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>

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

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

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

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

<u>and the "safety condition"</u>

(S): <u>there exists an infinite subset K of I such that</u>

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

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

(c) The <u>KOLMOGOROV</u> backward equations. Let $\{P(t)\}$ be an <u>honest</u> 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:

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

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

(i) <u>for each</u> i, <u>the series defining</u> $(\mathcal{Q}f)_i$ <u>converges absolutely</u>,

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

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

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

(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:

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

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

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

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

As in [QMP 1], we write ν_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) <u>is equivalent to the statement</u>:

$$(I^{Q}_{\to}): \qquad (\forall i) \quad \nu_i\{w_i : w_i(0+) \notin I\backslash i\} = 0.$$

This lemma is proved in Part 2.

Since ν_i has total mass q_i and

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

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

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

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

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

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

$(TI\Sigma)$. 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* $(TI\Sigma)$. *Then there exists an honest transition function* $\{P(t)\}$ *with generator* A *satisfying* $A \subseteq \mathfrak{Q}$.

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 $f \in \mathcal{D}(\mathfrak{Q})$, then $f^2 \in \mathcal{D}(\mathfrak{Q})$ so that $\mathcal{D}(\mathfrak{Q})$ is an *algebra*. An amusing corollary of Theorem 2 is that *if condition* (TI) *also holds, then* $\mathcal{D}(\mathfrak{Q})$ *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 $A \subseteq \mathfrak{Q}$. 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 $A \subseteq \mathfrak{Q}$.

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} $(i,j \in I\backslash b)$ be the usual *first-entrance* and *last-exit* functions occurring in the decompositions:

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

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

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

Introduce the <u>taboo transition function</u> $\{{}_bP(t)\}$ on $I\backslash b$ as usual:

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

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

(2) $$\sum_{j \neq b} {}_bp_{ij}(t) = 1 - F_{ib}(t).$$

It is standard that

(3) $$g_{bj}(t) \geq \sum_{i \neq b} q_{bi} \cdot {}_bp_{ij}(t).$$

This follows because $g_{b\cdot}(\cdot)$ is an <u>entrance law</u> for $\{{}_bP(t)\}$ and $g_{bj}(0+) = q_{bj}$.

PROPOSITION 1. <u>The condition</u>

$(b \overset{Q}{\rightarrow})$: $$\nu_b\{w_b : w_b(0+) \notin I\backslash b\} = 0$$

<u>holds if and only if</u>

(4) $$g_{bj}(t) = \sum_{i \neq b} q_{bi} \cdot {}_bp_{ij}(t) \qquad (\forall t > 0,\ j \in I\backslash b).$$

<u>Proof</u>. Set

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

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-HINČIN 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

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

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

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

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

<u>Condition</u> $(I \overset{Q}{\rightarrow})$ <u>of Lemma 1 therefore holds if and only if</u> (4) <u>holds for every</u> b <u>in</u> I.

Use the 'hat' notation:

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

for Laplace transforms. Thus (1) takes the form

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

and, for obvious probabilistic reasons,

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

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

<u>Proof that</u> (KBE) $\Rightarrow$ $(I \overset{Q}{\rightarrow})$. 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

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

yields

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

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 $(b \overset{Q}{\to})$.

Proof that $(I \overset{Q}{\to}) \Rightarrow$ (KBE). Assume that $(I \overset{Q}{\to})$ 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 $h = \hat{P}(\lambda)u \in \mathcal{D}(\mathfrak{Q})$ (you should check this carefully) and

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

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

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

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 ${}^{\beta|\underset{\sim}{N}}K1$ 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 ${}^{\beta|\underset{\sim}{N}}K1$ chain can be obtained by welding a certain strictly elementary chain onto an ${}^{\alpha|0}K1$ chain.

THE CHAIN $K1(b_n, a_n)$

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

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

Set

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

REUTER [14] gives an analytic proof that there exists a <u>unique</u> honest transition function $\{P(t)\}$ with $P'(0) = \mathbf{Q}$. 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 $k \in \underset{\sim}{N}$, X leaves k by jumping to 0. Hence, with the notation of Part 2,

(12) $$f_{i0}(t) = a_i e^{-a_i t} \qquad (i \in \underset{\sim}{N}),$$

(13) $${}_0p_{ij}(t) = \delta_{ij} e^{-a_j t} \qquad (i,j \in \underset{\sim}{N}).$$

Since $g_{0\cdot}(\cdot)$ is an entrance law for $\{{}_0P(t)\}$ and $g_{0j}(0+) = b_j \quad (j \in \underset{\sim}{N})$, we have

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

But now the various equations in Part 2 determine $\{P(t)\}$ uniquely from (12) - (14). Thus, for example, (9) and (14) give

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

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

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 $K1(b_n, a_n)$

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

$$w_0 : (0, \zeta_0(w_0)) \to \{j\} \quad \text{for some } j \text{ in } \underset{\sim}{N}$$

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 <u>Poisson point</u> process (with values in the space of excursions) with characteristic measure ν_0. We can therefore build X from ν_0.

THE CHAIN ${}^{\beta|\underset{\sim}{N}}K1(d_n, a_n - \beta)$

A ${}^{\beta|\underset{\sim}{N}}K1(b_n a_n - \beta)$ chain ${}^{\beta}Y$ 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 $\underset{\sim}{N}$ but not killed while it is at 0. Here $\beta > 0$ and the parameters a_n, b_n $(n \in \underset{\sim}{N})$ satisfy

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

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

$$\left(\begin{array}{c|cccc} 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & -\infty & b_1 & b_2 & \cdots \\ \beta & (a_1-\beta) & -a_1 & 0 & \cdots \\ \beta & (a_2-\beta) & 0 & -a_2 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdots \end{array}\right)$$

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

An excursion path $w_0(\cdot)$ of ${}^{\beta}Y$ from 0 will start at some value $w_0(0+) = j \in \underset{\sim}{N}$ and then will either die at some finite time $\zeta_0(w_0)$ because ${}^{\beta}Y$ 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 ${}^{\beta}\nu_0$ of ${}^{\beta}Y$ at 0 is specified by the two equations:

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

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

From (17), we see that

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

This means that

(19) the total time

$$\Gamma \equiv \text{meas.}\{t : {}^{\beta}Y(t) = 0\}$$

spent by ${}^{\beta}Y$ at 0 is exponentially distributed with rate α.

It is also clear from (17) that

(20) the probability that ${}^{\beta}Y$ jumps to Δ from state j is

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

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

Further, (16) and (17) imply that

(21) the expected total time spent by ${}^{\beta}Y$ in state $j \in \underset{\sim}{N}$ is

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

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 ${}^{\beta}\nu_0$ 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

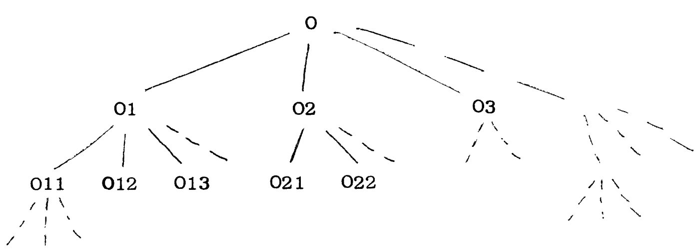
$$\gamma^* \equiv \inf\{t : L(t,0) > \Gamma^*\} .$$

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 ${}^{\beta}Y$. 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 $\underset{\sim}{N}$ 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

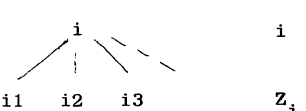
(22) given an exponentially distributed random variable Γ of rate α we can construct a ${}^{\beta|\underset{\sim}{N}}K1(b_n, a_n)$ chain ${}^{\beta}Y^*$ such that the time spent by ${}^{\beta}Y^*$ at 0 is EQUAL TO (not just identical in law to) Γ. Of course, we shall have to expand Ω by taking products $(\Omega \to \Omega \times \tilde{\Omega}$ (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 $\pi : I \setminus 0 \to I$ 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) $$c(i) \equiv \sum_{j \neq i} [q_{ij} - q^-_{ij}] < \infty$$

where

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

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

LEMMA 2. <u>There exists a probability measure</u> μ <u>on</u> I <u>such that</u>

(24) $$\Sigma c(i)\mu(i) < \infty$$

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

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 $\sigma_2, \sigma_3, \ldots$ in the obvious way. Then $\tilde{X}$, defined for $t < \sigma_\infty \equiv \lim \sigma_n$, is a Markov chain with generator $\tilde{\mathcal{A}} \subseteq \mathfrak{Q}$. 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 $\sigma_1, \sigma_2, \ldots$ of $\tilde{X}$ are stopping times relative to the family of σ-algebras $\tilde{\mathcal{F}}_t \equiv \sigma\{\tilde{X}_s : s \le t\}$ (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 GETOOR [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 $\mathcal{A} \subseteq \mathfrak{Q}$ 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),(TIΣ) and the further condition:

(Q↓) $$q_{ij} > 0 \iff j \in Z_1 .$$

(The "<=" condition in (Q↓) is easily shown to be harmless.)

Remarks (i) It is not surprising that the condition (Q↓) 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↓) and $(I \overset{Q}{\to})$ 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↓) and $(I \overset{Q}{\to})$ 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

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

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

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

and

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

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

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

$$0 < \mu^*_j \le \nu^*_j \quad (\forall j), \quad \mu^*_j / \mu^*(\underset{\sim}{N}) \le b^*_j \quad (\forall j) .$$

⟦Proof of proposition. Choose η such that $1 < \eta < b^*(\underset{\sim}{N})$. Let λ be a probability measure on $\underset{\sim}{N}$ with $0 < \lambda_k \le \eta^{-1} b^*_k$ $(\forall k)$. Choose K so that

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

Set

$$\mu_j^* \equiv \left(\min_{k\le K}\nu_k^*\right)\lambda_j \qquad (j \le K),$$
$$\equiv \left[\left(\min_{k\le K}\nu_k^*\right)\lambda_j\right] \wedge \nu_j^* \qquad (j > K).]]$$

THE CHAINS $X^{(i)}$

<u>Our matrix</u> Q <u>continues to satisfy</u> (DK),(TΣ) <u>and</u> ($Q\downarrow$). <u>Let</u> μ <u>be any probability measure on</u> I <u>satisfying</u> (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 $A \subseteq \mathfrak{Q}$ 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

$$i \leftrightarrow 0, i1 \leftrightarrow 1, i2 \leftrightarrow 2, \ldots .$$

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

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

(26) $X^{(i)}$ <u>is of type</u> ${}^{\beta_i|Z_i}K1(b_j, a_j : j \in Z_i)$ $\quad (i \in I\backslash 0)$;

(27) $\{a_j : j \in I\backslash 0\}$ <u>is defined recursively via</u>

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

(28) $\{\beta_i : i \in I\backslash 0\}$ <u>is defined via the consistency condition</u>:

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

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

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

and this is exactly guaranteed by 24(ii).

<u>SPLICING THE CHAINS $X^{(i)}$ TO OBTAIN X</u>

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

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

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. (<u>Important</u>. <u>We start</u> $X_{[0]}$ <u>at</u> 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

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

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

Formula (29) confirms DOEBLIN's interpretation of the fact that μ restricted to $0 \cup I_1$ 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_i) 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 $0 \cup I_1 \cup I_2$ for which states in $0 \cup I_1$ are instantaneous and states in I_2 are stable. For each path,

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

where

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

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),

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

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

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

with invariant measure μ restricted to $\cup\{I_k : k \le n+1\}$. The sequence $(X_{[n]} : n = 0,1,2,\ldots)$ 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 $I_0 \cup I_1 \cup \dots I_{n+1}$.

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

Define

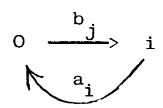
$$\xi_j \equiv \beta_{\pi(j)}/a_j \ , \quad \eta_j \equiv 1 - \xi_j \qquad (j \in I \setminus 0).$$

Suppose

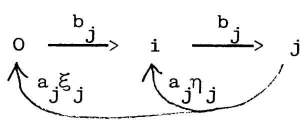
$$i \in I_1, \quad j \in I_2, \quad k \in I_3,$$
$$\pi(j) = i, \quad \pi(k) = j.$$

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. <u>The following pictures explain why we chose the</u> $X^{(i)}$ <u>as we did</u>. (The actual calculations of the $Q_{[n]}$ are left as amusing exercises.)

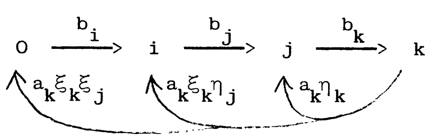
$Q_{[0]}$:



$Q_{[1]}$:



$Q_{[2]}$:



Recall that Q has the picture

$$Q: \qquad 0 \xrightarrow{b_i} i \xrightarrow{b_j} j \xrightarrow{b_k} k \longrightarrow \dots .$$

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

<u>FREEDMAN's convergence theorem</u>, Theorem (1.88) in [4], <u>now identifies</u> Q <u>as the Q-matrix of</u> 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 $\mathcal{A} \subseteq \mathfrak{Q}$. 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. <u>Every excursion</u> w_i <u>from</u> i <u>made by</u> $X_{[n]}$ <u>will satisfy</u> $w_i(0+) \in Z_i$ <u>and we shall have</u>

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

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

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

but who will hesitate to accept that we can "let $n \to \infty$ 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 $p(t;i,j), Q(i,j), \ldots$ (instead of $p_{ij}(t), q_{ij}$) refer to X and let $p_n(t;i,j)$, $Q_n(i,j), \ldots$ refer to X_n. We must prove that

$$Q_n(i,j) \to Q(i,j) \qquad (n \to \infty).$$

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

Since

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

we have

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

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),

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

Hence, from (7) and (8),

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

But, from (3),

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

Let $n \to \infty$ to find that

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

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

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

<u>Discussion of</u> (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,

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

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

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

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 $\mathcal{A}$ of our chain X. Compare KENDALL [7].

THOUGHT ON BRANCH-POINTS OF X

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

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

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

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. GETOOR, 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 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 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, Com. 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.