SOME PROBLEMS RELATING TO MARKOV GROUPS

JANE M. O. SPEAKMAN
UNIVERSITY OF CAMBRIDGE

1. Introduction

This paper forms a sequel to one by Professor D. G. Kendall [2]. We shall use the notation of that paper and assume the results in it. There are two parts: the first investigates the g-function in more detail and reduces the possible region for (γ, Γ), and the second establishes the existence of points on the (Γ, γ) diagram with $\Gamma \neq \gamma$.

The countable state space will be denoted by I.

2. Some properties of the g-function

We recall from [2] that a Markov semigroup has property (F) if and only if there is some $t > 0$ such that $g(t) > \frac{1}{2}$. We are particularly interested in whether there are any non-(U)-semigroups with property (F). The following simple lemma will prove very useful.

LEMMA 1. *For all u and t such that $0 < u < t$ and for all i and $k \in I$, there exist states j and l such that $p_{j,i}(u) \geq p_{k,i}(t)$ and $p_{l,i}(u) \leq p_{k,i}(t)$.*

We have $p_{k,i}(t) = \sum_{h \in I} p_{k,h}(t - u)p_{h,i}(u)$. This is a convex combination of the $p_{h,i}(u)$ $(h \in I)$, and so there exist states j and l as required.

DEFINITION. *Let $S_i(t) = \sup_j p_{j,i}(t)$ and $s_i(t) = \inf_j p_{j,i}(t)$.*

COROLLARY. *For each i, $S_i(t)$ is a nonincreasing function of t, and $s_i(t)$ is nondecreasing.*

THEOREM 1. (a) *On the set $\{t: g(t) \geq \frac{1}{2}\}$, $g(t)$ is a nonincreasing function.*

(b) *A direct sum, $P(t)$, of semigroups, $P_r(t)$, each having the property that its g-function, $g_r(t)$, is continuous on the left, has $g(t) \leq \frac{1}{2}$ for all $t > 0$ unless it enjoys property (U).*

(c) *If $g(t) = m > \frac{1}{2}$, then, for each $u < t$, either $g(u) \geq m$ or $g(u) \leq 1 - m$.*

When $g(t) \geq \frac{1}{2}$, $p_{i,i}(t) \geq \frac{1}{2}$ for all i. Since the row sums of $P(t)$ are 1, each off-diagonal element is less than or equal to $\frac{1}{2}$, and so, for such values of t, $S_i(t) = p_{i,i}(t)$ for every i. Statement (a) now follows from the above corollary.

To prove (b), first note that $g(t) = \inf_r g_r(t)$. We assume that for some $t_0 > 0$, $g(t_0) = m > \frac{1}{2}$. Then $g_r(t_0) \geq m$ for all r. The proof rests on the fact that for each r, $g_r(t)$ is monotone nonincreasing on $[0, t_0]$. By (a), it will be sufficient to show that for $t \in (0, t_0)$, $g_r(t)$ is never less than $\frac{1}{2}$. Assume on the contrary that $g_r(T) < \frac{1}{2}$ for some $T \in (0, t_0)$. The semicontinuity of $g_r(t)$ implies that the set

$S = \{t: g_r(t) < \frac{1}{2}\}$ is open. Now consider the right-hand endpoint, U, of the connected component of S which contains T. Then $U \leq t_0$ and $U \notin S$. By (a), $g_r(U) \geq m$ and so the left continuity of $g_r(t)$ is contradicted at U. To conclude the proof we note that $g(t)$ is monotone nonincreasing on $[0, t_0]$, therefore $\lim_{t \to 0+} g(t)$ exists and is greater than or equal to m. By Reuter's theorem (2), this limit is 1; hence (U) holds.

For the proof of (c), suppose that $g(u) < m$. Then there exists a state i such that $p_{i,i}(u) < m$. By lemma 1 there exists a state j such that $P_{j,i}(u) \geq m$. This state j cannot be i, and so $p_{j,j}(u) \leq 1 - p_{j,i}(u) \leq 1 - m$ and $g(u) \leq 1 - m$.

NOTE: the hypothesis of (b) is satisfied when $P(t)$ is a direct sum of (U)-semigroups. For each r, $g_r(t) = \inf_i p_{i,i}^{(r)}(t)$ and the functions $p_{i,i}^{(r)}(t)$ are continuous uniformly in i. This implies that $g_r(t)$ is a continuous function.

Mr. J. F. C. Kingman strengthened the original version of the following theorem.

THEOREM 2. (a) *For each m such that $\frac{1}{2} < m < 1$, the set $\{t: g(t) \geq m\}$ is closed, and it is nowhere dense unless (U) holds.*

(b) *If there is an instantaneous state, then $g(t) \leq \frac{1}{2}$ for all $t > 0$.*

The Chapman-Kolmogorov equation shows that for all $t, h \geq 0$,

$$(1) \qquad p_{i,i}(t+h) - p_{i,i}(t) = -(1 - p_{i,i}(h))p_{i,i}(t) + \sum_{j \neq i} p_{i,j}(h)p_{j,i}(t)$$
$$\leq -(1 - p_{i,i}(h))g(t) + (1 - p_{i,i}(h))(1 - g(t)),$$

since, if $j \neq i$, $p_{j,i}(t) \leq 1 - g(t)$. Thus

$$(2) \qquad \frac{p_{i,i}(t+h) - p_{i,i}(t)}{h} \leq -\frac{1 - p_{i,i}(h)}{h}(2g(t) - 1).$$

If $g(t) > \frac{1}{2}$ and i is not an absorbing state, the right side is negative.

If i is an instantaneous state, then as $h \to 0+$, the right side of (2) tends to $-\infty$, and so $p_{i,i}(t)$ has an infinite derivative at t. This is impossible ([1], addenda, theorem 4); hence, (b) is proved.

To prove (a), we first note that $\{t: g(t) \geq m\} = \bigcap_i \{t: p_{i,i}(t) \geq m\}$ and so is a closed set. It is sufficient then to show that, when (U) does not hold, this set contains no interval of positive length. Suppose that it does, that is, that $g(t) \geq m > \frac{1}{2}$ for $t \in [t_0, t_1] = J$, where $0 \leq t_0 < t_1$. Then for all i and all $t \in J$,

$$(3) \qquad m \leq p_{i,i}(t) \leq 1.$$

Since (U) does not hold, the semigroup is not q-bounded and there exists $i \in I$ such that $q_i > (1 - m)/\{(t_1 - t_0)(2m - 1)\}$. Now let $h \to 0+$ in (2). We obtain,

$$(4) \qquad p'_{i,i}(t) \leq -q_i(2g(t) - 1) \leq -q_i(2m - 1) < -(1 - m)/(t_1 - t_0).$$

The function $p_{i,i}(t)$ which we know to have continuous derivatives in J cannot satisfy both (3) and (4), and so we arrive at a contradiction.

THEOREM 3. *When (U) does not hold, then $g(t) \leq \frac{2}{3}$ for all $t > 0$.*

For the proof we shall require the following lemma.

LEMMA 2. *If $g(t) \geq \frac{2}{3}$, then for each $u < t$ and for each $i \in I$, there exists at most one j such that $p_{j,i}(u) > \frac{2}{3}$.*

Suppose that there exist $j, l \in I, j \neq l$, such that $p_{j,i}(u) > \frac{2}{3}$ and $p_{l,i}(u) > \frac{2}{3}$. Without loss of generality we may suppose that $p_{j,i}(u) \leq p_{l,i}(u)$. Then

$$(5) \qquad p_{j,j}(t) = p_{j,i}(u)p_{i,j}(t-u) + \sum_{k \neq i} p_{j,k}(u)p_{k,j}(t-u) < p_{j,i}(u)p_{i,j}(t-u) + \tfrac{1}{3},$$

since $\sum_{k \neq i} p_{j,k}(u) = 1 - p_{j,i}(u) < \frac{1}{3}$. Also, $p_{j,j}(t) \geq \frac{2}{3}$; therefore, we have $p_{j,i}(u)p_{i,j}(t-u) > \frac{1}{3}$.

Again, $p_{l,j}(t) \geq p_{l,i}(u)p_{i,j}(t-u) \geq p_{j,i}(u)p_{i,j}(t-u) > \frac{1}{3}$. But $l \neq j$ and $p_{l,l}(t) \geq \frac{2}{3}$ so that $p_{l,j}(t) \leq \frac{1}{3}$. This gives a contradiction and establishes the lemma.

We now give the proof of theorem 3. Suppose that for some $t_0 > 0$, $g(t_0) > \frac{2}{3}$. Since (U) does not hold, Reuter's theorem shows that $\gamma \leq \frac{1}{2}$. Certainly there exists a positive $u < t_0$ such that $g(u) < \frac{2}{3}$, and so there is a state i such that $p_{i,i}(u) < \frac{2}{3}$. By lemma 1, there exists a state j such that $p_{j,i}(u) > \frac{2}{3}$ and j cannot be i. Consider $t_1 = \inf \{t > 0\colon p_{j,i}(t) > \frac{2}{3}\}$. Now $t_1 > 0$ since $p_{j,i}(0) = 0$ and $p_{j,i}$ is continuous. By continuity, $p_{j,i}(t_1) = \frac{2}{3}$ and so there exists a state $l \neq j$ such that $p_{l,i}(t_1) > \frac{2}{3}$. By the continuity of $p_{l,i}$, there is an $\epsilon > 0$ such that, for $t \in [t_1, t_1 + \epsilon)$, $p_{l,i}(t) > \frac{2}{3}$. From the definition of t_1 we see that there is a point in this interval for which also $p_{j,i}(t) > \frac{2}{3}$. This contradicts lemma 2 and establishes the theorem.

It follows from this theorem that $\Gamma \leq \frac{2}{3}$ except when $\gamma = \Gamma = 1$. The next theorem further restricts Γ.

THEOREM 4. *When (U) does not hold, then $\Gamma \leq (\sqrt{5} - 1)/2 = .618\ldots$.*

We write $\sigma = (\sqrt{5} - 1)/2$. Suppose that $\Gamma > \sigma$. For any number m such that $\sigma < m < \Gamma$, the set $S = \{t\colon g(t) \geq m\}$ is closed and has a limit point at the origin.

LEMMA 3. *Every point T of S is a limit point of S from the right (unless S is bounded and T is its largest member).*

Since $p_{i,i}(s)p_{i,i}(t) \leq p_{i,i}(s+t)$ for all $s, t \geq 0$ and all $i \in I$, we have $g(s)g(t) \leq p_{i,i}(s+t)$ for all $s, t \geq 0$ and all $i \in I$ and so $g(s)g(t) \leq g(s+t)$. By letting s tend to 0 suitably, we obtain $\Gamma g(t) \leq \limsup_{s \to 0+} g(s+t)$ for all $t \geq 0$.

It follows that for each $\delta > 0$, there exists $t_n \downarrow T$ such that $g(t_n) \geq \Gamma m - \delta$.

Suppose that there exists a T_1 in S greater than T. Then for each n such that $t_n < T_1$ either $g(t_n) \geq m$ or $g(t_n) \leq 1 - m$ (theorem 1(c)). If the first alternative holds for infinitely many n, the lemma is proved. We show that if δ is sufficiently small, the second alternative is excluded. For we should have $m^2 - \delta < m\Gamma - \delta \leq 1 - m$ and if $\delta < m^2 + m - 1$ (which is positive since $m > \sigma$, the larger root of the equation $x^2 + x - 1 = 0$) this yields a contradiction.

We conclude the proof of theorem 4. Choose a positive element t_0 of S. Consider $(0, t_0)\backslash S$. It is open, and by Reuter's theorem, nonempty. Choose $u \in (0, t_0)\backslash S$. The left-hand endpoint U of that connected component of $(0, t_0)\backslash S$

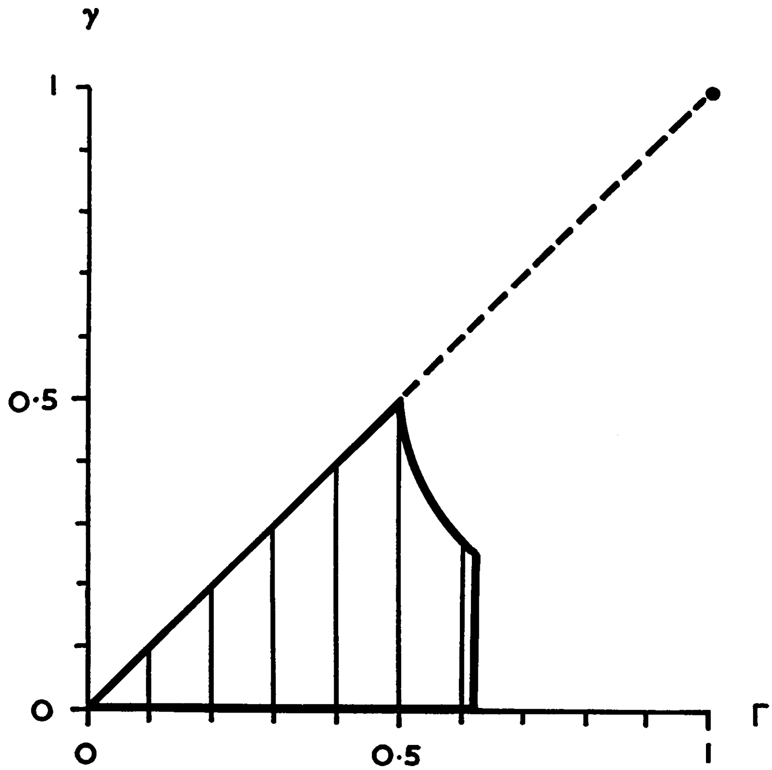

FIGURE 1

The revised (Γ, γ) diagram.

which contains u is in S. Even though $U < t_0$, it does not satisfy the conclusion of lemma 3. This contradicts the initial assumption and establishes the theorem.

The theorem does not rule out the possibility that, for some $t > 0$, $g(t) > \sigma$. But when this is the case we can use it together with theorem 1(c) to show that $\Gamma \leq 1 - g(t)$ if the semigroup is non-(U).

The above results place quite severe restrictions on the behavior of $g(t)$ when it is greater than $\frac{1}{2}$. In particular, if it could be shown that $g(t)$ were continuous on one side at all points greater than 0 for which $g(t) > \frac{1}{2}$, then theorem 2(a) would imply that for a non-(U)-process $g(t)$ is never greater than $\frac{1}{2}$.

3. The construction of some Markov semigroups having $\gamma \neq \Gamma$

We start with a (U)-semigroup $\{P(t)\colon t \geq 0\}$ which has the following property:

$$m \equiv \inf_{t>0} g(t) < \liminf_{t\to\infty} g(t) \equiv M \tag{6}$$

(for example, any three-state Markov chain whose Q-matrix has two complex eigenvalues).

When (6) holds, $M \le \frac{1}{3}$. To see this, first note that there are no inessential states (in the terminology of Chung [1]). For, if i were an inessential state, $\pi_i = \lim_{t\to\infty} p_{i,i}(t)$ would be 0, and M, which can be no greater than π_i, would be 0, and (6) would be contradicted. Next, if i is in some essential class containing only one or two elements, then $p_{i,i}(t)$ is nonincreasing. Thus, if (6) is to hold, there must be an essential class with three or more elements. Let i be an element of such a class. Then $\sum_{j\in C(i)} \pi_j = 1$ ([1], theorem I.7.1), hence, for some $j \in C(i)$, $\pi_j \le \frac{1}{3}$. Since $M \le \pi_j$, we have $M \le \frac{1}{3}$ as required.

We shall construct a semigroup, $R(t)$, which has $\gamma = m$ and $\Gamma = K$, where K is an arbitrary number in $(m, M]$.

The function, $g(t)$ being $\inf_i p_{i,i}(t)$ where the $p_{i,i}(t)$ are continuous uniformly in i, is itself continuous. Because of this and because it is bounded away from m in some (N, ∞), it attains the value m at some point, t_0, in $[0, N]$. For the moment, let us suppose that K is strictly less than M. Then $\{t: g(t) = K\} \cap [0, t_0]$ is nonempty and closed, and hence, contains its greatest lower bound which we shall call t_1. This is the first value of t for which $g(t) = K$. Similarly, $g(t)$ attains the value K for a last time at some point $t_2 > t_0$. The final semigroup will be a direct sum of "speeded-up" versions of $P(t)$, which itself forms the first component. For the second we take $P(kt)$ where $k = t_2/t_1 > 1$ and for the n-th $P(k^{n-1}t)$. The resulting g-function is $\inf_{r=0,1,2,\cdots} g(k^r t)$. This has $\gamma = m$ because it is bounded below by m and attains the value m at the points $k^{-r}t_0 (r = 0, 1, 2, \cdots)$. Also $\Gamma = \Gamma_1$ is at least K because it attains the value K at the points $k^{-r}t_2 (r = 0, 1, 2, \cdots)$. To ensure that Γ is exactly K, we add as a final component a semigroup having $\gamma = \Gamma = K$. Section 4 of [2] tells us that such semigroups exist as K is certainly less than $\frac{1}{2}$. Now $\gamma = \min(m, K) = m$ and $\Gamma = \min(\Gamma_1, K) = K$ as required.

If K had been chosen to be M, we should have taken a sequence (K_n) which increases to M and have used K_n to determine the "speeding-up" factor at the n-th step instead of K itself.

The rest of this section consists of examples employing this construction.

The region that can be obtained in this way when the basic semigroup is a three-state cyclic Markov chain has been investigated in some detail. The Q-matrix of such a semigroup has the form

$$\begin{pmatrix} -a & a & 0 \\ 0 & -b & b \\ c & 0 & -c \end{pmatrix} \tag{7}$$

and $P(t) = e^{Qt}$. If any of a, b, or c is 0, then the diagonal transition functions are all monotone and (6) is not satisfied. Since m and M are unchanged when Q is multiplied by a positive constant we can normalize Q and relabel the states if necessary to obtain Q in the form

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & -b & b \\ c & 0 & -c \end{pmatrix} \tag{8}$$

where $0 < b, c \le 1$. We write $bc = \rho$, $b + c = \sigma$.

The value $\pi_i = \lim_{t\to\infty} p_{i,i}(t)$ is given by the i-th coordinate of the unique positive vector $y = (y_i)$ satisfying $Qy = 0$ and $\sum y_i = 1$ ([1], theorem I.7.1). Clearly, $y = (bc, c, b)/(bc + b + c)$. The smallest coordinate, π_i, is M and it has the value $\rho/(\sigma + \rho)$. Now let $F(x)$ be the distribution function of the sum of three independent exponential random variables with expectations 1, $1/b$, and $1/c$. Then

$$p_{i,i}(t) = e^{-q_i t} + \sum_{r=1}^{\infty} \int_0^t e^{-q_i(t-s)}\, dF^{(r)}(s) \qquad (i = 1, 2, 3), \tag{9}$$

where $F^{(r)}$ is the r-th convolution power of F. Since $e^{-at} \le e^{-bt}$ if $t \ge 0$ and $a \ge b$ we see that, for all $t > 0$, $p_{1,1}(t) \le p_{2,2}(t)$, $p_{3,3}(t)$ and so m, which is $\inf_{t>0} g(t)$, is equal to $\inf_{t>0} p_{1,1}(t)$.

The determinant of $\lambda I - Q$ is $\lambda^3 + (1 + \sigma)\lambda^2 + (\sigma + \rho)\lambda$ and so the eigenvalues of Q are 0 and $-(1 + \sigma)/2 \pm \sqrt{(1 + \sigma)^2/4 - (\sigma + \rho)}$. We write $s = 2\sqrt{|(1 + \sigma)^2/4 - (\sigma + \rho)|}$. Zero cannot be a multiple eigenvalue because we have assumed that both b and c are positive, and so the coefficient of λ in the determinant is positive. There are three possible cases.

CASE I. *The last two eigenvalues are real and distinct.* It is clear that they are both negative. Suppose they are $-f$ and $-g$. Then $p_{1,1}(t) = h + ke^{-ft} + le^{-gt}$ for some constants h, k, l. Such a function has at most one turning point, and if t is positive there this point gives the absolute minimum since we know that $p'_{1,1}(0) < 0$. The constants h, k, and l are determined by the equations $\lim_{t\to\infty} p_{1,1}(t) = \rho/(\sigma + \rho)$, $p_{1,1}(0) = 1$, $p'_{1,1}(0) = -1$. It turns out that there always is a positive turning point and that

$$m = p_{1,1}(t) = \frac{\rho}{\sigma + \rho} - \frac{1}{2}\,\frac{1 - \sigma + s}{\sigma + \rho}\left(\frac{1 - \sigma + s}{1 - \sigma - s}\right)^{-(1+\sigma+s)/2s}. \tag{10}$$

CASE II. *The last two eigenvalues coincide and are $-f$, say.* In this case $p_{1,1}(t) = h + (k + lt)e^{-ft}$. Again such a function has at most one turning point. The constants are determined as in I and we obtain a positive turning point. At this point

$$m = p_{1,1}(t) = \frac{\rho}{\sigma + \rho} - \frac{2(1 - \sigma)}{(1 + \sigma)^2}\, e^{-(1+\sigma)/(1-\sigma)}. \tag{11}$$

CASE III. *The last two eigenvalues form a pair of complex conjugates,* $(-(1 + \sigma)/2 \pm is)$. In this case $p_{1,1}(t) = h + (k \sin st + l \cos st)e^{-(1+\sigma)t/2}$. The constants can be determined as before. Also $p_{1,1}(t)$ may be written in the form $h + Ae^{-(1+\sigma)t/2} \cos (st + B)$. If u is a minimum turning point for a function $h(t)$ of this form, then $h(u) \le f(t)$ for all $t > u$. We know also that $p'_{1,1}(0) < 0$, and so the absolute minimum of $p_{1,1}(t)$ for $t > 0$ is attained at its first positive turning point. Its value is

$$m = \frac{\rho}{\sigma + \rho} - \frac{\sqrt{\rho}}{\sigma + \rho} \exp\left\{-\frac{1 + \sigma}{s} \tan^{-1} \frac{s}{1 - \sigma}\right\}, \tag{12}$$

where the inverse tangent is taken to be in the interval $(0, \pi)$.

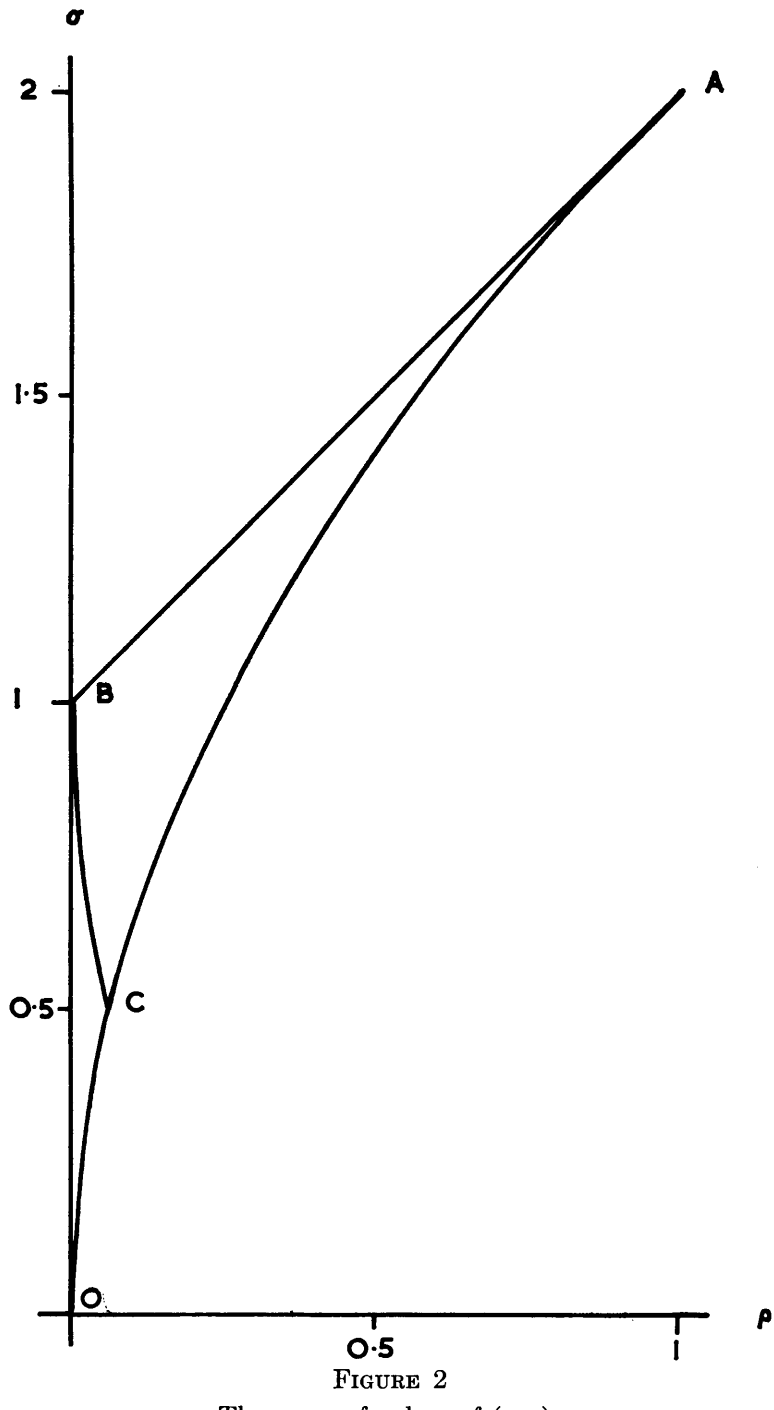

FIGURE 2
The range of values of (ρ, σ).

The possible region for (σ, ρ) is that for which the equation $x^2 - \sigma x + \rho = 0$ has two real roots in the interval (0, 1]. It consists of the points satisfying the conditions $\sigma^2 \geq 4\rho$, $\rho > 0$, $\sigma \leq 1 + \rho$, $\sigma > 0$ and is shown in figure 2. The condition for equal eigenvalues of Q is that $(1 - \sigma)^2 = 4\rho$—the line BC with the point B excluded. In the region ABC (with the line BC excluded) there are complex eigenvalues, and in the region OBC, with BC excluded, distinct real ones. The points on the (Γ, γ)-diagram which can be obtained from semigroups

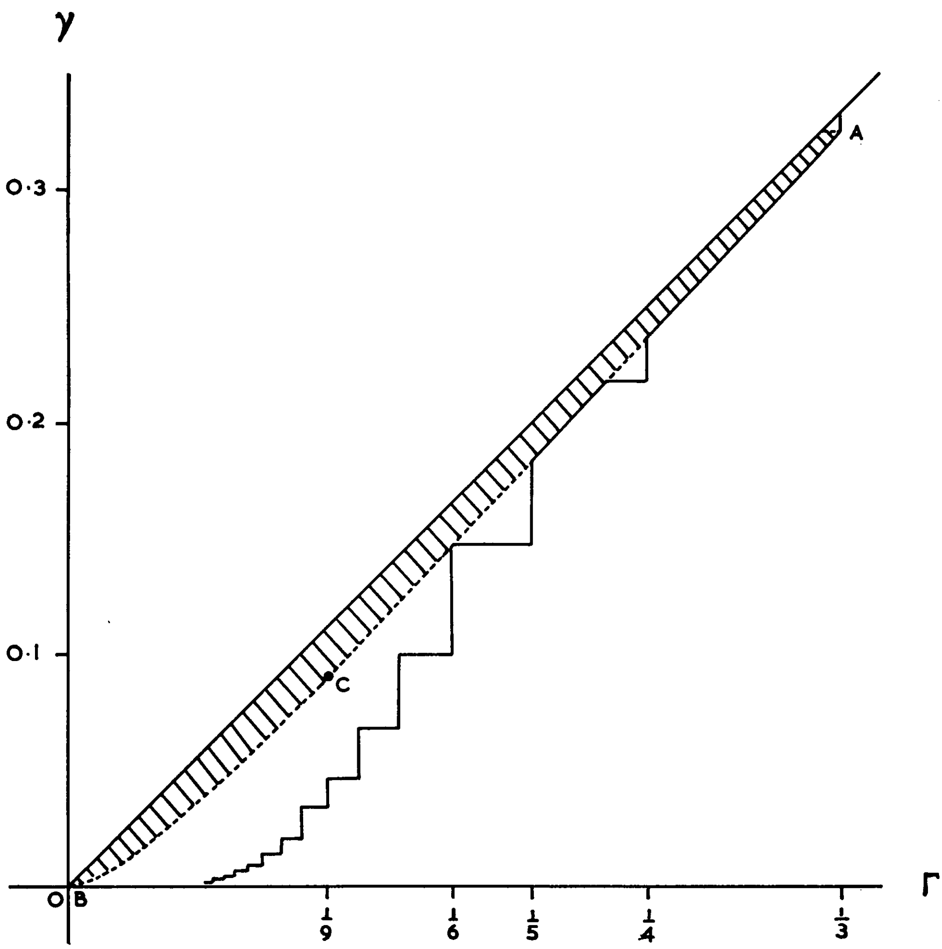

FIGURE 3

The part of the (Γ, γ)-diagram which can be obtained by the construction from certain special semigroups. The shaded area is what can be obtained by using three-state cyclic semigroups, and the triangular regions are those that can be obtained as described in the text when $n = 3, 4, \cdots, 16$.

corresponding to points on the boundary have been plotted. I have been unable to prove analytically that no point in the interior gives a value of (Γ, γ) outside this curve: several such points have been plotted and there has been no indication that this was the case.

If we have two semigroups with $(\Gamma, \gamma) = (\Gamma_1, \gamma_1)$ and (Γ_2, γ_2), respectively, then we can construct a third with $\Gamma = \min(\Gamma_1, \Gamma_2)$ and $\gamma = \min(\gamma_1, \gamma_2)$ by taking their direct sum. Thus, if we have one with $(\Gamma, \gamma) = (a, b)$, then we can find others with $(\Gamma, \gamma) = (x, b)$ where x is any number in the interval $[b, \min(\frac{1}{2}, a)]$ by using Kingman's examples described in section 4 of [2]. The region obtainable in this way by direct sums of three-state cyclic semigroups is (at least) that shaded in figure 2.

Looking for further examples, I considered the n-state semigroups with Q-matrices of the form

$$\begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{13}$$

The matrix $Q + I$ is a permutation matrix and so

$$[(Q + I)^m]_{i,i} = \begin{cases} 1 & \text{if } n \text{ divides } m. \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

Thus,

$$\begin{aligned} p_{i,i}(t) = (\exp(Qt))_{i,i} &= (\exp(-It)\exp\{(Q + I)t\})_{i,i} \\ &= e^{-t}\{1 + t^n/n! + t^{2n}/(2n)! + \cdots\} \end{aligned} \tag{15}$$

for all i. We shall assume for the moment that for $n \geq 3$ the first turning point of $p_{1,1}(t) = g(t)$ exists, and that at that point g attains its absolute minimum

TABLE I

N	Γ	γ
3	.33333	.32447
4	.25000	.21860
5	.20000	.14791
6	.16667	.10025
7	.14286	.06805
8	.12500	.04626
9	.11111	.03150
10	.10000	.02147
11	.09091	.01466
12	.08333	.01001
13	.07692	.00685
14	.07143	.00469
15	.06667	.00321
16	.06250	.00220
17	.05882	.00151
18	.05556	.00103
19	.05263	.00071

for the only time. Then at this point $p_{1,1}(t) < \lim_{t\to\infty} p_{1,1}(t) = 1/n$. The case where $n = 3$ has been dealt with above. For n, from 5 to 19 inclusive, the series could be truncated after three terms, and the value at the first minimum was computed on TITAN to four significant figures. (For $n \geq 10$, two terms would have been sufficient for this accuracy.) I am grateful to Mr. J. G. Basterfield for assistance in preparing the program. It was found that for $n = 4$ the contribution from the higher order terms was not negligible and the exact formula $g(t) = \frac{1}{4} + \frac{1}{2}e^{-t}\cos t + \frac{1}{4}e^{-2t}$ was preferred. The figures obtained are shown in table I.

In the case of this type of Q-matrix with $(\Gamma, \gamma) = (a, b)$, we can also obtain semigroups having $(\Gamma, \gamma) = (a, x)$ where x is any number in the interval $[b, a)$. We do this by combining $P(t)$ in a certain sense with the semigroup $R(t)$ which has Q-matrix

$$(16)\qquad S = \begin{pmatrix} -1 & 1/(n-1) & \cdots & 1/(n-1) \\ 1/(n-1) & -1 & \cdots & 1/(n-1) \\ \cdot & \cdot & \cdots & \cdot \\ 1/(n-1) & 1/(n-1) & \cdots & -1 \end{pmatrix},$$

and monotone transition functions;

$$(17)\qquad \begin{aligned} r_{i,i}(t) &= (1 + (n-1)e^{-nt/(n-1)})/n, \\ r_{i,j}(t) &= (1 - e^{-nt/(n-1)})/n, \qquad (i \neq j). \end{aligned}$$

The matrices Q and S commute, since S can be expressed as a polynomial in $I + Q$, and so ${}_{\lambda}P(t) = \exp\{\lambda Q + (1-\lambda)St\} = P(\lambda t)R((1-\lambda)t)$ and ${}_{\lambda}P(t)$ gives a Markov semigroup when $\lambda \in [0, 1]$. Now

$$(18)\qquad \begin{aligned} {}_{\lambda}p_{1,1}(t) &= \sum_j p_{1,j}(\lambda t) r_{j,1}((1-\lambda)t) \\ &= p_{1,1}(\lambda t)(1 + (n-1)e^{-(1-\lambda)nt/(n-1)})/n \\ &\quad + (1 - p_{1,1}(\lambda t))(1 - e^{-(1-\lambda)nt/(n-1)})/n \\ &= 1/n + (p_{1,1}(\lambda t) - 1/n)e^{-(1-\lambda)nt/(n-1)}. \end{aligned}$$

Suppose now that $\lambda \neq 0$. Then ${}_{\lambda}p'_{1,1}(t) = 0$ where

$$(19)\qquad -(1-\lambda)(p_{1,1}(\lambda t) - 1/n)n/(n-1) + \lambda p'_{1,1}(\lambda t) = 0.$$

Let the first point for which $p_{1,1}(t) = 1/n$ be t_0 and its first minimum turning point be t_1. Then $0 < t_0 < t_1$. Both terms on the left of (19) increase for $t \in [0, t_1/\lambda]$. The expression is negative for $t = t_0/\lambda$ and nonnegative for $t = t_1/\lambda$ and so there is a unique turning point, which we shall call $t_2(\lambda)$, for ${}_{\lambda}p_{1,1}(t)$ in the interval $(t_0/\lambda, t_1/\lambda]$ and this is the first turning point of ${}_{\lambda}p_{1,1}$. We show that ${}_{\lambda}p_{1,1}$ attains its absolute minimum at $t_2(\lambda)$. According to (18),

$$(20a)\qquad 1/n - {}_{\lambda}p_{1,1}(t_2(\lambda)) = (1/n - p_{1,1}(\lambda t_2(\lambda)))e^{-(1-\lambda)nt_2(\lambda))/(n-1)}$$

and since there is no other turning point in $[t_2, t_1/\lambda]$, this is not less than

(20b) $$(1/n - p_{1,1}(t_1))e^{-(1-\lambda)nt_1/(n-1)\lambda}$$
$$> (1/n - p_{1,1}(s\lambda))e^{-(1-\lambda)ns/(n-1)} \quad \text{for all} \quad s > t_1/\lambda$$
$$= 1/n - {}_\lambda p_{1,1}(s) \quad \text{for all} \quad s > t_1/\lambda.$$

This is also true (with a weak inequality) for $s \in [0, t_1/\lambda]$ since we have shown that there is only one turning point in that interval.

If $\lambda \in (0, 1]$, the point at which (19) is satisfied varies continuously with λ and tends to infinity as $\lambda \to 0+$ since it is greater than t_0/λ. The value of ${}_\lambda p_{1,1}(t) = g(t)$ at its absolute minimum varies continuously taking all values in the interval $[p_{1,1}(t_1), 1/n)$. Therefore, all these values are possible for γ when $\Gamma = 1/n$.

By the construction discussed before we can now obtain (Γ, γ) anywhere within the triangular regions of figure 3. It is also clear that by this construction based on a genuinely n-state semigroup we could not hope to obtain points with $\Gamma > 1/n$, and certainly not points with $\Gamma > \frac{1}{2}$.

In conclusion, I sketch a proof of the fact that $p_{1,1}(t)$ attains its absolute minimum for the only time at its first turning point:

$$p_{n-j+2,1} = p_{1,j} = e^{-t}\left(\frac{t^{j-1}}{(j-1)!} + \frac{t^{n+j-1}}{(n+j-1)!} + \frac{t^{2n+j-1}}{(2n+j-1)!} + \cdots\right), \tag{21}$$

$(j = 1, \cdots, n)$ and $p_{i,j}(0) = \delta_{i,j}$. There is some $T > 0$ such that, in (0, T), $p_{1,n}(t) < p_{1,n-1}(t) < \cdots < p_{1,1}(t)$. To see this consider the derivatives at the origin. On $(0, T]$, $s(t) \equiv s_1(t) = p_{1,n}(t)$. The corollary to lemma 1 shows that $s(t)$ is monotone increasing. We shall need a lemma to show that $p_{1,n}$ intersects $p_{1,1}$ before it intersects any other $p_{i,j}$ after $t = 0$. Then it will follow that there exist numbers T_1, T_2 such that $0 < T_1 < T_2$ and $s(t) = p_{1,n}(t)$ for $t \in [0, T_1]$ and $s(t) = p_{1,1}(t)$ for $t[T_1, T_2]$. The differential equations of the semigroup state that $p'_{i,j} = p_{i,j-1} - p_{i,j}$ (addition of suffixes being modulo n) and so, in particular, $p_{1,1}$ has its first turning point at T_1. Also $p_{1,1}$ is analytic and not constant, and so $s(t)$ is strictly increasing in $[T_1, T_2]$. Thus for all $u > T_1$, $p_{1,1}(u) > s(T_1)$. This is also true for $u \in [0, T_1)$ because T_1 is the first turning point.

LEMMA 4. *If $n \geq 3$, $p_{1,n}$ intersects $p_{1,1}$ before it intersects any other $p_{1,j}$ after $t = 0$.*

We assume this is not true and that $p_{1,n}$ intersects $p_{1,j}$ at T before it intersects $p_{1,1}$. Then there exists h, $j \leq h < n$ such that $p_{1,h}$ intersects $p_{1,h+1}$ before $p_{1,j}$ meets $p_{1,n}$. Choose k to be the largest integer with this property and t_0 the smallest $t > 0$ for which $p_{1,k}(t) = p_{1,k+1}(t)$. None of the functions $p_{1,l}(l > k)$ has intersected any other function in the interval $(0, t_0)$. The differential equations show that at t_0, $p_{1,k+1}$ has its first turning point. To the left of t_0, $p_{1,k} > p_{1,k+1}$ and so $p'_{1,k}(t_0) \leq 0$. This implies that at some point in $(0, t_0]$, $p_{1,k}$ has had a turning point. Let t_1 be the smallest such turning point. We can work backwards in this way until we reach t_{k-1} where $p_{1,2}$ has its first turning point.

Now $p_{1,1}$ is decreasing throughout $[0, T]$ because it has never intersected $p_{1,n}$ and its derivative has always been negative. At t_{k-1}, $p_{1,2} = p_{1,1}$ and $p_{1,2}$ is greater than $p_{1,1}$ in $(t_{k-1}, T]$ because otherwise it would have cut $p_{1,1}$ again with zero gradient and this is impossible. Similarly, $p_{1,r} > p_{1,r-1}$ in the interval $(t_{k-r+1}, T]$, $(r = 3, \cdots, k)$, and thus $p_{1,j} > p_{1,1}$ in $(t_{k-j+1}, T]$. This contradicts the initial assumption.

4. Acknowledgments

I am greatly indebted to Professor D. G. Kendall who first put forward this problem and has made many valuable suggestions throughout. I must also thank Miss S. E. Prior for assistance in computing the results in section 3 and for preparation of the diagrams. I hold a Graduate Studentship at Girton College and a Scientific Research Council Studentship.

REFERENCES

[1] K. L. Chung, *Markov Chains*, Berlin, Springer, 1960.

[2] D. G. Kendall, "On Markov Groups," to appear in the *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1966, Vol. II, Part II, pp. 165–173.