SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

DAVID WILLIAMS

The Q-matrix problem

Séminaire de probabilités (Strasbourg), tome 10 (1976), p. 216-234

<http://www.numdam.org/item?id=SPS_1976__10__216_0>

© Springer-Verlag, Berlin Heidelberg New York, 1976, tous droits réservés.

L'accès aux archives du séminaire de probabilités (Strasbourg) (http://portail.mathdoc.fr/SemProba/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Université de Strasbourg
Séminaire de Probabilités

THE Q-MATRIX PROBLEM

by

David Williams

Part 1. Introduction

(a) One object of this paper is to prove a theorem (announced in [20]) which solves the "Q-matrix problem for Markov chains" for the case when all states are instantaneous. The solution for that case is amazingly simple.

That is all very well, but if this paper has any value, it derives from the following considerations. The proof of the 'necessity' part of our theorem is rather 'new-fangled' in that it involves RAY compactifications. By contrast, the proof of the 'sufficiency' part is in the 'bare-hands' spirit of (what many would regard as) the 'good old days' when probabilistic intuition, free from technicalities, guided LEVY's work and the 'analytic' work of FELLER, KENDALL, NEVEU and REUTER. The fact that our theorem is tight (that is: our conditions are both necessary and sufficient) does something to fuse the 'old' and 'new' traditions. It is interesting that we have to go back to the quite astonishing 1958 paper [8] by KENDALL in order to find the necessary escape route from the tyranny of last-exit decompositions modulo a finite set.

It is obvious that the 'tightness' of our theorem means that "nothing can be thrown away at any stage of the argument". However, I respectfully ask the reader to bear this in mind throughout. (Thus, for example, the local character condition, which is an essential feature of the RAY topology, must be reflected in the later 'bare-hands' construction.)

Acknowledgements. P.D SEYMOUR proved the very important combinatorial Lemma 9. I could only prove it under an extra hypothesis.

G.E.H. REUTER's comments and questions on early drafts of this paper have been a great help.

P.A. MEYER and J. NEVEU invited me to talk in France and thereby gave a sense of urgency to my attempts at the 'if' part of the theorem. Otherwise, I would have taken my time about it. (I knew the 'only if' part in 1967 - it is implicit in [16].)

(b) Let I be a countably infinite set. Let $(P(t))$ be a (subMarkovian) transition function (TF) on I with coefficients

$$p_{ij}(t) \equiv P(t;i,j) \equiv P(t;i,\{j\}).$$

(The symbol "$\equiv$" will signify "is <u>defined</u> to be equal to".) We always make the usual assumption that $(P(t))$ is "standard" in CHUNG's sense:

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

The Q-<u>matrix</u>, Q, of $(P(t))$ is the componentwise derivative $Q \equiv P'(0)$. We normally write q_{ij} for $Q(i,j)$ and q_i for $-q_{ii}$. In detail,

$$q_i \equiv -q_{ii} \equiv \lim_{t\downarrow 0} t^{-1}[1-p_{ii}(t)] \leq \infty;$$

$$q_{ij} \equiv \lim_{t\downarrow 0} t^{-1} p_{ij}(t) \qquad (i \neq j).$$

It is well known that the DOOB-KOLMOGOROV limits q_{ij} exist and satisfy the (DK) conditions:

DK(1) $\qquad 0 \leq q_{ij} < \infty \qquad (i \neq j),$

DK(2) $\qquad \sum_{j\neq i} q_{ij} \leq q_i \leq \infty.$

A state i in I is called <u>stable</u> (for Q, $(P(t))$, or the associated chain X) if $q_i < \infty$ and <u>instantaneous</u> if $q_i = \infty$.

We shall concentrate on the case of chains which are <u>totally instantaneous</u>, that is, which satisfy the hypothesis

(TI): $\qquad q_i = \infty \qquad (\forall\, i).$

One detail of notation before we state our theorem: for $i \in I$ and $J \subseteq I\setminus i$, we write

$$Q(i,J) \equiv \sum_{j\in J} q_{ij}.$$

<u>Note</u>. The sense in which $Q(\cdot,\cdot)$ is a <u>restriction</u> of the true LEVY kernel of X will be explained shortly.

THEOREM. <u>Suppose that</u> Q <u>is an</u> $I \times I$ <u>matrix satisfying</u>

(TI): $\qquad q_i \equiv -q_{ii} = \infty \qquad (\forall\, i)$

<u>and</u>

DK(1): $\qquad 0 \leq q_{ij} < \infty \qquad (i \neq j).$

(<u>Condition</u> DK(2) <u>is then automatically satisfied</u>.) <u>Then</u> $Q = P'(0)$ <u>for</u>

some subMarkovian TF(P(t)) if and only if the following two conditions (N) and (S) hold:

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

(S): there exists an infinite subset K of I such that

$$Q(i,K\backslash i) < \infty \qquad (\forall\, i).$$

Further, (P(t)) may be chosen to be "honest" in the sense that

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

Notes. (i) For us, the old term "honest" is much better than "(strictly) Markovian". We wish to be able to say that "X is honest" (which amounts to saying that X has almost surely infinite lifetime). "Conservative" has a special meaning in chain theory (see Proposition 8).

(ii) The "N" is in deference to NEVEU in whose work condition (N) is implicit. The "S" stands for "safety factor": there has to be a big set K which is fairly safe from hits.

(c) This section (c) of the Introduction can be skipped on a first reading. It summarises in coded form the present state of knowledge about "Q-matrix problems". Here is the key to the code:

($\mathscr{E}$): existence; ($\mathscr{U}$): uniqueness; ($\mathscr{H}$): honest;

(TI): totally instantaneous; (TS): totally stable.

Thus, for example, Problem ($\mathscr{E}\mathscr{U}\mathscr{H}$|TS) reads: "find a necessary and sufficient condition on an I × I matrix Q satisfying

(TS): $$q_i < \infty \qquad (\forall i)$$

for there to exist ($\mathscr{E}$) a unique ($\mathscr{U}$) honest ($\mathscr{H}$) transition function with Q-matrix Q".

($\mathscr{E}$|TS), ($\mathscr{E}\mathscr{U}\mathscr{H}$|TS) were solved long ago by FELLER.

($\mathscr{E}\mathscr{H}$|TS), ($\mathscr{E}\mathscr{U}$|TS) were recently solved in HOU [6]. REUTER [14] gives a much better proof of HOU's results.

($\mathscr{E}$|TI), ($\mathscr{E}\mathscr{H}$|TI) are solved by our theorem.

($\mathscr{E}\mathscr{U}$|TI), ($\mathscr{E}\mathscr{U}\mathscr{H}$|TI) are nonsensical because there is no possibility of uniqueness under hypothesis (TI).

The ($\mathscr{E}\mathscr{U}$) problem is completely solved because it reduces to the ($\mathscr{E}\mathscr{U}$|TS) problem solved by HOU. I have 'effectively' solved the ($\mathscr{E}\mathscr{U}\mathscr{H}$) problem and shall discuss it and the general ($\mathscr{E}$) problem elsewhere. The nice ($\mathscr{E}$) problem under the assumption that X is purely discontinuous and Q is the full LEVY kernel of X will also be treated.

Part 2. Proof of the 'only if' part of the theorem

MARKOV CHAINS AS RAY PROCESSES

Let $(P(t))$ be a ("standard") strictly honest transition function on the countable set I. Let U^{α} be the resolvent of $(P(t))$. Let $\mathcal{S}$ be the smallest $\wedge$-stable cone, stable under the kernels U^{α}, containing the constant function 1 and the functions $U^{\alpha}(\cdot,\{j\})$. Then $\mathcal{S}$ separates points of I. Let $\bar{E}$ be the compactification of I relative to $\mathcal{S}$. Then $(P(t))$ extends to a RAY transition function on 'the' RAY-KNIGHT (RK) compactification $\bar{E}$ of I relative to $(P(t))$. Let

$$E \equiv \{x \in \bar{E} : P(t;x,I) = 1, \ \forall t > 0\},$$

and let D be the set of non-branch points in E. Then D is a Polish space containing I.

The space E is the RAY-NEVEU-KNIGHT-DOOB- ... state-space of a right process:

$$X = (X_t, \theta_t, \mathcal{F}_t^0, \mathcal{F}_t, \ldots, P^x : x \in E)$$

with transition function $(P(t))$ such that

1(i) $$X_{t+}(\omega) = X_t(\omega) \in D \qquad (\forall t \geq 0, \ \forall \omega),$$

1(ii) $$X_{t-}(\omega) \text{ exists in } E \qquad (\forall t > 0, \ \forall \omega).$$

The fundamentals of the theory of RAY processes and right processes have now attained their definitive form in GETOOR [4] and GETOOR-SHARPE [5]. Brief indication of the equivalence of RAY-KNIGHT, NEVEU [13] and DOOB [2] E-spaces for chains is given in my paper [17].

Definition. Points of $D \setminus I$ are called fictitious states of X.

[[Two confessions. (i) There is an obvious error in the definition of V in the discussion of R-K compactification on page 298 of my expository paper [19]. (ii) In my earlier paper [17], the definition of the RAY-NEVEU (= RK) topology is correct but my conjecture about its probabilistic significance is little short of idiotic. LAMB and ARCHINARD quickly gave counter-examples.]]

Q-MATRICES AND EXCURSIONS

X and $(P(t))$ have the same meanings as above.

The probabilistic significance of $Q \equiv P'(0)$ is perhaps best explained as follows:

2(i) $$q_i = \nu_i(V_i),$$

2(ii) $$q_{ij} = \nu_i\{w_i \in V_i : w_i(0+) = j\} \qquad (j \neq i).$$

Here V_i is the space of excursions w_i from i. A typical element w_i of V_i is a map

$$w_i : (0, \zeta_i(w_i)) \to D \setminus \{i\}$$

from a 'random' interval $(0,\zeta_i(w_i))$ to $D\setminus\{i\}$ which satisfies conditions analogous to those at (1):

3(i) $w_i(t+) = w_i(t) \qquad (0 < t < \zeta_i(w_i) \le \infty)$,

3(ii) $w_i(0+)$ exists in D and $w_i(t-)$ exists in E for $0 < t < \zeta_i(w_i)$.

We define the σ-cylinder algebra $\mathcal{V}_i$ on V_i as usual. ITO ([7]) showed that in terms of the local time $L(\cdot,i)$ at i:

(4) $$L(t,i) \equiv \text{meas}\{s \le t : X_s = i\},$$

the excursions from i form a 'Poisson' point process with values in $(V_i,\mathcal{V}_i)$. He defined the ITO excursion law ν_i at i as the 'characteristic measure' of this point process. MAISONNEUVE [9] has a fine treatment of excursions and of the important related concept of incursions.

Equation 2(ii) is equivalent to FREEDMAN's result ([3]) on 'pseudo-jumps'.

(5) [[EXAMPLE (FELLER-MCKEAN-LEVY). Let I be the set of rational numbers. Let m be a probability measure on I with $m(i) > 0$, $\forall i$. Let B be Brownian motion on $\mathbb{R}$ and let ℓ be its jointly continuous local time Put

$$\gamma(t) \equiv \sum_{i\in I} \ell(t,i)m(i), \qquad X_t \equiv B\circ\gamma^{-1}(t).$$

Then X is a Markov chain on I and, from (2),

$$q_i = \infty \quad (\forall i), \qquad q_{ij} = 0 \quad (\forall i,j : i \neq j).]]$$

The interpretation (2) allows us to derive some simple necessary conditions for a matrix to be a Q-matrix. Here is an example. For $i \neq j$,

$$\nu_i\{w_i : w_i(s) = j, \forall s \le t\} = q_{ij}e^{-q_jt}.$$

Hence

$$\nu_i\{\zeta_i > t\} \ge \sum_{j\neq i} q_{ij}e^{-q_jt}.$$

But $\nu_i\circ\zeta_i^{-1}$ is the classical LEVY-BLUMENTHAL-GETOOR measure associated with the local time $L(i)$ at the regular state i of X. Hence

(6) $$\sum_{j\neq i} q_{ij}(q_j+1)^{-1} \le \int_0^\infty e^{-t}\nu_i\{\zeta > t\}dt < \infty.$$

Results like (6) explain why it is natural to impose hypothesis (TI) in our theorem.

Every result which may be obtained by the type of argument just used is an immediate consequence of NEVEU's analytic work on entrance laws and excursion laws. Result (N) is a more subtle consequence of NEVEU's theory. We shall soon see why this result reflects the Hausdorff property of the RK topology.

THE LEVY SYSTEM OF X

The theory of LEVY kernels provides a simple probabilistic interpretation of the off-diagonal elements q_{jk} $(j \neq k)$ of Q.

For $j, k \in I$ with $j \neq k$, define

$$J_t(j,k) \equiv \#\{s \leq t : X_{s-} = j, X_s = k\},$$

so that $J_t(j,k)$ is the number of jumps from j to k during time t. Then for $i \in I$,

$$E^i J_t(j,k) = \int_0^t p_{ij}(s)\, q_{jk}\, ds .$$

Important discussion. The full LEVY kernel N of X is defined on $E \times D$. (See BENVENISTE-JACOD [1] for the latest and best account of LEVY systems.) It is extremely important that the kernel N carries much more information than the Q-matrix which is the restriction of N to $I \times I$. The way in which (in general) we force

$$N(i, D \setminus I) > 0$$

in our construction for the 'if' part of the theorem should be especially noted.

THE 'LOCAL CHARACTER' CONDITION

The following simple lemma, which is reminiscent of local-character conditions for establishing continuity or right-continuity of paths, is the key to the application of the RK topology to the study of Q-matrices.

(7) LEMMA. Let G be an open subset of E and let $h \in G \cap I$. Put $G^c \equiv E \setminus G$. Then

$$Q(h, G^c \cap I) < \infty.$$

Note. The analogous result for the LEVY kernel N would be more natural.

Proof. Since G is open, the equations

$$X_t^G \equiv X_t \qquad (t < T_{G^c} \equiv \inf\{s > 0 : X_s \in G^c\}),$$

$$\equiv \partial \qquad (t \geq T_{G^c}),$$

define a "standard" Markov chain X^G on $(G \cap I) \cup \partial$.

On applying DK(1) to the Q-matrix of X^G, we obtain

$$\infty > \lim_{t \downarrow 0} t^{-1} P^h[X_t^G = \partial] \geq \liminf_{t \downarrow 0} t^{-1} P(t; h, G^c \cap I)$$

$$\geq Q(h, G^c \cap I).$$

Proof of (N). Let a,b be distinct points of I. By the Hausdorff property for E, there exist disjoint open subsets G_a, G_b of E with $a \in G_a$, $b \in G_b$. But then

$$\sum_{j \notin \{a,b\}} q_{aj} \wedge q_{bj} \leq Q(a, G_a^c) + Q(b, G_b^c) < \infty .$$

(8) PROPOSITION. Suppose that H is a finite subset of I such that

8(i) $$\liminf_{j} \sum_{h \in H} q_{hj} > 0 .$$

Then every i in $I\setminus H$ is both stable and 'conservative' in the sense that

$$\sum_{j \neq i} q_{ij} = q_i < \infty \qquad (\forall\ i \in I \setminus H).$$

Proof. It is clearly enough to consider the case when H is minimal subject to the requirement 8(i). Then every state in H is instantaneous and, by right continuity of paths, is an accumulation point of I. Let G be an open subset of E which contains H. Then Lemma 7 and hypothesis 8(i) imply that G^c contains only finitely many points of I. Thus I is homeomorphic to the disjoint union of $|H|$ copies of $\mathbb{N} \cup \{\infty\}$ and is already compact: $I = \bar{E}$. In particular, X takes all its values in I.

Now fix i in $I\setminus H$. Then i is isolated in the RK topology. By right continuity of paths, i is stable. An excursion path w_i from i must satisfy $w_i(0+) \in I$ because $I = \bar{E}$. Further $w_i(0+) \neq i$ (for obvious topological reasons and for the probabilistic reason that $w_i(0+) = i$ would contradict the strong Markov property). Hence $w_i(0+) \in I\setminus\{i\}$ and

$$q_i = \nu_i(V_i) = \sum_{j\neq i} \nu_i \{w_i : w_i(0+) = j\} = \sum_{j\neq i} q_{ij}.$$

In other words, i is conservative.

Notes. (i) One can adapt NEVEU's methods to prove that 8(i) implies that each i in $I\setminus H$ is stable. However, I can not give any 'analytic' proof that each i in $I\setminus H$ is conservative.

(ii) A number of 'analytic' results about Markov chains can be obtained by similar use of the RK topology. Among these are the result quoted at the very end of my paper [18] on Markov groups and Theorems 1, 2 and 3 of REUTER-RILEY [15].

(iii) The above proof of Proposition 8 modernises that given in my 1967 paper [16].

Proof of (S). Now assume that (TI) holds. Then by Proposition 8, the following statement is true: (S^*) : for every finite subset H of I,

$$\liminf_{j} \sum_{h\in H} q_{hj} = 0.$$

Condition (S^*) is easily shown to be equivalent to condition (S). For imagine I labelled as the set $\mathbb{N}$ of natural numbers. Then condition (S^*) implies that there is an infinite set

$$K = \{k(1), k(2), k(3), \ldots\} \subseteq I$$

such that

$$\sum_{i\leq n} q_{i,k(n)} < 2^{-n} \qquad (\forall n).$$

Then

(S): $$Q(i, K\setminus i) < \infty \qquad (\forall n).$$

That $(S) \Rightarrow (S^*)$ is trivial.

Notes. (i) We have now proved the 'only if' part of our theorem under the assumption that $(P(t))$ is honest. But all the more is it true if $(P(t))$ is not honest!

(ii) The full strength of Proposition 8 was not needed for proving (S) but will become important during proof of the 'if' part of the theorem. See the remark following (10) below.

Part 3. Sketched proof of the 'if' part of the theorem.

Proof of the 'if' part of the theorem is much more difficult, both for technical reasons and for reasons of greater substance. I give here the gist (but not the technical details) of a proof based on "KENDALL's branching procedure". To be honest, I have to say that my proof of the more technical parts (I shall indicate these below) is too clumsy to inflict upon readers.

I hope to publish later in Proc. London Math Soc. either a complete and tidy account of the entire branching-procedure proof or a totally different proof based on diffusion theory. My first attempt at a proof via diffusion theory failed but I am determined to have another shot at it after Professor MEYER independently suggested the diffusion approach.

As I see things now, the combinatorial Lemma 9, for which the present branching-procedure proof provides the motivation, will have to be used in any diffusion approach. Lemma 9 provides the only way that I can see of picking out the 'correct' RAY topology from Q. The hope that a 'diffusion' proof will work rests in part upon the theorem that every countable metric space without isolated points is homeomorphic to the rationals. Lemma 9 guarantees that we can imbed I in $\underset{\sim}{R}$ in a manner consistent with the local character condition.

We now suppose that I is a countably infinite set and that Q is an $I \times I$ matrix satisfying conditions (TI), DK(1), (N) and (S) of our theorem. We construct explicitly an honest chain X with Q-matrix Q.

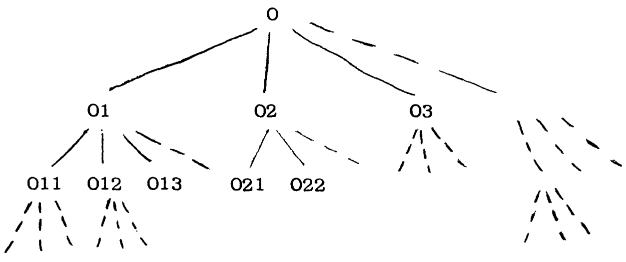
Though a desire for clarity has persuaded me to separate out the statement of the combinatorial Lemma 9, I have presented the remainder of the proof in the order in which I thought it through. I apologise to those readers who would have preferred the 'systematic' approach, but I never did like those elementary analysis books which begin proofs: "Suppose $\varepsilon > 0$ and choose

$$\delta < \varepsilon(163\pi^{\frac{1}{4}} + \hbar e/c^2)^{-137} \ldots" .$$

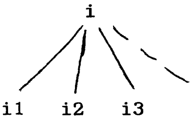
A COMBINATORIAL LEMMA

The motivation for Lemma 9 will become clear in the next section.

We shall say that I is tree-labelled if I is labelled as the set of vertices of the infinitely ramified tree:



We shall then write $Z(i)$ for the set of <u>immediate successors</u> of i so that we have the following local picture of $i \cup Z(i)$:



$Z(i)$

(9) LEMMA (P.D. SEYMOUR). I <u>may be tree-labelled in such a way that for every</u> i <u>in</u> I,

9(i): <u>there exists an infinite subset</u> $K(i)$ <u>of</u> $Z(i)$ <u>such that</u>

$$Q(i,K(i)) < \infty ;$$

9(ii)

$$c(i) \equiv \sum_{j \neq i} [q_{ij} - \bar{q}_{ij}] < \infty ,$$

<u>where</u>

$$\bar{q}_{ij} \equiv q_{ij} \text{ if } j \in i \cup Z(i) ,$$

$$\equiv 0 \text{ otherwise.}$$

The point of 9(i) is that the safety set K of hypothesis S is big enough to allow <u>local</u> safety. Note that the only "I to I jumps" permitted under $\bar{Q}$ are from a state to an immediate successor. Condition 9(ii) states that the pair (Q,I) may be "approximately tree-ordered": Q differs only "finitely" from $\bar{Q}$. The function $c(\cdot)$ on I should be remembered as a <u>correction</u> term.

SEYMOUR's proof of Lemma 9 is deferred until Part 4(b) so as not to interrupt the probabilistic construction. <u>We therefore assume now that</u> I <u>is already tree-labelled in accordance with Lemma 9</u>. We write I_n for the n-th level of the tree so that

$$I_0 \equiv \{0\} , \quad I_1 = \{01,02,03,\dots\} ,$$

and generally for $n \geq 1$,

$$I_n = \{0m_1m_2 \dots m_n : m_k \in \underset{\sim}{N}, \ \forall k \leq n\} .$$

THE BASIC IDEA: KENDALL'S BRANCHING PROCEDURE

For each i in I we shall later construct an honest chain $X^{(i)}$ (of a certain special "LEVY-FELLER (LF)" type) having minimal state-space $i \cup Z(i)$, and with Q-matrix $Q^{(i)}$ satisfying

10(i) $\quad q_i^{(i)} = \infty, \quad q_{ij}^{(i)} = q_{ij} \qquad (j \in Z(i))$,

10(ii) $\quad q_j^{(i)} < \infty, \quad q_{jk}^{(i)} = 0 \qquad (j \in Z(i), k \in [i \cup Z(i)] \setminus j)$.

Note that by Proposition 8, condition 9(i) is necessary for the existence of an honest chain with Q-matrix $Q^{(i)}$ satisfying conditions (10). It follows from arguments similar to those used to prove Proposition 8 that $X^{(i)}$ is necessarily a very complicated chain with infinitely many fictitious states. But let us not look for difficulties.

Let us

(i) assume that "suitable" $X^{(i)}$ exist,

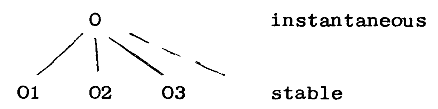
(ii) see how to piece them together via KENDALL's branching procedure,

(iii) collect together all the properties which the various $X^{(i)}$ must have in order that the branching procedure will work,

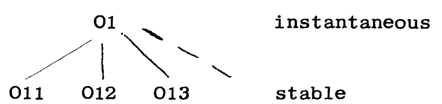
(iv) finally prove that suitable $X^{(i)}$ do exist.

By KENDALL's branching procedure we mean the ingenious probabilistic idea which motivated the analysis in KENDALL's 1958 paper [8]. That remarkable paper should be compulsory reading for all who are interested in LEVY systems and their relation to infinitesimal generators. The paper has a nice RAY-KNIGHT compactification too!

The chains $X^{(i)}$ are our basic building-blocks. We piece them together as instructed by KENDALL. Begin with $X^{(0)}$ which has minimal state-space

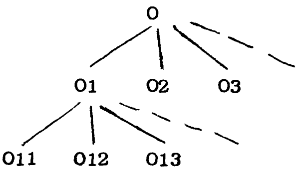


(Important: we shall start $X^{(0)}$ at 0, so we always work with the $P^{(0)}$ law of $X^{(0)}$.) Each visit by $X^{(0)}$ to its stable state 01 will be exponentially distributed with some rate a_{01}. During each such visit to 01, replace $X^{(0)}$ by a chain with minimal state-space



and with the $P^{(01)}$ law of $X^{(01)}$. (In effect, each of these $P^{(01)}$ chains is killed at rate a_{01} but it is essential that for each ω the "lives of the

$P^{(01)}$ chains are exactly the visits by $X^{(0)}$ to 01. This is a matter of paths, not merely of laws.) The resulting 'wedge' or 'join' of $X^{(0)}$ and $X^{(01)}$ is a Markov chain with minimal state-space



of which states 0 and 01 (and only these) are instantaneous. (I take the Markov property just asserted as intuitively obvious. It is!) Note that the wedging operation has "transferred a certain proportion of the time originally spent at 01 down to the set $Z(01)$".

We now apply the obvious induction to fill out the entire tree I. We must ensure however that (almost) no time is pushed right off the end of the tree. (This statement is formulated precisely in the section below entitled "Invariant Measures: filling out the time".) Then the resulting limit process X^- should be a Markov chain with minimal state-space I and with Q-matrix Q^-. Of course, equations (10) arrange that the various $Q^{(i)}$ fit together to produce Q^-.

INCREASING THE LEVY SYSTEM

Before we begin an attempt to make the branching procedure rigorous, we must take account of the final step in the proof of the theorem. This involves "increasing the LEVY kernel" so as to produce from X^- (which has Q-matrix Q^-) a chain X with Q-matrix Q.

Define

$$\varphi(t) \equiv \int_0^t c \circ X_s^- ds \tag{11}$$

where $c(\cdot)$ is the correction function at 9(ii). We shall have to choose the $X^{(i)}$ so as to guarantee that φ is a (finite-valued) CAF of X^-. Define a new process $\tilde{X}$ which agrees with X^- up to the time σ_1 of the first "new" jump of $\tilde{X}$, where

$$P[\sigma_1 > t \mid X^-] = \exp[-\varphi(t)]$$

$$P[\tilde{X}(\sigma_1) = j \mid \tilde{X}(\sigma_1 -) = i] = c(i)^{-1}[q_{ij} - q^-_{ij}].$$

Introduce further "new" jumps $\sigma_2, \sigma_3, \ldots$ in the obvious way. Then $\tilde{X}$, defined for $t < \sigma_\infty$ ($\sigma_\infty = \lim \sigma_n$), will be a Markov chain with Q-matrix Q.

Provided that we do not wish to insist that X is honest, we are home: take $X = \tilde{X}$. The only way that I can see of obtaining an honest X involves an elaborate trick. I now sketch it for those who are interested. (Those who are not can proceed to the next section.)

Adjoin to I a new (instantaneous) state α to produce $I^* \equiv I \cup \{\alpha\}$. Extend Q to an $I^* \times I^*$ matrix Q^* by

$$q^*_\alpha = \infty, \quad q^*_{\alpha i} = q^*_{i\alpha} = 0 \qquad (i \in I).$$

Then Q^* on I^* satisfies the hypotheses of the theorem. By Lemma 9, we can tree-label I^* and find c^*, Q^{*-} etc., etc.. We can obtain X^{*-} (with Q-matrix Q^{*-}) with minimal state-space I^*. Define

$$\varphi^*(t) \equiv 2\int_0^t c^* \circ X^{*-}_s \, ds.$$

Now produce a definitely honest chain X^* via

$$P[\sigma^*_1 > t \mid X^{*-}] = \exp[-\varphi(t)],$$

$$P[X^*(\sigma^*_1) = j \mid X^*(\sigma^*_1-) = i] = \tfrac{1}{2}c^*(i)^{-1}[q^*_{ij} - q^{*-}_{ij}],$$

$$P[X^*(\sigma^*_1) = \alpha \mid X^*(\sigma^*_1-) = i] = \tfrac{1}{2},$$

etc.. Then X^* does not necessarily have Q-matrix Q^* but the Q-matrix of X^* will agree with Q on $I \times I$. Now time-transform X^* by ignoring the time spent by X^* in state α. This produces an honest chain X with Q-matrix Q. The state α becomes a regular <u>fictitious</u> state of X.

INVARIANT MEASURES: FILLING OUT THE TIME

We now return to the problem of ensuring that "almost no time is pushed off the end of the tree". First we must formulate this idea precisely.

Let $\mathfrak{X}_{[n]}$ be the chain (<u>starting at</u> 0) with minimal state-space

$$I_0 \cup I_1 \cup \dots \cup I_n \cup I_{n+1}$$

which is obtained after applying KENDALL's branching procedure down to the n-th level. The states in $I_0 \cup I_1 \cup \dots \cup I_n$ are instantaneous for $\mathfrak{X}_{[n]}$ and the states in I_{n+1} are stable for $\mathfrak{X}_{[n]}$. Note that

(12) $$\mathfrak{X}_{[n]}(t) \in \bigcup_{k\le n} I_k \Rightarrow \mathfrak{X}_{[m]}(t) = \mathfrak{X}_{[n]}(t), \ \forall m \ge n.$$

Thus, on the set

$$\mathcal{J} \equiv \bigcup_n \{t : \mathfrak{X}_{[n]}(t) \in \bigcup_{k\le n} I_k\}$$

we may define

$$X^-(t) \equiv \lim_n \mathfrak{X}_{[n]}(t),$$

the limit existing in the discrete topology of I. What we have to do is to choose the $X^{(i)}$ in a way which guarantees that

(13) (<u>a.s</u>) $\mathcal{J}$ <u>is of full measure</u>.

Define

$$\ell_{[n]}(t,i) \equiv \text{meas}\{s \le t : \mathfrak{X}_{[n]}(s) = i\} \qquad \left(i \in \bigcup_{k\le n+1} I_k\right).$$

Then, for $i \in I_n$,

$$\ell_{[n]}(t,i) = \ell_{[m]}(t,i) = \ell^-(t,i) \quad \text{(say)} \quad (\forall m \geq n),$$

and, for $j \in I_{n+1}$,

$$(14) \qquad \ell^-(t,j) \leq \ell_{[n]}(t,j).$$

It is clear that (13) amounts to the same thing as the statement:

$$(15) \qquad \sum_{j \in I_{n+1}} \ell_{[n]}(t,j) \downarrow 0 \quad (n \uparrow \infty), \ \underline{\text{a.s.}}.$$

Set

$$\tau \equiv \inf\{t : \ell^-(t,0) > 1\}.$$

If we can ensure that

$$(16) \qquad \sum_{j \in I_{n+1}} E\ell_{[n]}(\tau,j) \downarrow 0 \quad (n \uparrow \infty),$$

then (15) is guaranteed and moreover (if for the moment we take the Markov property of X^- for granted) X^- will exist as a <u>positive recurrent</u> Markov chain. DOEBLIN's result that

$$\mu\{j\} \equiv E\ell^-(\tau,j) \qquad (j \in I)$$

<u>then defines the unique invariant measure</u> μ <u>for</u> X^- <u>normalised via the condition</u>

$$\mu\{0\} = 1$$

tells us how to obtain the necessary control over the $X^{(i)}$: <u>control their invariant measures</u>! We need to know that we can do this.

(17) LEMMA. <u>Fix</u> i <u>in</u> I. <u>Let</u> $\mu^{(i)}$ <u>be any totally finite measure on</u> $i \cup Z(i)$ <u>such that</u>

$$\mu^{(i)}\{i\} = 1, \quad \mu^{(i)}\{j\} > 0 \qquad (j \in Z(i)).$$

<u>Then there exists an irreducible, positive-recurrent chain</u> $X^{(i)}$ <u>with</u> Q-<u>matrix</u> $Q^{(i)}$ <u>and invariant measure</u> $\mu^{(i)}$.

The proof of this lemma (which is the most illuminating part of the paper) is deferred to Part 4(a).

We are now free to choose the $\mu^{(i)}$ in any way we wish. Can we choose them so that (16) holds and (11) does define a CAF of X^-? Yes, we can. Absolutely straightforward calculations confirm the formula

$$(18) \qquad E\ell_{[n]}(\tau,j) \leq \mu^{(0)}\{0j_1\}\mu^{(0j_1)}\{0j_1j_2\}\cdots\mu^{(0j_1j_2\cdots j_n)}\{j\},$$

where

$$j_1, j_2, \ldots, j_{n+1} \in \underset{\sim}{N}$$

and

$$j = 0j_1j_2\cdots j_{n+1} \in I_{n+1}.$$

The intuitive reason for (18) is clear. From (18), we can easily see (and prove) that the $\mu^{(i)}$ may be chosen so as to arrange both that (16) holds and (recall (14)) that

$$(19) \qquad \Sigma c(i)\, E\ell^-(\tau,i) < \infty.$$

The inequality (19) guarantees that φ at (11) is a CAF of X^-.

USE OF FREEDMAN'S METHOD

How can we prove rigorously that X^- is indeed a Markov chain with Q-matrix Q^-?

We can certainly do this by utilising the method in Chapter 3 of FREEDMAN's book [3]. (Note that our use of (16) mirrors an idea used very effectively by FREEDMAN.) However, FREEDMAN's method is unnecessarily complicated for our situation in which the logic of (12) makes things simple.

Giving a neat rigorous proof of the Markov property of X^- is one of the main "technicalities" which I would like to think further about before publishing a more complete account of Q-matrices. It would be advantageous

(i) to exploit martingales,

(ii) to consider superposition operations more general than the wedging operation of the branching procedure.

Note. The conjecture about the (γ) problem for Q-matrices at the end of §3.2 in [3] ("The following statements have a good chance to be right ...") is very wrong.

Exercise. Give a counter-example to FREEDMAN's conjecture by considering the non-existent FELLER-McKEAN chain (see (5) above) based on the inallowable measure m on the set I of rationals with

$$m(i) \equiv 1 \qquad (\forall i).$$

Part 4. Proofs of Lemmas 17 and 9

a) PROOF OF LEMMA 17

Now everything becomes clear.

Recall the set-up. Fix i in I. Let $\mu^{(i)}$ be any totally finite measure on $i \cup Z(i)$ such that

$$\mu^{(i)}\{i\} = 1, \ \mu^{(i)}\{j\} > 0 \qquad (j \in Z(i)).$$

We wish to prove that there exists an irreducible, positive-recurrent chain $X^{(i)}$ with $\mu^{(i)}$ as invariant measure and with Q-matrix $Q^{(i)}$ satisfying (10):

$$q_i^{(i)} = \infty, \ q_{ij}^{(i)} = q_{ij} \qquad (j \in Z(i)),$$
$$q_j^{(i)} = \infty, \ q_{jk}^{(i)} = 0 \qquad (j \in Z(i): \ k \in [i \cup Z(i)] \setminus j).$$

We know that we can only hope to do this because the local safety condition holds: there exists an infinite subset $K(i)$ of $Z(i)$ such that

$$Q^{(i)}(i, K(i)) < \infty.$$

Set up a one-one correspondence ρ of $i \cup Z(i)$ with the entire set $\mathbb{Q}^+$ of non-negative rationals in such a way that $\rho(i) = 0$ and $\rho(Z(i) \setminus K(i))$ has no point of accumulation in $[0,\infty]$ other than 0. We can do this precisely because $K(i)$ is infinite.

Let us regard ρ as an identification. We may then think of

$$\mu^{(i)} \text{ as the measure } \tilde{\mu} \equiv \mu^{(i)} \circ \rho^{-1} \text{ on } \mathcal{Q}^{+},$$
$$q_{i\cdot}^{(i)} \text{ as the function } \tilde{q}_{0\cdot} \equiv q_{i\cdot}^{(i)} \circ \rho^{-1} \text{ on } \mathcal{Q}^{++}.$$

(Notation. We write $\mathcal{Q}^{++}$ for the set of strictly positive rationals.)

By our choice of ρ, we have arranged that $\tilde{q}_{0\cdot}$ satisfies the local character condition with respect to the Euclidean topology of $\mathcal{Q}^{+}$:

$$(20) \qquad \tilde{Q}(0,[x,\infty)) \equiv \sum_{j\in \mathcal{Q}^{+}\cap[x,\infty)} q_{0j} < \infty \qquad (\forall x > 0).$$

We can now drop annoying superscripts (i) by thinking about $Y = \rho \circ X^{(i)}$ and reformulating Lemma 17 as follows.

LEMMA 17*. Let $\tilde{\mu}$ be a totally finite measure on $\mathcal{Q}^{+}$ such that

$$\tilde{\mu}\{0\} = 1, \quad \tilde{\mu}\{j\} > 0 \qquad (j \in \mathcal{Q}^{++}).$$

Let $\tilde{q}_{0\cdot}$ be a function on $\mathcal{Q}^{++}$ which satisfies the local character condition (20). Then there exists an irreducible positive-recurrent chain Y with minimal state-space $\mathcal{Q}^{+}$, with invariant measure $\tilde{\mu}$ and with Q-matrix Q^{Y} satisfying

$$q_{0}^{Y} = \infty,\ q_{0j}^{Y} = \tilde{q}_{0j} \qquad (j \in Q^{++}),$$
$$q_{j}^{Y} < \infty,\ q_{jk}^{Y} = 0 \qquad (j \in Q^{++},\ k \in \mathcal{Q}^{+}).$$

Proof. The full state-space of Y will be the entire half-time $[0,\infty)$. First, we choose the full LEVY kernel

$$N(\cdot) = N^{Y}(0,\cdot)$$

of Y at 0.

Choose for N any measure on $((0,\infty), \mathcal{B}(0,\infty))$ such that

21(i) $\quad 0 < N[x,\infty) < \infty \qquad (x > 0),$

21(ii) $\quad N(0,\infty) = \infty,$

21(iii) $\quad N\{j\} = \tilde{q}_{0j} \qquad (j \in \mathcal{Q}^{++}).$

We can choose such an N because of (20).

Next define

$$(22) \qquad q_{j}^{Y} \equiv \tilde{\mu}\{j\}^{-1} N[j,\infty).$$

It follows from standard theory that there exists a simple (but not strong!) Markov process Y with state-space $[0,\infty)$ such that

23(i) Y spends almost all its time in $\mathcal{Q}^{+}$ and so is a chain with minimal state-space $\mathcal{Q}^{+}$;

23(ii) each j in $\mathcal{Q}^{++}$ is a stable state of Y with rate q_{j}^{Y};

23(iii) the paths of Y are continuous and non-increasing on the set $\{t : Y_t \neq 0\}$;

23(iv) state 0 is an instantaneous (but not fictitious) state of Y ;

23(v) $N^Y(0,\cdot) = N(\cdot)$.

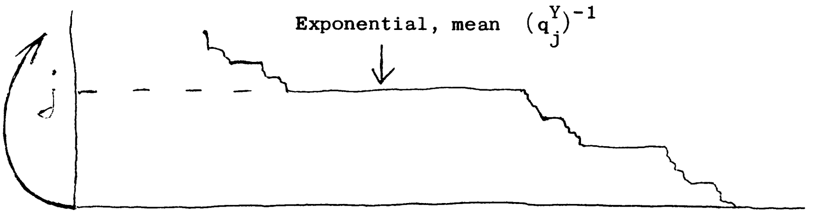
Properties 23 completely characterise the law of Y . The idea of using chains with random Cantor functions as paths is due to LEVY. The way to build-in the LEVY system at 0 is due to FELLER (in 'analytic' work which was developed by NEVEU and REUTER).

We 'cooked' $q^Y_{0j} = \tilde{q}_{0j}$ $(j \in \mathbb{Q}^{++})$ via 21(iii). Since Y can move from j in $\mathbb{Q}^{++}$ to $k(\neq j)$ in $\mathbb{Q}^{+}$ only via countably many other rationals, it follows that

(24) $$q^Y_{jk} = 0 \qquad (j \in \mathbb{Q}^{++}, k \in \mathbb{Q}^{+}) .$$

The picture

LEVY measure $N[j,\infty)$



and DOEBLIN's description (see the discussion before (17)) of invariant measures make it clear that the unique invariant measure μ^Y of Y satisfying $\mu^Y\{0\} = 1$ is given by

$$\mu^Y\{j\} = N[j,\infty)(q^Y_j)^{-1} .$$

Thus (22) 'cooks' the result: $\mu^Y = \tilde{\mu}$.

The proof of Lemma 17 is complete, but one further comment might be helpful. It is easy to "split each rational" so as to obtain the RAY-KNIGHT version of Y . On leaving a "true" state of j in $\mathbb{Q}^{++}$, the RK version of Y will jump to a non-branch fictitious state in the D-space for Y . Of course, (24) rules out the possibility of a jump from j to $\mathbb{Q}^{+}\setminus\{j\}$. The fact that in general we are forced to choose

$$N([0,\infty)\setminus \mathbb{Q}^{+}) > 0$$

further emphasises that the construction only works because we allow jumps to fictitious states.

b) SEYMOUR'S PROOF OF LEMMA 9.

Since the diagonal elements of Q play only a nuisance rôle, we may as well modify things a little (and simplify notation too).

So suppose that I is the countable set $\{0,1,2,\ldots\}$ of non-negative integers. Suppose that Q is an $I \times I$ matrix with

$$0 \le q_{ij} < \infty \text{ for all pairs } (i,j) .$$

We shall now use the kernel notation $Q_i(\cdot)$ in preference to $Q(i,\cdot)$.

Our assumptions are the following:

(N) $$\sum_j q_{aj} \wedge q_{bj} < \infty \qquad (\forall a,b : a \neq b);$$

(S) there exists an infinite subset K of I with

$$Q_i(K) < \infty, \qquad \forall i.$$

We may and shall assume that

$$0 \notin K.$$

We wish to show that I may be tree-labelled in such a way that for each i,

25(i) there is an infinite subset K_i of $Z_i \equiv Z(i)$ such that $Q_i(K_i) < \infty$;

25(ii) $Q_i(I\backslash Z_i) < \infty$.

(Recall that $Z_i \equiv Z(i)$ is the set of immediate successors of i.)

Proof. For $a \neq b$, put

$$T_{ab} \equiv \{j : q_{aj} < q_{bj}\}.$$

Then, by property (N), for $a \neq b$,

$$\infty > \sum_j q_{aj} \wedge q_{bj} = Q_a(T_{ab}) + Q_b(I\backslash T_{ab}),$$

so that

$$Q_a(T_{ab}) < \infty, \quad Q_b(I\backslash T_{ab}) < \infty.$$

Hence (for $a \neq b$) there exists $S_{ab} \subseteq T_{ab}$ with $T_{ab}\backslash S_{ab}$ finite such that

$$Q_a(S_{ab}) < 2^{-b}, \quad Q_b(I\backslash S_{ab}) < \infty.$$

Put

$$W_a \equiv I\backslash \left[\bigcup_{c\neq a} S_{ac}\right].$$

Then

(26) $$Q_a(I\backslash W_a) \leq \sum_{c\neq a} Q_a(S_{ac}) < \infty$$

and for $b \neq a$

(27) $$Q_b(W_a) < Q_b(I\backslash S_{ab}) < \infty.$$

For each i, put

$$J_i \equiv (I\backslash K) \cap (W_i \cup \{i+1\}).$$

Then

$$\bigcup_i J_i \geq (I\backslash K) \backslash \{0\}$$

Note that (by (26)) for each i,

(28) $$Q_i(I\backslash (K\cup J_i)) \leq Q_i(I\backslash W_i) < \infty$$

and (by (27)) for each $i \neq h$,

(29) $$Q_i(J_i \cap J_h) \leq Q_i(J_h) \leq Q_i(W_h) + Q(i,h+1) < \infty.$$

Now put

$$J_i^- \equiv (J_i \cap \{j : j > i\}) \backslash \left[\bigcup_{h<i} J_h\right].$$

Then $J_0^-, J_1^-, J_2^-, \ldots$ are disjoint and

$$\bigcup_i J_i^- = I\backslash K.$$

Further, for each i,

(30) $$Q_i(J_i \backslash J_i^-) \le Q_i(\{j : j \le i\}) + \sum_{h<i} Q_i(J_i \cap J_h) < \infty$$

from (29).

Express K as a <u>disjoint</u> union

$$K = \bigcup_i K_i$$

where each K_i is <u>infinite</u> and

$$K_i \subseteq \{j : j > i\}.$$

Put

$$Z_i \equiv J_i^- \cup K_i \subseteq \{j : j > i\}.$$

Then $Z_0, Z_1, Z_2, \ldots$ are disjoint and

$$\bigcup Z_i = I \backslash \{0\}.$$

It is clear that the Z_i induce a tree-labelling of I (more precisely: a family of tree-labellings of I) and that all that remains is to prove that

$$Q_i(I \backslash Z_i) < \infty.$$

However,

$$Q_i(I \backslash Z_i) \le Q_i(K \backslash K_i) + Q_i(I \backslash (K \cup J_i^-))$$
$$\le Q_i(K) + Q_i(I \backslash (K \cup J_i)) + Q_i(J_i \backslash J_i^-) < \infty$$

from (28), (30) and our assumption that $Q_i(K) < \infty$.

BIBLIOGRAPHY

[1] A. BENVENISTE ans J. JACOD, Systèmes de Lévy des processus de Markov, Invent. Math. 21 (1973), pp. 183-198.

[2] J.L. DOOB, State-spaces for Markov chains, Trans. Amer. Math. Soc. 149 (1970), pp. 279-305.

[3] D. FREEDMAN, Approximating Markov chains, Holden-Day 1972.

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

[5] R.K. GETOOR and M.J. SHARPE. The Ray space of a right process, to appear in Ann. Inst. Fourier Grenoble.

[6] C.T. HOU, The criterion for uniqueness of a Q process, Scientia Sinca vol. XVII No. 2 (1974), pp. 141-159.

[7] K. ITO, Poisson point processes attached to Markov processes, Proc. 6th Berkeley Symposium, vol. III, (1971), pp. 225-240.

[8] D.G. KENDALL, A totally unstable denumerable Markov process, Quarterly J. Math., Oxford, vol. 9, No. 34 (1958), pp. 149-160.

[9] B. MAISONNEUVE, Systèmes régénératifs, Astérisque 15, Société Mathématique de France (1974).

[10] J. NEVEU, Lattice methods and subMarkovian processes, Proc. 4th Berkeley Symposium, vol. 2 (1960), pp. 347-391.

[11] J. NEVEU, Une généralisation des processus à accroissements positifs indépendants, Abh. Math. Sem. Univ. Hamburg 25 (1961), pp. 36-61.

[12] J. NEVEU, Sur les états d'entrée et les états fictifs d'un processus de Markov. Ann. Inst. Henri Poincaré 17 (1962), pp. 323-337.

[13] J. NEVEU, Entrance, exit and fictitious states for Markov chains, Proc. Aarhus Colloq. Combinatorial Probability (1962), pp. 64-68.

[14] G.E.H. REUTER, paper on HOU's uniqueness theorem (to appear in Z f W).

[15] G.E.H. REUTER and P.W. RILEY, The Feller property for Markov semigroups on a countable state-space, J. London Math. Soc. (2), 5(1972), pp. 267-275.

[16] D. WILLIAMS, A note on the Q-matrices of Markov chains, Z. Wahrscheinlichkeitstheorie verw. Gebiete 7 (1967), pp. 116-121.

[17] D. WILLIAMS, Fictitious states, coupled laws and local time, Z f W 11 (1969), pp. 288-310.

[18] D. WILLIAMS, On operator semigroups and Markov groups, Z f W 13 (1969), pp. 280-285.

[19] D. WILLIAMS, Brownian motions and diffusions as Markov processes, Bull. London Math. Soc. 6 (1974), 257-303.

[20] D. WILLIAMS, The Q-matrix problem for Markov chains (to appear).

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