SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

DAVID WILLIAMS

The Q-matrix problem 2 : Kolmogorov backward equations

Séminaire de probabilités (Strasbourg), tome 10 (1976), p. 505-520

http://www.numdam.org/item?id=SPS_1976__10__505_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/>

THE Q-MATRIX PROBLEM 2: KOLMOGOROV BACKWARD EQUATIONS

by

David Williams

Part 1. Introduction

(a) This paper is a sequel to [QMP 1] (= [16]). The main result of [QMP 1] is recalled as Theorem 1 below.

Here we introduce and study the KOLMOGOROV backward equations for arbitrary chains. Theorem 2 solves the existence problem for totally instantaneous chains which satisfy these equations. This theorem is therefore a kind of (dual!) analogue of the 'existence' part of the STROOCK-VARADHAN theorem ([15]) on diffusions.

Two of the chief methods in [QMP 1], SEYMOUR's lemma and KENDALL's branching procedure, again play a large part. However, because the chains constructed in [QMP 1] never satisfy the KOLMOGOROV backward equations, the branching procedure has been substantially modified along lines suggested by FREEDMAN's book [4]. We therefore arrive at the splicing procedure described in Part 4. The splicing technique provides a nice application of ITO's excursion theory.

I hope to show in [QMP 3] that the methods of [QMP 1, 2] may be used to make some slight impact on some altogether more profound and important problems on chains.

(b) Let I be a countably infinite set. Let Q be an $I \times I$ matrix satisfying the DOOB-KOLMOGOROV condition:

$$(DK): \quad 0 \leq q_{ij} < \infty \quad (\forall i, j: i \neq j).$$

For $i \in I$ and $J \subseteq I \setminus i$, write

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

(The symbol " $\equiv$ " signifies "is defined to be equal to".) As usual, define

$$q_i \equiv -q_{ii}.$$

We say that Q is a Q-matrix if there exists a ("standard") transition function $\{P(t)\}$ on I with $P'(0) = Q$. The matrix Q is then called the Q-matrix of $\{P(t)\}$ and of any chain X with minimal state-space I and transition function $\{P(t)\}$. We say that $\{P(t)\}$ (equivalently, X) is honest if $P(t)1 = 1, \forall t$, that is, if X has almost-surely-infinite lifetime.

THEOREM 1. Suppose that Q satisfies ((DK) and) the "totally instantaneous" condition

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

Then Q is a Q-matrix if and only if Q satisfies "NEVEU's condition"

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

and the "safety condition"

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

$$Q(i, K \setminus i) < \infty, \quad \forall i.$$

Further, we can then find an honest $\{P(t)\}$ with $P'(0) = Q$.

(c) The KOLMOGOROV backward equations. Let $\{P(t)\}$ be an honest transition function on I and define $Q = P'(0)$.

Let $B(I)$ be the Banach space of bounded functions on I with the usual supremum norm. With an eye to LEVY systems, define the operator $\mathcal{Q}$ on $B(I)$ as follows:

$$(\mathcal{Q}f)_i \equiv \sum_{j \neq i} q_{ij}(f_j - f_i)$$

on the domain $\mathcal{D}(\mathcal{Q})$ consisting of those f in $B(I)$ such that

- (i) for each i , the series defining $(\mathcal{Q}f)_i$ converges absolutely,
- (ii) $\mathcal{Q}f \in B(I)$.

We shall say that $\{P(t)\}$ satisfies the KOLMOGOROV backward equations

(KBE) if

$$(KBE)_1: \quad A \subseteq \mathcal{Q}$$

(that is: $\mathcal{D}(A) \subseteq \mathcal{D}(\mathcal{Q})$ and $A = \mathcal{Q}$ on $\mathcal{D}(A)$) where A is the strong infinitesimal generator of $\{P(t)\}$ acting on $B(I)$. Define the resolvent $\{\hat{P}(\lambda) : \lambda > 0\}$ of $\{P(t)\}$ as usual:

$$(\hat{P}(\lambda)f)_i \equiv \int_0^\infty e^{-\lambda t} (P(t)f)_i dt \quad (f \in B(I), i \in I).$$

It is standard that $A \subseteq \mathcal{Q}$ if and only if

$$(KBE)_2: \quad (\lambda - \mathcal{Q})\hat{P}(\lambda)f = f \quad (f \in B(I)).$$

Of course, $(KBE)_2$ must be read as implying that $\hat{P}(\lambda) : B(I) \rightarrow \mathcal{D}(\mathcal{Q})$.

As in [QMP 1], we write v_i for the ITO excursion law at i and w_i for a typical excursion path from i . It is easy to guess the following result from work of REUTER [13] and CHUNG [2] on the stable case.

LEMMA 1. (KBE) is equivalent to the statement:

$$(I^{\mathcal{Q}}): \quad (\forall i) \quad v_i\{w_i : w_i(0+) \notin I \setminus i\} = 0.$$

This lemma is proved in Part 2.

Since v_i has total mass q_i and

$$v_i\{w_i : w_i(0+) = j\} = q_{ij} \quad (i \neq j),$$

condition $(I^{\mathcal{Q}})$ implies that

$$(\Sigma) \quad q_i = \sum_{j \neq i} q_{ij} \quad (\leq \infty) \quad (\forall i).$$

If $\{P(t)\}$ satisfies (KBE) and (TI), it therefore follows that $Q \equiv P'(0)$ satisfies (DK), (N) and

$$(TI\Sigma): \quad q_i = \sum_{j \neq i} q_{ij} = \infty \quad (\forall i).$$

Suppose conversely that Q is an $I \times I$ matrix satisfying (DK), (N) and

(TIE). Then Q automatically satisfies condition (S), so that there certainly exists an honest $\{P(t)\}$ with $P'(0) = Q$. Recall however that the methods of [QMP 1] never produce a $\{P(t)\}$ satisfying (KBE). Still, everything works out right.

THEOREM 2. Suppose that Q is an $I \times I$ matrix satisfying (DK), (N) and (TIE). Then there exists an honest transition function $\{P(t)\}$ with generator A satisfying $A \subseteq Q$.

Note. In [QMP 1], the proof of the apparent 'detail' that $\{P(t)\}$ in Theorem 1 can be chosen to be honest was proved by a trick. Since that trick would not work for Theorem 2, we are forced to give the proper (and very much shorter!) proof this time. All that is needed is a direct application of the quasi-left-continuity property in the form for RAY processes.

(d) Let Q be an $I \times I$ matrix satisfying (DK) and (Σ). Note that if $f \in \mathcal{D}(Q)$, then $f^2 \in \mathcal{D}(Q)$ so that $\mathcal{D}(Q)$ is an algebra. An amusing corollary of Theorem 2 is that if condition (TI) also holds, then $\mathcal{D}(Q)$ separates points of (I) if and only if condition (N) holds. This corollary is amusing for two reasons: (i) I can not prove it directly; (ii) it is false if condition (TI) is dropped! Is it possible that the corollary is more than merely amusing?

(e) Our construction will make it clear that the $\{P(t)\}$ in Theorem 2 can not possibly be unique.

The lack of uniqueness of $\{P(t)\}$ in Theorem 2 will be obvious to devotees of the Strasbourg school for the following reasons. Let Q be as in Theorem 2 and let X be a RAY chain with generator A satisfying $A \subseteq Q$. Since X is totally instantaneous, the Baire Category Theorem implies that X almost surely visits uncountably many fictitious states during any time-interval. The set of fictitious states is therefore non-semi-polar and so (DELLACHERIE [3]) contains a (non-semi-polar) finely perfect set. This finely perfect set is the fine support of a continuous additive functional ϕ (DELLACHERIE [3], AZEMA [1]) and we can use ϕ to change the LEVY system of X without destroying the condition $A \subseteq Q$.

Part 2. Proof of Lemma 1

Let $\{P(t)\}$ be an arbitrary ("standard") honest transition function on I and set $Q \equiv P'(0)$. Let X be a good (RAY) chain with minimal state-space I and with transition function $\{P(t)\}$.

Let b be a point of I . Let $f_{ib}, g_{bj}(i, j \in I \setminus b)$ be the usual first-entrance and last-exit functions occurring in the decompositions:

$$(1) \quad p_{ib}(t) = \int_0^t f_{ib}(s) p_{bb}(t-s) ds, \quad p_{bj}(t) = \int_0^t p_{bb}(s) g_{bj}(t-s) ds.$$

See, for example, CHUNG [2]. Let T_b be the hitting time of b . Then

$$F_{ib}(t) \equiv P^i[T_b \leq t] = \int_0^t f_{ib}(s) ds \quad (i \neq b).$$

Introduce the taboo transition function $\{ {}_b^P(t) \}$ on $I \setminus b$ as usual:

$${}_b^P(t) \equiv P^i[T_b > t; X(t) = j].$$

Since $\{P(t)\}$ is honest,

$$(2) \quad \sum_{j \neq b} {}_b^P(t) = 1 - F_{ib}(t).$$

It is standard that

$$(3) \quad g_{bj}(t) \geq \sum_{i \neq b} q_{bi} \cdot {}_b^P(t).$$

This follows because $g_b(\cdot)$ is an entrance law for $\{ {}_b^P(t) \}$ and $g_{bj}(0+) = q_{bj}$.

PROPOSITION 1. The condition

$$(b \xrightarrow{Q}): \quad \nu_b \{ w_b : w_b(0+) \notin I \setminus b \} = 0$$

holds if and only if

$$(4) \quad g_{bj}(t) = \sum_{i \neq b} q_{bi} \cdot {}_b^P(t) \quad (\forall t > 0, j \in I \setminus b).$$

Proof. Set

$$(5) \quad g_b(t) \equiv \sum_{j \neq b} g_{bj}(t).$$

Let $\zeta_b(w_b)$ denote the lifetime of excursion w_b from b . Then $\nu_b \circ \zeta_b^{-1}$ is the classical LEVY-HINCIN measure of the subordinator associated with inverse local time at b . Hence from standard theory (NEVEU [12], KINGMAN [9]) based on (9) below,

$$\nu_b \{ \zeta_b > t \} = g_b(t).$$

Because

$$\nu_b \{ w_b : w_b(0+) = i \} = q_{bi} \quad (i \neq b),$$

it is clear that $(b \xrightarrow{Q})$ holds if and only if

$$(6) \quad g_b(t) = \sum_{i \neq b} q_{bi} [1 - F_{ib}(t)].$$

Proposition 1 now follows on comparing (2), (3) and (6).

Condition $(I \xrightarrow{Q})$ of Lemma 1 therefore holds if and only if (4) holds for every b in I .

Use the 'hat' notation:

$$\hat{c}(\lambda) \equiv \int_0^\infty e^{-\lambda t} c(t) dt \quad (\lambda > 0)$$

for Laplace transforms. Thus (1) takes the form

$$(7) \quad \hat{p}_{ib}(\lambda) = \hat{f}_{ib}(\lambda) \hat{p}_{bb}(\lambda), \quad \hat{p}_{bj}(\lambda) = \hat{p}_{bb}(\lambda) \hat{g}_{bj}(\lambda),$$

and, for obvious probabilistic reasons,

$$(8) \quad \hat{p}_{ij}(\lambda) = \hat{p}_{ij}(\lambda) - \hat{f}_{ib}(\lambda) \hat{p}_{bj}(\lambda).$$

Further, since $\{P(t)\}$ is honest,

$$1 = \lambda \sum_j \hat{p}_{bj}(\lambda) = \lambda \hat{p}_{bb}(\lambda) [1 + \hat{g}_b(\lambda)]$$

so that

$$(9) \quad \hat{p}_{bb}(\lambda)^{-1} - \lambda = \lambda \hat{g}_b(\lambda).$$

Proof that $(KBE) \Rightarrow (I \xrightarrow{Q})$. Assume that (KBE) holds. Take b in I . Set $u \equiv \chi_{\{b\}} \in B(I)$. ($\chi_{\{b\}}$ is the characteristic function of $\{b\}$.) Then the equation

$$(\lambda - \hat{\mathcal{Q}})\hat{P}(\lambda)u = u$$

yields

$$\begin{aligned} (10) \quad \lambda \hat{p}_{bb}(\lambda) - 1 &= \sum_{i \neq b} q_{bi} [\hat{p}_{ib}(\lambda) - \hat{p}_{bb}(\lambda)] \\ &= \hat{p}_{bb}(\lambda) \sum_{i \neq b} q_{bi} [\hat{f}_{ib} - 1]. \end{aligned}$$

From (9) and (10),

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

so that (6) holds and $(b \in \mathcal{Q})$.

Proof that $(I \xrightarrow{\mathcal{Q}}) \Rightarrow (KBE)$. Assume that $(I \xrightarrow{\mathcal{Q}})$ holds. Take b in I . Then from (4), (7) and (8) it follows that for $u \in B(I)^+$ and $h = \hat{P}(\lambda)u$,

$$\hat{p}_{bb}(\lambda)^{-1} h_b - u_b = \sum_{i \neq b} q_{bi} [h_i - \hat{f}_{ib}(\lambda) h_b].$$

But from (9) and (6),

$$\hat{p}_{bb}(\lambda)^{-1} h_b - \lambda h_b = \sum_{i \neq b} q_{bi} [1 - \hat{f}_{ib}(\lambda)] h_b$$

so that

$$\lambda h_b - u_b = \sum_{i \neq b} q_{bi} [h_i - h_b].$$

Thus $h = \hat{P}(\lambda)u \in \mathcal{D}(\hat{\mathcal{Q}})$ (you should check this carefully) and

$$(\lambda - \hat{\mathcal{Q}})\hat{P}(\lambda)u = u.$$

Note. I leave the problem of giving the correct interpretation of (KBE) in the form

$$\frac{d}{dt} P(t) = \hat{\mathcal{Q}} P(t)$$

to people who are more expert (and more interested!) in analysis.

Part 3. KOLMOGOROV's chain "K1"

There is a substantial literature on K1. The paper [8] by KENDALL and REUTER gives a most exhaustive analysis which is taken up in CHUNG's book [2]. See also FREEDMAN [4]. REUTER [14] uses K1 very effectively to obtain results on the rate of convergence of $p(t)$ to 1 as $t \downarrow 0$ for Markov p -functions.

ITO's excursion theory allows us to rephrase the (LEVY-) KENDALL-REUTER-CHUNG description of K1. For K1 itself, ITO's idea provides no more than a rephrasing. However, excursion theory gives the natural language for the "splicing procedure" of Part 4. For Part 4, we need the modified form $\beta \mid_N K1$ of K1 described later in this part. We can use ITO's idea effectively only because of the path-decomposition result which explains how a $\beta \mid_N K1$ chain can be obtained by welding a certain strictly elementary chain onto an $\alpha \mid_0 K1$ chain.

THE CHAIN $K1(b_n, a_n)$

Let I be the set $\{0, 1, 2, \dots\}$. Pick (finite) $b_k > 0$ ($k \in \mathbb{N}$) and (finite) $a_k > 0$ ($k \in \mathbb{N}$) such that $\sum b_k = \infty$ and

$$(11) \quad \sum b_k (a_k + \lambda)^{-1} < \infty \quad (\forall \lambda > 0).$$

Set

$$Q \equiv \begin{pmatrix} -\infty & b_1 & b_2 & b_3 & \dots \\ a_1 & -a_1 & 0 & 0 & \dots \\ a_2 & 0 & -a_2 & 0 & \dots \\ a_3 & 0 & 0 & -a_3 & \dots \\ \cdot & \cdot & \cdot & \cdot & \dots \end{pmatrix}.$$

REUTER [14] gives an analytic proof that there exists a unique honest transition function $\{P(t)\}$ with $P'(0) = Q$. He mentions that CHUNG and I had been able to provide probabilistic proofs of this fact. I guess that CHUNG's proof is essentially the same as mine and goes like this.

Suppose that a RAY chain X with Q -matrix Q exists. Then we see that for $k \in \underline{N}$, X leaves k by jumping to 0 . Hence, with the notation of Part 2,

$$(12) \quad f_{i0}(t) = a_i e^{-a_i t} \quad (i \in \underline{N}),$$

$$(13) \quad {}^0P_{ij}(t) = \delta_{ij} e^{-a_j t} \quad (i, j \in \underline{N}).$$

Since $g_0(\cdot)$ is an entrance law for $\{{}^0P(t)\}$ and $g_{0j}(0+) = b_j$ ($j \in \underline{N}$), we have

$$(14) \quad g_{0j}(t) = b_j e^{-a_j t} \quad (j \in \underline{N}).$$

But now the various equations in Part 2 determine $\{P(t)\}$ uniquely from (12) -

(14). Thus, for example, (9) and (14) give

$$(15) \quad \hat{p}_{00}(\lambda) = [\lambda + \lambda \sum_{j \in \underline{N}} b_j (a_j + \lambda)^{-1}]^{-1}.$$

The existence of $\{P(t)\}$ follows 'constructively' and we see that (11) is exactly the right restriction on $(b_n, a_n : n \in \underline{N})$.

The standard RAY-KNIGHT compactification $\bar{E}$ of I for X (see Part 2 of [QMP 1]) may contain points not in I (this will happen if and only if $\liminf a_n < \infty$). However, we shall always have

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

Thus, almost surely,

$$X(t) \in I, \forall t \geq 0; X(t-) \in I, \forall t > 0.$$

THE ITO DESCRIPTION OF $KI(b_n, a_n)$

The discussion above shown that we can restrict excursion paths $w_0(\cdot)$ from 0 to constant functions with

$$w_0 : (0, \zeta_0(w_0)) \rightarrow \{j\} \quad \text{for some } j \text{ in } \underline{N}$$

and that

$$\nu_0\{w_0 : w_0(0+) = j, \zeta_0(w_0) \in dt\} = a_j b_j e^{-a_j t} dt.$$

ITO [6] and MAISONNEUVE [11] expand on the idea that, in terms of the local time

$$L(t, 0) \equiv \text{meas}\{s \leq t : X(s) = 0\},$$

the excursions from 0 form a Poisson point process (with values in the space of excursions) with characteristic measure ν_0 . We can therefore build X from ν_0 .

THE CHAIN $\beta |_{\mathbb{N}} K1(d_n, a_n - \beta)$

A $\beta |_{\mathbb{N}} K1(b_n, a_n - \beta)$ chain β_Y is a chain identical in law to a $K1(b_n, a_n - \beta)$ chain which is killed at rate β while it is in $\mathbb{N}$ but not killed while it is at 0. Here $\beta > 0$ and the parameters a_n, b_n ($n \in \mathbb{N}$) satisfy

$$\sum b_n = \infty, \sum b_n / a_n < \infty, a_n > \beta \quad (\forall n).$$

If we adjoin a coffin state Δ and put β_Y in Δ from the killing-time on, we obtain β_Y as an honest chain on $\{\Delta, 0, 1, 2, \dots\}$ with Q-matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 0 & -\infty & b_1 & b_2 & \dots \\ \beta & (a_1 - \beta) & -a_1 & 0 & \dots \\ \beta & (a_2 - \beta) & 0 & -a_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

(The dotted lines separate out the components involving Δ .) Again the Q-matrix determines a unique honest transition function on $\{\Delta, 0, 1, 2, \dots\}$. We shall always work with the P^0 law of β_Y : that is, we suppose that β_Y starts at 0.

An excursion path $w_0(\cdot)$ of β_Y from 0 will start at some value $w_0(0+) = j \in \mathbb{N}$ and then will either die at some finite time $\zeta_0(w_0)$ because β_Y jumps to 0 or will jump to Δ at some finite time $\zeta_\Delta(w_0)$ in which case $\zeta_0(w_0) = \infty$. The excursion law β_{ν_0} of β_Y at 0 is specified by the two equations:

$$(16) \quad \beta_{\nu_0}\{w_0 : w_0(0+) = j; \zeta_0(w_0) \in dt\} = b_j(a_j - \beta)e^{-a_j t},$$

$$(17) \quad \beta_{\nu_0}\{w_0 : w_0(0+) = j; \zeta_\Delta(w_0) \in dt\} = b_j \beta e^{-a_j t}.$$

From (17), we see that

$$(18) \quad \beta_{\nu_0}\{w_0 : \zeta_0(w_0) = \infty\} = \alpha \equiv \beta \sum_{j \in \mathbb{N}} b_j / a_j.$$

This means that

$$(19) \quad \text{the total time}$$

$$\Gamma \equiv \text{meas.}\{t : \beta_Y(t) = 0\}$$

spent by β_Y at 0 is exponentially distributed with rate α .

It is also clear from (17) that

$$(20) \quad \text{the probability that } \beta_Y \text{ jumps to } \Delta \text{ from state } j \text{ is}$$

$$\mu_j / \mu(\mathbb{N}) = \beta \mu_j / \alpha$$

where μ is the measure on $\mathbb{N}$ with $\mu_j \equiv \mu(\{j\}) \equiv b_j / a_j$.

Further, (16) and (17) imply that

$$(21) \quad \text{the expected total time spent by } \beta_Y \text{ in state } j \in \mathbb{N} \text{ is}$$

$$\beta^{-1} \mu_j / \mu(\mathbb{N}) = \alpha^{-1} \mu_j.$$

A PATH-DECOMPOSITION RESULT

Define

$$\gamma \equiv \sup\{t : \beta_Y(t) = 0\}.$$

Construct a process X starting at 0 with ITO excursion law at 0 which

is the restriction of β_{ν_0} to the set $\{\zeta_0(w_0) < \infty\}$. Then X will be a $K1(b_n - \beta b_n/a_n, a_n)$ chain. Let $L(\cdot, 0)$ denote the 'local' time spent at 0 by X . With (19) in mind, let Γ^* denote an exponentially distributed variable independent of X and with rate α . Set

$$\gamma^* \equiv \inf\{t: L(t, 0) > \Gamma^*\}.$$

Then $\{X(t): t < \gamma^*\}$ is identical in law to $\{\beta_Y(t): t < \gamma\}$. We can therefore construct a chain identical in law to the chain $\{\beta_Y(t): t < \gamma\}$ by inserting appropriate excursions into the interval $[0, \Gamma]$ which represents the growth of local time at 0 for β_Y . The chain $\{\beta_Y(t+\gamma): t \geq 0\}$ is independent of the chain $\{\beta_Y(t): t < \gamma\}$ and is easily described. Indeed, the chain $\{\beta_Y(t+\gamma): t \geq 0\}$ starts at a point j of $\mathbb{N}$ chosen according to the distribution in (20), stays at j for an exponentially distributed time of rate a_j , and then jumps to and stays in Δ . Hence (22) given an exponentially distributed random variable Γ of rate α we can construct a $\beta|_{\mathbb{N}} K1(b_n, a_n)$ chain β_Y^* such that the time spent by β_Y^* at 0 is EQUAL TO (not just identical in law to) Γ . Of course, we shall have to expand Ω by taking products $(\Omega \rightarrow \Omega \times \tilde{\Omega} \text{ (say)})$ in this construction but we must extend Γ by $\Gamma(\omega, \tilde{\omega}) = \Gamma(\omega)$.

Part 4. Proof of Theorem 2

We say that I is tree-labelled if I is labelled as the set of vertices of the tree

We then write Z_i for the set of immediate successors of i so that we have the following local picture of $i \cup Z_i$:

We also write $\pi: I \setminus 0 \rightarrow I$ for the immediate predecessor map so that $Z_i = \pi^{-1}\{i\}$.

SEYMOUR's lemma (Lemma 9 in [QMP 1]) implies that under the hypotheses of Theorem 2, I may be tree-labelled in such a way that

$$(23) \quad c(i) \equiv \sum_{j \neq i} [q_{ij} - q_{ij}^-] < \infty$$

where

$$q_{ij}^- \equiv q_{ij} \text{ if } j \in i \cup Z_i \\ \equiv 0 \text{ otherwise.}$$

We now suppose that the hypotheses of Theorem 2 hold and that I is already tree-labelled as just described.

LEMMA 2. There exists a probability measure μ on I such that

$$(24) \quad \sum c(i) \mu(i) < \infty$$

and a positive recurrent chain X^- (with minimal state-space I) with μ as an invariant measure and with generator A^- satisfying $A^- \subseteq \mathcal{Q}^-$.

EXTENDING THE LEVY SYSTEM

Before proving Lemma 2, let us see why it implies Theorem 2.

Define

$$\varphi(t) \equiv \int_0^t c \circ X_s^- ds,$$

where c is defined at (23). From (24), it follows 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

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

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

Define further "new" jumps $\sigma_2, \sigma_3, \dots$ in the obvious way. Then $\tilde{X}$, defined for $t < \sigma_\infty \equiv \lim_n \sigma_n$, is a Markov chain with generator $\tilde{A} \subseteq \mathcal{Q}$. If $\sigma_\infty = \infty$ (almost surely), then $\tilde{X}$ is honest and Theorem 2 is proved.

Note that

$$\sigma_1 = \inf \{t : \tilde{X}(t) \notin \tilde{X}(t-) \cup Z_{\tilde{X}(t-)}\}.$$

Hence the "new" jump times $\sigma_1, \sigma_2, \dots$ of $\tilde{X}$ are stopping times relative to the family of σ -algebras $\tilde{\mathcal{F}}_t \equiv \sigma\{\tilde{X}_s : s \leq t\}$ (completed in the usual way). Suppose that $\tilde{X}$ is made into an honest process $\tilde{X}^\Delta$ by the usual adjunction of a coffin state Δ . Then

$$\tilde{X}^\Delta(\sigma_\infty) = \Delta \text{ on } \{\sigma_\infty < \infty\}.$$

But, in the standard RAY-KNIGHT compactification of I associated with $\tilde{X}^\Delta$ (see [QMP 1]),

$$\tilde{X}^\Delta(\sigma_\infty-) = \lim_n \tilde{X}^\Delta(\sigma_n)$$

exists and satisfies

$$1 = \tilde{P}[\tilde{X}^\Delta(\sigma_\infty) = \Delta \mid \tilde{\mathcal{F}}(\sigma_\infty-)] = \tilde{P}(0; \tilde{X}^\Delta(\sigma_\infty-), \{\Delta\})$$

on $\{\sigma_\infty < \infty\}$. (This follows from the quasi-left-continuity property appropriate to RAY processes. See GETTOOR [5].) Hence $\tilde{X}^\Delta(\sigma_\infty-) = \Delta$ on $\{\sigma_\infty < \infty\}$. We can therefore modify $\tilde{X}$ to an honest process X with generator $A \subseteq \mathcal{Q}$ by making X agree with $\tilde{X}$ up to time σ_∞ , putting (say) $X(\sigma_\infty) = 0$ on $\{\sigma_\infty < \infty\}$, and letting X run again (when necessary).

Proof of Lemma 2

The proof of Lemma 2 takes up the remainder of the paper.

We may as well simplify notation by writing Q instead of Q^- . We therefore suppose that Q is an $I \times I$ matrix satisfying (DK), (TIS) and the further condition:

$$(Q \downarrow) \quad q_{ij} > 0 \Leftrightarrow j \in Z_1.$$

(The " $\leq$ " condition in $(Q \downarrow)$ is easily shown to be harmless.)

Remarks (i) It is not surprising that the condition $(Q \downarrow)$ determines the crucial case of Theorem 2. Readers unfamiliar with FREEDMAN's book [4] might find it rather difficult to arrange for a chain satisfying $(Q \downarrow)$ and $(I \overset{Q}{\rightarrow})$ to be able to return to state 0 (more or less immediately!) after leaving it. It is in puzzling out such things that much of the charm of chain theory remains.

(ii) I have an alternative proof of Lemma 2 based on the properties of branch-points of RAY processes. This alternative proof makes it easier to understand intuitively how certain chains satisfying $(Q \downarrow)$ and $(I \overset{Q}{\rightarrow})$ are able to return to 0. However, I believe that the present proof is 'better' (in a sense which I hope to clarify in [QMP 3]). The alternative proof is no shorter than the one given here.

CHOICE OF INVARIANT MEASURE μ

Define

$$b_i \equiv Q(\pi(i), i), \quad i \in I \setminus 0.$$

Let c be a given non-negative function on I . (Of course, this function c now plays the role of the 'correction term' c in (23).) Then

(24) there exists a probability measure μ on I such that

$$(24i) \quad \mu_k > 0 \quad (\forall k), \quad \sum_i c_i \mu_i < \infty,$$

and

$$(24ii) \quad \frac{\mu_j}{\mu(Z_{\pi(j)})} < \frac{b_j \mu_{\pi(j)}}{b_{\pi(j)} \mu_{\pi \circ \pi(j)}}, \quad \forall j \in I \setminus [0 \cup Z_0].$$

To prove (24), first choose a totally finite measure ν on I with $\nu_k > 0$ ($\forall k$) and such that $\sum_i c_i \nu_i < \infty$. Then make an obvious recursive use of the following elementary proposition.

PROPOSITION. Suppose that ν^* and b^* are measures on $\mathbb{N}$ with $\nu_k^* > 0, b_k^* > 0$ ($\forall k \in \mathbb{N}$) and $1 < b^*(\mathbb{N}) \leq \infty$. Then there exists a measure μ^* on $\mathbb{N}$ such that

$$0 < \mu_j^* \leq \nu_j^* \quad (\forall j), \quad \mu_j^* / \mu^*(\mathbb{N}) \leq b_j^* \quad (\forall j).$$

[Proof of proposition.] Choose η such that $1 < \eta < b^*(\mathbb{N})$. Let λ be a probability measure on $\mathbb{N}$ with $0 < \lambda_k \leq \eta^{-1} b_k^*$ ($\forall k$). Choose K so that

$$\lambda(\{1, 2, \dots, K\}) > \eta^{-1}.$$

Set

$$\begin{aligned}\mu_j^* &\equiv \left(\min_{k \leq K} v_k^* \right) \lambda_j \quad (j \leq K), \\ &\equiv \left[\left(\min_{k \leq K} v_k^* \right) \lambda_j \right] \wedge v_j^* \quad (j > K). \quad \square\end{aligned}$$

THE CHAINS $X^{(i)}$

Our matrix Q continues to satisfy $(DK), (TE)$ and $(Q\downarrow)$. Let μ be any probability measure on I satisfying (24 ii). By splicing together various chains $X^{(i)}$, we shall construct a positive recurrent chain X with minimal state-space I , with generator A satisfying $A \subseteq \mathfrak{A}$ and with (necessarily unique) invariant probability measure μ .

$X^{(i)}$ will be a chain on $i \cup Z_i$ but we may consider $i \cup Z_i$ as naturally labelled via the correspondence

$$i \leftrightarrow 0, i1 \leftrightarrow 1, i2 \leftrightarrow 2, \dots$$

This labelling allows us the obvious interpretation of the following set-up:

- (25) $X^{(0)}$ is of type $K1(b_j, a_j : j \in Z_0)$;

(26) $X^{(i)}$ is of type $\beta_i | Z_i K1(b_j, a_j : j \in Z_i)$ ($i \in I \setminus 0$) ;

(27) $\{a_j : j \in I \setminus 0\}$ is defined recursively via

$$\frac{b_j}{a_j} = \frac{\mu_j}{\mu_{\pi(j)}} ;$$

(28) $\{\beta_i : i \in I \setminus 0\}$ is defined via the consistency condition:

$$a_i = \alpha_i \equiv \beta_i \sum_{j \in Z_i} b_j / a_j .$$

For $i \in I \setminus 0$, we now regard $X^{(i)}$ as a killed chain with state-space $i \cup Z_i$ (not as an honest chain with state-space $i \cup Z_i \cup \Delta$). For (26) to make sense, we must have

$$a_j > \beta_i \quad (j \in Z_i)$$

and this is exactly guaranteed by 24(ii).

SPLICING THE CHAINS $X^{(i)}$ TO OBTAIN X

Define $I_0 \equiv \{0\}$, $I_1 \equiv Z_0$, and, generally,

$$I_{n+1} = \pi^{-1} I_n \quad (n \geq 0).$$

Define $X_{[0]} \equiv X^{(0)}$. The state-space of $X_{[0]}$ is $0 \cup I_1$, of which state 0 is instantaneous and states in I_1 are stable. (Important. We start $X_{[0]}$ at 0, so we always work with the $p^{(0)}$ law of $X_{[0]}$.)

Each visit by $X_{[0]}$ to a state i in I_1 is exponentially distributed with rate a_i defined by (27). Define

$$L_{[0]}(t, k) \equiv \text{meas}\{s \leq t : X_{[0]}(s) = k\} \quad (k \in 0 \cup I_1)$$

and

$$\tau_{[0]} \equiv \inf\{t : L_{[0]}(t, 0) > 1\}.$$

The number of visits by $X_{[0]}$ to a state i in I_1 before time $\tau_{[0]}$ has (the Poisson distribution of) mean b_i . Hence

$$(29) \quad EL_{[0]}(\tau_{[0]}, i) = b_i/a_i = \mu_i/\mu_0 \quad (i \in I_1).$$

Formula (29) confirms DOEBLIN's interpretation of the fact that μ restricted to $O \cup I_1$ is the (unique modulo constant multiples) invariant measure for the positive recurrent chain $X_{[0]}$.

As already mentioned, each i-interval ($i \in I_1$) of $X_{[0]}$ (that is: each visit made by $X_{[0]}$ to state i) is exponentially distributed with rate a_i . Because of (19), the consistency formula (28) arranges that under the $p^{(i)}$ law of $X^{(i)}$, the total time spent by $X^{(i)}$ at i also has the exponential distribution of rate a_i .

Because of the path-decomposition result described at the end of Part 3, we can therefore build up from any i-interval ($i \in I_1$) of $X_{[0]}$ a chain with the $p^{(i)}$ law of $X^{(i)}$ by inserting suitable excursions (into Z_1) throughout this i-interval. It is important that one excursion has to be inserted immediately after the right-hand end-point of the i-interval.

We now assume that for each i in I_1 , each i-interval of $X_{[0]}$ is built into a chain with the $p^{(i)}$ law of $X^{(i)}$ in the manner just described. This operation produces a chain $X_{[1]}$ on $O \cup I_1 \cup I_2$ for which states in $O \cup I_1$ are instantaneous and states in I_2 are stable. For each path,

$$(30) \quad X_{[0]}(t) = X_{[1]}(\gamma_{01}(t)),$$

where

$$\begin{aligned} \gamma_{01}(t) &\equiv \inf\{s: L_{[1]}(s, I_0 \cup I_1) > t\}, \\ L_{[1]}(t, J) &\equiv \text{meas}\{u \leq t: X_{[1]}(u) \in J\} \end{aligned}$$

for $J \subseteq I_0 \cup I_1 \cup I_2$.

Set

$$\tau_{[1]} \equiv \inf\{t: L_{[1]}(t, 0) > 1\}.$$

Then for $i \in I_1$, $L_{[1]}(\tau_{[1]}, i) = L_{[0]}(\tau_{[0]}, i)$, so that from (29),

$$EL_{[1]}(\tau_{[1]}, i) = \mu_i/\mu_0 \quad (i \in I_1).$$

An easy calculation based on (21) confirms that this last equation also holds

for $i \in I_2$. Thus the restriction of μ to $I_0 \cup I_1 \cup I_2$ is invariant for $X_{[1]}$.

Proceed in the obvious inductive fashion to produce a chain

$$X_{[n]} \text{ on } \underbrace{I_0 \cup I_1 \cup \dots \cup I_n}_{\text{instantaneous}} \cup I_{n+1}^{\text{stable}}$$

with invariant measure μ restricted to $\cup\{I_k: k \leq n+1\}$. The sequence

$(X_{[n]}: n = 0, 1, 2, \dots)$ is time-projective in the obvious sense which generalises (30), and we have arranged that

$$\sum_n \sum_{I_n} EL_{[n]}(\tau_{[n]}, i) = \mu(I)/\mu_0 < \infty.$$

I now claim by analogy (!!!) with the situation studied by FREEDMAN in Chapter 3 of

[4] - and if you will not accept analogy, you can systematically reduce our case to that considered by FREEDMAN - that the projective limit chain X on I exists. The chain X is positive recurrent with unique invariant probability measure μ and $X_{[n]}$ is simply X observed while it is in $I_0 \cup I_1 \cup \dots \cup I_{n+1}$.

PROOF THAT X SATISFIES $A \subseteq \mathcal{Q}$

Define

$$\xi_j \equiv \beta_{\pi(j)} / a_j, \quad \eta_j \equiv 1 - \xi_j \quad (j \in I \setminus 0).$$

Suppose

$$i \in I_1, \quad j \in I_2, \quad k \in I_3, \\ \pi(j) = i, \quad \pi(k) = j.$$

Let us draw (the off-diagonal elements of) the Q -matrix $Q_{[n]}$ of $X_{[n]}$ for $n = 0, 1, 2$. The general pattern will then be clear. The following pictures explain why we chose the $X^{(i)}$ as we did. (The actual calculations of the $Q_{[n]}$ are left as amusing exercises.)

Recall that Q has the picture

We see that $Q_{[n]} \rightarrow Q$ (componentwise) as $n \rightarrow \infty$.

FREEDMAN's convergence theorem, Theorem (1.88) in [4], now identifies Q as the Q -matrix of X . (For the reader's convenience, we provide a simple direct proof of FREEDMAN's theorem in the next section.)

We do not need Freedman's convergence theorem because we can argue directly the desired stronger result that $A \subseteq \mathcal{Q}$. The pictures of $Q_{[0]}, Q_{[1]}, Q_{[2]}, \dots$ are not necessary either but they may help clarify the following argument.

Suppose that $i \in I_n$ ($n \geq 1$). Then each excursion from i made by $X_{[n-1]}$ will begin at some predecessor of i . The splicing which takes $X_{[n-1]}$ to $X_{[n]}$ will remove the possibility of a jump from i to a predecessor of i . Every excursion w_i from i made by $X_{[n]}$ will satisfy $w_i(0+) \in Z_i$ and we shall have

$$\nu_i\{w_i(0+) = j\} = q_{ij} \quad (j \in Z_i)$$

for the process $X_{[n]}$. Further splittings $X_{[n]} \rightarrow X_{[n+1]} \rightarrow \dots$ will not change the measure $\nu_i \circ w_i(0+)^{-1}$. Hence X satisfies $A \subseteq \mathcal{Q}$.

AN ANALYTIC APPROACH

There may be readers who are prepared to accept that for $b \in I_n, X_{[m]}$ ($m \geq n$) satisfies

$$(31) \quad \nu_b\{w_b(0+) \notin Z_b\} = 0, \quad \nu_b\{w_b(0+) = j\} = q_{bj},$$

but who will hesitate to accept that we can "let $n \rightarrow \infty$ to deduce that (31) holds for X ". In such circumstances, we can resort to analytic methods which leave no room for doubt. (CHUNG, FREEDMAN and I believe however that it is best to tighten the probabilistic reasoning.) We shall deal analytically with the problem of (31) in a moment. First, let us test out the analysis by giving a short direct proof of FREEDMAN's convergence theorem.

[[Proof of FREEDMAN's convergence theorem. Let X be any chain on a countable set I . Let (J_n) be an increasing sequence of subsets of I with union I . Let X_n be " X observed only while it is in J_n ". Let $p(t; i, j), Q(i, j), \dots$ (instead of $p_{ij}(t), q_{ij}$) refer to X and let $p_n(t; i, j), Q_n(i, j), \dots$ refer to X_n . We must prove that

$$Q_n(i, j) \rightarrow Q(i, j) \quad (n \rightarrow \infty).$$

We know that

$$\int_0^t p(s; i, j) ds$$

is the $P^{(i)}$ -expected time that X spends at j before X -time t . Hence

$$(32) \quad \int_0^t p_n(s; i, j) ds \downarrow \int_0^t p(s; i, j) ds, \quad (n \uparrow).$$

Since

$$(33) \quad Q(i, j) = \lim_{\lambda \uparrow \infty} \lambda [\hat{p}(\lambda; i, j) - \delta_{ij}]$$

we have

$$Q_n(i, j) \downarrow Q_\infty(i, j) \geq Q(i, j) \quad (n \uparrow)$$

By an obvious 'holding-time' argument, $Q_\infty(i, i) = Q(i, i), \forall i$. It is therefore enough to prove that $Q(b, j) \geq Q_\infty(b, j)$ when $j \neq b$.

From (32),

$$\hat{p}_n(\lambda; i, j) \rightarrow \hat{p}(\lambda; i, j).$$

Hence, from (7) and (8),

$$\hat{p}_n(\lambda; i, j) \rightarrow \hat{p}(\lambda; i, j), \quad \hat{g}_n(\lambda; b, j) \rightarrow \hat{g}(\lambda; b, j).$$

But, from (3),

$$\hat{g}_n(\lambda; b, j) \geq Q_n(b, j) \cdot \hat{p}_n(\lambda; j, j).$$

Let $n \rightarrow \infty$ to find that

$$\lambda \hat{g}(\lambda; b, j) \geq Q_\infty(b, j) \cdot \hat{p}(\lambda; j, j)$$

and now let $\lambda \uparrow \infty$ to get the desired result. See KINGMAN [10] for a deeper convergence theorem.]

Warning. It is very important that the monotonicity in (32) only takes effect after n is so large that $i, j \in J_n$. (Otherwise, one could prove some extraordinary results.)

Discussion of (31). Assume that $X_{[m]}$ satisfies the appropriate version of (KBE) for each m . Fix b and j and restrict attention to those m such that both b and j belong to $\cup\{I_k : k < m\}$. By Proposition 1,

$$\hat{g}_{[m]}(\lambda; b, j) = \sum_{i \in Z_b} q_{bi} \cdot \hat{p}_{[m]}(\lambda; i, j).$$

As $m \uparrow$, we have strict monotonicity (see Warning above) on the right-hand-side.

Hence

$$(34) \quad \hat{g}(\lambda; b, j) = \sum_{i \in Z_b} q_{bi} \cdot \hat{p}(\lambda; i, j).$$

Since (34) holds for all b and j , X satisfies (KBE).

We can of course try to carry the analysis the whole way by defining explicitly the generator A of our chain X . Compare KENDALL [7].

THOUGHT ON BRANCH-POINTS OF X

Suppose that $i(0) = 0, i(1), i(2), \dots \in I$ and that

$$i(k+1) \in Z_{i(k)}, \quad \forall k.$$

It seems intuitively plausible from our pictures of the $Q_{[n]}$ that if

$$\prod_{n \geq 2} \xi_{i(n)} > 0,$$

then, in the RAY-KNIGHT compactification of X , the sequence $(i(n))$ converges to a branch-point x of X with

$$P(0; x, \{0\}) = \prod_{n \geq 2} \xi_{i(n)},$$

$$P(0; x, \{i(k)\}) = \eta_{i(k+1)} \prod_{k \geq n+2} \xi_{i(k)} \quad (k \geq 1).$$

BIBLIOGRAPHY

- [1] J. AZEMA, Une remarque sur les temps de retour, trois applications, Séminaire de Prob. Strasbourg VI, Lect. Notes vol. 258, 1972.
- [2] K.L. CHUNG, Markov chains with stationary transition probabilities, Springer, Berlin, (2nd edition), 1967.
- [3] C. DELLACHERIE, Ensembles épais: applications aux processus de Markov, C.R. Acad. Sci. Paris 266, 1258-1261, 1968.
- [4] D. FREEDMAN, Approximating Markov chains, Holden-Day, San Francisco, 1971.
- [5] R.K. GETTOOR, Markov processes: Ray processes and right processes, Lect. Notes vol. 440, 1975.
- [6] K. ITO, Poisson point processes attached to Markov processes, Proc. 6th Berkeley Symposium, vol. III, 225-240, 1971.
- [7] D.G. KENDALL, A totally unstable denumerable Markov process, Quart. J. Math. Oxford 9, 149-160, 1958.

- [8] D.G. KENDALL and G.E.H. REUTER, Some pathological Markov processes with a denumerable infinity of states and the associated semigroups of operators on $\mathcal{L}$, Proc. Intern. Congress Math. 1954 (Amsterdam) 3, 377-415, 1956.
- [9] J.F.C. KINGMAN, Regenerative phenomena, Wiley, London, New York, 1972.
- [10] J.F.C. KINGMAN, A property of the derivatives of Markov transition properties, Quart. J. Math. Oxford (2) 26, 121-128, 1975.
- [11] B. MAISONNEUVE, Systèmes régénératifs, Astérisque 15, Société Mathématique de France, 1974.
- [12] J. NEVEU, Une généralisation des processus à accroissements positifs indépendants, Abh. Math. Sem. Univ. Hamburg 25, 36-61, 1961.
- [13] G.E.H. REUTER, Denumerable Markov processes and the associated contraction semi-groups on $\mathcal{L}$, Acta. Math. 97, 1-46, 1957.
- [14] G.E.H. REUTER, Remarks on a Markov chain example of Kolmogorov, Z. Wahrscheinlichkeitstheorie 13, 315-320, 1969.
- [15] D.W. STROOCK and S.R.S. VARADHAN, Diffusion processes with continuous coefficients: I, II, Co m. Pure Appl. Math. XXII, 345-400, 479-530, 1969.
- [16] D. WILLIAMS, The Q-matrix problem, Séminaire de Prob. Strasbourg X.

Note. In connection with [15] and the remarks at the beginning of Part 3 of [QMP 1], see also STROOCK's very important paper "Diffusion processes associated with Levy generators", Z. Wahrscheinlichkeitstheorie 32, 209-244 (1975). However it now looks as if the methods of [QMP 1,2] are the right ones for chains.

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