $$(a^*_{Omn})^{-1}\gamma^*_{Omn} < \rho^*(Omn, Om\,\overline{n+1}) ,$$

$(H^*_{Om} : 3)$ $$(a^*_{Omn})^{-1}\gamma^*_{Omn} \sum_{k \le n} Q^*(Om, Omk) < 2^{-(n+2)} ,$$

and such that there exists a finite subset $W(^*Omn)$ of $S(^*Omn)$ such that

$(H^*_{Om} : 4)$ $$a^*_{Omn} = \sum_{j \in W(^*Omn)} Q(^*Omn, j) .$$

Proceed inductively in the obvious way.

(4f) DROPPING THE STARS

Let us now forget that the old labelling ever existed, drop the stars (including those in Figure 1), and summarise things as they stand in the new notation.

We have the following properties:

$(H_O : 1)$ $$Q(0,01) > 0 ;$$

$(H_O : 2)$ $$a^{-1}_{On}\gamma_{On} < \rho(On, O\overline{n+1}) ;$$

$(H_O : 3)$ $$a^{-1}_{On}\gamma_{On} \sum_{k \le n} Q(0, 0k) < 2^{-(n+2)} ;$$

$(H_O : 4)$ $$a_{On} = \sum_{j \in S(\overline{On+1})} Q(On, j) ;$$

and the condition which replaces the old $(Q \in S\downarrow)$:

$(H_O : 5)$ $$Q(On, j) > 0 \Rightarrow j \in S(On) \cup S(\overline{On+1}) .$$

The general form of $(H_i : k)$ $(i \in I ; 1 \le k \le 5)$ is (I hope!) now obvious.

(4g) THE CHAINS $X_{[n]}$.

(E,ρ) is the compactification of I derived from the new labelling. Let

$$I_{[n]} = \{0\} \cup I_1 \cup I_2 \cup \ldots \cup I_n \cup I_{n+1}$$

("in the new labelling" is now understood). Then $I_{[n]}$ is a compact subset of E. Figure 1 suggests that we choose for $X_{[n]}$ a strong FELLER, stochastically continuous chain on $I_{[n]}$ with natural generator $\mathcal{N}_{[n]}$ defined as follows:

(19.i) $$\mathcal{N}_{[n]} f(i) \equiv \sum_{j \ne i} q_{ij}[f(j) - f(i)] , \quad i \in I_{[n-1]}$$

(19.ii) $$\mathcal{N}_{[n]} f(i) \equiv a_i[f(Oi_1i_2\ldots i_n\overline{i_{n+1}+1}) - f(i)]$$

for $i = Oi_1i_2\ldots i_ni_{n+1} \in I_{n+1}$; the domain $\mathcal{D}(\mathcal{N}_{[n]})$ of $\mathcal{N}_{[n]}$ is exactly the set of those f in $C(I_{[n]})$ such that

(19.iii) for each i in $I_{[n-1]}$, the series at (19.i) converges absolutely,

(19.iv) $$\mathcal{N}_{[n]} f \in B(I_{[n]}) .$$

That the choice of parameters 'guarantees the existence' of such an $X_{[n]}$ is implicit in Sections (4i), (4j). That $\mathcal{N}_{[n]}$ is an <u>extension</u> of the natural generator of the chain $X_{[n]}$ described by the general form of Figure 1 follows from results in [QMP2]. That $\mathcal{N}_{[n]}$ is <u>exactly</u> the natural generator then follows because $\mathcal{N}_{[n]}$ satisfies the <u>minimum principle</u>:

$$\mathcal{N}_{[n]} f \geq 0 \ \underline{\text{at a global minimum of}} \ f .$$

(See Theorem 5.5 of DYNKIN [1].) Of course $\mathcal{A}_{[n]}$ is obtained by restricting the range of $\mathcal{N}_{[n]}$ to fall in $C(I_{[n]})$.

<u>Note the intuitive idea that 'in the limit' as</u> $n \to \infty$, (19.i) <u>takes the form</u> $\mathcal{N} \subset \mathfrak{A}$ <u>while</u> (19.ii-iv) <u>help provide the boundary conditions</u>.

(4h) TIME-PROJECTIVE PROPERTY

Suppose throughout (4h) that for some n, the function f in (19) depends only on the first n coordinates: that is, f is constant on $i \cup S(i)$ for each i in I_n. Then

$$\mathcal{N}_{[n]} f = 0 \ \text{on} \ I_{n+1} ,$$

$$\mathcal{N}_{[n]} f(i) = \sum_{j \neq i} q_{ij} [f(j) - f(i)] \ \text{on} \ I_{[n-2]} ,$$

and for $i = 0 i_1 i_2 \ldots i_n \in I_n$,

$$\mathcal{N}_{[n]} f(i) = a_i [f(0 i_1 i_2 \ldots i_{n-1} \overline{i_n + 1}) - f(i)]$$

because of (H: 5) and the consistency condition (H: 4). Thus, in an obvious sense,

$$\mathcal{N}_{[n]} f = 0 \ \underline{\text{on}} \ I_{n+1} ;$$

$$\mathcal{N}_{[n]} f = \mathcal{N}_{[n-1]} f \ \underline{\text{on}} \ I_{[n-1]} .$$

That these conditions correspond to the time-projective property has long been known. The reader might care to provide himself with a martingale-problem explanation.

That we can arrange the time-projective property in the <u>sample-path</u> sense is perhaps best seen by utilising excursion theory in the manner described in [QMP2]. <u>We assume this done</u>.

(4i) PROJECTIVE LIMIT

We can establish the existence of a projective limit chain X which 'time-projects' onto each $X_{[n]}$ by FREEDMAN's method. The point is that conditions $(H:1)$ and $(H:3)$ (and the fact that $\gamma \geq 1$) imply that X is irreducible, positive recurrent with totally-finite invariant measure μ satisfying

$$\mu(in)/\mu(i) < 2^{-(n+2)}.$$

Similar arguments abound in [QMP1,2].

(4j) USE OF LEMMAS 2 AND 3

For $j \in I$, define

$$\beta_j \equiv E^{01}\int_0^{T_{02}} [\gamma \circ X(s)]\chi_j \circ X(s)ds = \gamma_j\, E^{01}\int_0^{T_{02}} \chi_j \circ X(s)ds.$$

Then

$$\beta_j > 0 \Rightarrow j \in D_{01} \cup (D_{02} \setminus \{02\}).$$

Recall that D_i is the subtree of i with vertex i. If jn denotes a typical element of $S(j)$, then elementary calculations (performed in [QMP1]) show that

$$\beta_{jn} = \beta_j \gamma_j^{-1} \gamma_{jn} a_{jn}^{-1} \sum_{k \leq n} Q(j,jk) \leq \beta_j 2^{-(n+2)}$$

from $(H:3)$ and the fact that $\gamma_j^{-1} \leq 1$. Hence, for j in $D_{01} \cup (D_{02} \setminus \{02\})$,

$$\sum_{k \in S(j)} \beta_k \leq \tfrac{1}{2}\beta_j, \quad \text{whence} \quad \sum_{k \in D(j)} \beta_k \leq 2\beta_j.$$

Thus

$$E^{01}\int_0^{T_{02}} \gamma \circ X(s)ds \leq 2\beta_{01} + 2\sum_{n \in \mathbb{N}} \beta_{02n} \leq 2\gamma_{01}a_{01}^{-1} + 2\sum_{n \in \mathbb{N}} a_{02n}^{-1} \leq 4\rho(01,02)$$

by $(H:2)$. The same argument clearly gives for every i in I,

$$E^{in}\int_0^{T_{i\overline{n+1}}} \gamma \circ X(s)ds \leq 4\rho(in, i\overline{n+1}). \tag{20}$$

Now let $i,j \in I$ with $i > j$ in the $[0,1]$ identification. Then (20) and some obvious additive properties imply that

$$E^{i}\int_0^{T_j} \gamma \circ X(s)ds \leq 4\rho(i,j)$$

if i is not higher than j in the tree. If ($i > j$ in the $[0,1]$ identification and) i is higher than j in the tree, let k be the unique element level with i such that $j \in D(k)$; then

$$E^i\int_0^{T_k}\gamma\circ X(s)ds + E^j\int_0^{T_k}\gamma\circ X(s)ds \le 16\rho(i,j)$$

by simple geometry.

Since $\gamma \ge 1$, the proof of Lemma 2 applies with trivial modifications to show that the resolvent $\{R(\lambda)\}$ of X satisfies

$$R(\lambda) : B(I) \to C(E) .$$

Since $\gamma \ge v$, the proof of Lemma 3 applies with trivial modifications to show that once we have proved Theorem 2 for Q, we can deduce Theorem 2 for $\tilde{Q}$. (Condition (7) is easily verified in the present situation.)

(4k) USE OF LEMMA 1

The resolvent of X has already been shown to satisfy condition (2).

To show that condition (3) is satisfied, we must show that for each i in I there exists f^i in $B(I)$ with

$$R(1)f^i = \chi_{D(i)} \quad \text{on} \quad I: \quad f^i = 1 \quad \text{on} \quad D(i)\setminus\{i\} .$$

Write $\chi^{[n]}_{D(i)}$ for the restriction of $\chi_{D(i)}$ to $I_{[n]}$. Then (see (19)) for $n \ge m$, $\chi^{[n]}_{D(i)} \in \mathcal{D}(\mathcal{N}_{[n]})$ and

$$R_{[n]}(1)f^i_{[n]} = \chi^{[n]}_{D(i)}, \quad \text{where} \quad f^i_{[n]} \equiv (1-\mathcal{N}_{[n]})\chi^{[n]}_{D(i)} .$$

Note that for $n \ge m$,

$$f^i_{[n]} = f^i_{[n+1]} \mid I_{[n]}, \quad f^i_{[n]} = 1 \quad \text{on} \quad I_{[n]} \cap (D(i)\setminus\{i\}) .$$

Set $f^i(j) \equiv \lim_n f^i_{[n]}(j)$ so that $f^i = 1$ on $D(i) \cap \{i\}$.

If we extend $R_{[n]}$ to $B(I)$ via

$$R_{[n]}(1;j,k) \equiv \delta(j,k), \quad (j,k) \in (I\times I)\setminus(I_{[n]}\times I_{[n]}),$$

then

$$R_{[n]}(1)f^i = \chi_{D(i)} \qquad (\forall n) .$$

However, it is standard (and very easy to prove from (32), page 518 of [QMP2] and SCHEFFE's Lemma) that

$$\lim_{n\to\infty}\sum_k | R_{[n]}(1;j;k) - R(1;j;k)| = 0, \ \forall j .$$

Hence condition (3) of Lemma 1 is satisfied:

$$R(1)f^i = \chi_{D(i)} ; \tag{21}$$

that $f^i = 1$ on $D(i)\setminus\{i\}$, we already know. The conclusions of Lemma 1 therefore apply to X.

The argument of [QMP2] shows that X satisfies (KBE). All that remains is to show that X satisfies (FLK). Suppose that (FLK) fails, so that (for some initial law)

$$P\{\exists t : X_{t-}(\omega) \neq X_t(\omega) ;\ (X_{t-}(\omega), X_t(\omega)) \notin I \times I\} > 0 .$$

Then (by (KBE) and obvious topology) for some i,

$$P\{\exists t : X_{t-}(\omega) \in K(i)\setminus I ;\ X_t(\omega) \in E\setminus K(i)\} > 0 ,$$

so that $\exists j \in D(i)\setminus\{i\}$ such that

$$P^j\{X \text{ hits } E\setminus K(i) \text{ before } i\} > 0 .$$

Conclusion (III) of Lemma 2 rules this out.

Theorem 2 is finally proved.

Note that the main point of the approximation argument leading to (21) is to show that $\chi_{D(i)} \in \mathcal{D}(\mathcal{N})$, whence (recall that $\mathcal{N} \subset \mathfrak{A}$)

$$\chi_{D(i)} \circ X(t) - \int_0^t \mathfrak{A}\chi_{D(i)} \circ X(s)ds$$

is a martingale. One can use this fact to rephrase the proof of part (III) of Lemma 1.

(41) THE BOUNDARY CONDITIONS FOR $\mathcal{A}$

The coup de grâce for our chain X is delivered by the very simple exact formula for its strong generator $\mathcal{A}$.

For $f \in \mathcal{D}(\mathcal{A})$ and $i \in I$, we must have

$$(22.\mathrm{i}) \qquad (\mathcal{A}f)(i) = \sum_{j \neq i} q_{ij}[f(j) - f(i)]$$

If $\xi = 0i_1i_2\ldots \in I_\infty$, write $\xi(n) = 0i_1i_2\ldots i_n$. Then, for $f \in \mathcal{D}(\mathcal{A})$, DYNKIN's formula and condition (III) of Lemma 1 imply that

$$(22.\mathrm{ii}) \qquad (\mathcal{A}f)(\xi) = \lim_n b_n(\xi)^{-1}[f \circ \xi - f \circ \xi(n)] ,$$

where $b_n(\xi) \equiv E^{\xi}T_{\xi(n)}$ (which may be explicitly computed without too much trouble).

It is clear that for $f \in \mathcal{D}(\mathcal{A})$,

(22.iii) the series defining $\mathcal{A}f(i)$ converges absolutely for every i, (22.iv) the limit at (22.ii) exists for each ξ in I_∞,

(22.v) $Af \in C(E)$.

Recall that we must have

(22.vi) $\mathfrak{D}(A) \subset C(E)$.

Conditions (22) may be regarded as specifying a certain <u>extension</u> A^+ (say) of A. However, this operator A^+ satisfies the <u>minimum principle</u>. Since $\{P(t)\}$ is <u>FELLER</u> and E is <u>compact</u>, it therefore follows (Theorem 5.5. of DYNKIN [1] again) that $A^+ = A$. Thus

(THEOREM 3) <u>conditions (22) exactly specify the strong generator</u> A (<u>and its domain</u>).

BIBLIOGRAPHY

[1] E.B. DYNKIN. Markov processes (2 volumes), (English translation). Springer 1965.

[2] R.K. GETOOR. Markov processes: Ray processes and right processes. Lecture Notes 440. Springer 1975.

[3] D.G. KENDALL. A totally unstable denumerable Markov process. Quart. J. Math. Oxford 9(34) 1958, pp 149-160.

[4] P.A. MEYER. Démonstration probabiliste de certaines inégalités de Littlewood-Paley. Seminaire de Probabilités X. Lecture Notes 511 1976, pp 125-183.

[5] J. NEVEU Sur les états d'entrée et les états fictifs d'un processus de Markov. Ann. Inst. Henri Poincaré 17 (1962) 323-336.

[6] =[QMP1]. D. WILLIAMS. The Q-matrix problem. Séminaire de Probabilités X. Lecture Notes 511 1976 pp 216-234.

[7] =[QMP2] D. WILLIAMS, The Q-matrix problem, 2: Kolmogorov backward equations. ibid, pp 505-520.

[8] D. WILLIAMS. Some Q-matrix problems. To appear in proceedings of 1976 Amer. Math. Soc. Probability Symposium held at Urbana-Champaign.

Department of Pure Mathematics
University College
Swansea SA2 8PP
Great Britain