THE QUASISTATIONARY DISTRIBUTIONS OF HOMOGENEOUS QUASI-BIRTH-AND-DEATH PROCESSES

N.G. Bean[1], P.K. Pollett[2] and P.G. Taylor[1]

1 *Department of Applied Mathematics, University of Adelaide, South Australia 5005.*
2 *Department of Mathematics, University of Queensland, Queensland 4072.*

ABSTRACT

For evanescent Markov processes with a single transient communicating class, it is often of interest to examine the stationary probabilities that the process resides in the various transient states, conditional on absorption not having taken place. Such distributions are known as quasistationary distributions. In this paper we consider the determination of all the quasistationary distributions of a general homogeneous quasi-birth-and-death process (QBD). These distributions are shown to have a form analogous to the quasistationary distributions exhibited by birth-and-death processes. We discuss methods for the computation of these quasistationary distributions.

1 Introduction

Consider a discrete-time Markov chain $(X_n; n \in \mathbb{Z}_+)$ on a countable state space $\mathcal{S} = \{0, 1, \ldots\}$ with transition matrix P. Assume (X_n) has an absorbing state 0 and an irreducible and aperiodic communicating class $\mathcal{C} \equiv \mathcal{S} \setminus \{0\}$. We shall assume that absorption is certain from one (and then all) states $i \in \mathcal{C}$. Let T denote the time until absorption of the process.

For many evanescent Markov processes, T is usually very large. However, over any reasonable period of time, the process appears to reach an equilibrium. A quasistationary distribution, $\boldsymbol{\pi}$, is a stationary distribution of the process conditioned to stay in the recurrent class; that is, if $P(X_0 = j) = \pi_j$, $j \in \mathcal{C}$, then

$$P(X_n = j | T > n) = \pi_j, \qquad j \in \mathcal{C},$$

for all $n \geq 1$. In other words, conditional on the chain being in $\mathcal{C}$ the state probabilities do not vary with time.

A nontrivial, nonnegative row vector $\boldsymbol{m}(\beta)$ that satisfies

$$\boldsymbol{m}(\beta) = \beta \boldsymbol{m}(\beta) \widehat{P} \tag{1.1}$$

is called a β-invariant measure. Here, and throughout, $\widehat{P}$ denotes the restriction of P to $\mathcal{C}$. It is elementary to show that $\boldsymbol{\pi}$ is a quasistationary distribution if and only if, for some $\beta > 1$, it is a β-invariant measure, in which case

$$\beta^{-1} = 1 - \sum_{i \in \mathcal{C}} \pi_i p_{i0},$$

represents the probability (under the quasistationary distribution) that the process remains within the recurrent class at the next time step.

For each Markov process there is a maximum value of β for which a quasistationary distribution can exist. This critical parameter is called the convergence radius and is denoted by α. In certain circumstances an α-invariant measure may also have a limiting-conditional interpretation: that is,

$$\lim_{n \to \infty} P(X_n = j | X_0 = i, T > n) = \pi_j, \qquad j \in \mathcal{C},$$

no matter what the initial state i.

The convergence radius can be rigorously characterized as follows. For $z \in \mathbb{R}$, let $N_{ij}(z)$ be defined by

$$N_{ij}(z) = \sum_{n=0}^{\infty} z^n \widehat{P}_{ij}^{(n)}\ , \tag{1.2}$$

where $\widehat{P}_{ij}^{(n)}$ is the (i,j)th entry of $\widehat{P}^n$. Theorem 6.1 of Seneta (1981) states that, for a given value of z, either $N_{ij}(z)$ is finite for all (i,j) or $N_{ij}(z)$ is infinite for all (i,j). The convergence radius associated with $\widehat{P}$ is defined as

$$\alpha = \sup \{z : N_{ij}(z) \text{ is finite}\}\ .$$

There are very few substochastic chains for which a full quasistationary analysis is available. Historical exceptions were finite-state processes, the Galton-Watson branching process, simple birth-and-death chains and the work of Kyprianou (1972a) on GI/M/1 queues (see also Kyprianou (1972b) for an analysis under conditions of heavy traffic).

Recently, a notable advance was made by Kijima (1993) who gave an algebraic equation for the convergence radius of PH/PH/1 queues (in fact, more generally, for processes of M/G/1 and GI/M/1 type). In the queueing context considered in Kijima (1993), this equation can be solved by use of the Laplace-Stieltjes transform of the interarrival and service time distributions. Kijima also gave the form of the quasistationary distribution for the special cases of the M/PH/1 and PH/M/1 queues. This work was extended by Makimoto (1993) who gave an explicit representation of the quasistationary distribution for PH/PH/c queues in terms of solutions to a matrix equation. Makimoto did not, however, discuss methods of solution for this equation in the general case. For a nice survey of this area see Kijima and Makimoto (1995).

In Bean *et al.* (1995) the results of Kijima (1993) and Makimoto (1993) were extended by examining the limiting-conditional behaviour of general quasi-birth-and-death processes (QBDs), which includes the PH/PH/c queues as a subclass.

An algorithm for the explicit numerical computation of the convergence radius, α, and the limiting-conditional distribution was also presented.

We extend the results of Bean *et al.* (1995) by finding all the quasistationary distributions, not just the limiting-conditional distribution. In Bean *et al.* (1995) the limiting-conditional distribution is written in such a way that it is not an obvious generalization of the limiting-conditional distribution for a birth-and-death process. In contrast, here we present all the quasistationary distributions as natural extensions of the quasistationary distributions for ordinary birth-and-death processes. We also discuss methods for their computation.

The results of this paper can also be applied to *continuous-time* QBDs. For details, see Bean *et al.* (1995).

2 Absorbing QBDs and the Convergence Radius

In this section we summarise the results of Sections 3 and 4 of Bean *et al.* (1995) in order to define the convergence radius α and establish some fundamental concepts. These sections extend the matrix geometric theory of QBDs, as developed by Neuts (1981), to absorbing QBDs. Throughout, a matrix is termed finite (respectively infinite) if all its entries are finite (infinite).

Assume that (X_n) is a quasi-birth-and-death process. This can be regarded as a two-dimensional Markov chain with $\mathcal{C} = \{(k,j) : k \geq 1, 1 \leq j \leq M\}$ and whose transition matrices are of the block-partitioned form

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ A_2\boldsymbol{e} & A_1 & A_0 & 0 & 0 & \cdots \\ 0 & A_2 & A_1 & A_0 & 0 & \cdots \\ 0 & 0 & A_2 & A_1 & A_0 & \cdots \\ 0 & 0 & 0 & A_2 & A_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \text{and} \quad \widehat{P} = \begin{pmatrix} A_1 & A_0 & 0 & 0 & \cdots \\ A_2 & A_1 & A_0 & 0 & \cdots \\ 0 & A_2 & A_1 & A_0 & \cdots \\ 0 & 0 & A_2 & A_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Here the partitioning corresponds to distinguishing subsets of states called levels. Level k is defined by $l(k) = \{(k,j) : 1 \leq j \leq M\}$ for $k \geq 1$ and level 0 is the absorbing state 0. Throughout, $\boldsymbol{e}$ denotes a column vector of ones.

The equations (1.1) which define the β-invariant measures can now be written

$$\boldsymbol{m}_1(\beta) = \beta\left[\boldsymbol{m}_1(\beta)A_1 + \boldsymbol{m}_2(\beta)A_2\right], \tag{2.1}$$

$$\boldsymbol{m}_k(\beta) = \beta\left[\boldsymbol{m}_{k-1}(\beta)A_0 + \boldsymbol{m}_k(\beta)A_1 + \boldsymbol{m}_{k+1}(\beta)A_2\right], \quad k \geq 2, \tag{2.2}$$

where the M-vector $\boldsymbol{m}_k(\beta)$ is the restriction of $\boldsymbol{m}(\beta)$ to level k.

Let $N_{11}(\beta)$ denote the $M \times M$ matrix whose (i,j)th entry is $N_{(1,i)(1,j)}(\beta)$ as defined in (1.2). Define

$$R(\beta) = \beta A_0 N_{11}(\beta). \tag{2.3}$$

The entry $R_{ij}(\beta)$ can be interpreted as the expected total discounted reward for visits to state $(2,j)$ before returning to level 1, conditional on starting in state

$(1, i)$ with a discount factor β. In the rest of this paper, we shall consider only the situation where β is greater than or equal to one. Since (X_n) is homogeneous on all levels greater than 1, the interpretation given above also holds when levels 1 and 2 are replaced by levels k and $k+1$, respectively.

The following lemma can be shown in a similar way to Lemmas 1.2.2 and 1.2.3 in Neuts (1981).

Lemma 1 *If the matrix $R(\beta)$ is finite then it is the minimal nonnegative solution to the matrix-quadratic equation*

$$S = \beta\left[A_0 + SA_1 + S^2 A_2\right]. \tag{2.4}$$

Consider the solutions of equation (2.4). For $0 \le z \le 1$, let $\chi(z)$ be the maximal eigenvalue of the matrix

$$A(z) = A_0 + zA_1 + z^2 A_2 \tag{2.5}$$

and $\boldsymbol{u}(z)$ and $\boldsymbol{v}(z)$ the corresponding left and right eigenvectors normalized so that $\boldsymbol{u}(z)\boldsymbol{e} = 1 = \boldsymbol{u}(z)\boldsymbol{v}(z)$.

By Theorem 1.3.2 of Neuts (1981), the condition that absorption is certain is equivalent to the condition that $\chi'(1^-) > 1$ and $\chi(0) > 0$.

Let $\eta(\beta)$ be the maximal eigenvalue of $R(\beta)$; note that $\boldsymbol{u}(\eta(\beta))$ is the associated left eigenvector and we denote by $\boldsymbol{w}(\eta(\beta))$ the associated right eigenvector. In Kijima (1993) (see also Bean *et al.* (1995)) the following theorem was established.

Theorem 2 *The convergence radius α associated with (X_n) is given by*

$$\alpha = \left[\boldsymbol{u}(z_0)\left[A_1 + 2z_0A_2\right]\boldsymbol{v}(z_0)\right]^{-1}, \tag{2.6}$$

where $\eta(\alpha) = z_0$ is the unique solution to

$$\chi'(z)z = \chi(z) \tag{2.7}$$

in the interval (0,1), and $\boldsymbol{u}(z_0), \boldsymbol{v}(z_0)$ are the Perron-Frobenius left and right eigenvectors of $A(z_0)$ respectively.

3 The One-dimensional Manifold

In this section we change emphasis from the parameter β to z. In order to do this, we investigate the one-dimensional manifold of solutions to

$$z = \beta\chi(z), \tag{3.1}$$

which is the equation relating equations (2.4) and (2.5). Figure 1 shows the form of the manifold in the general case.

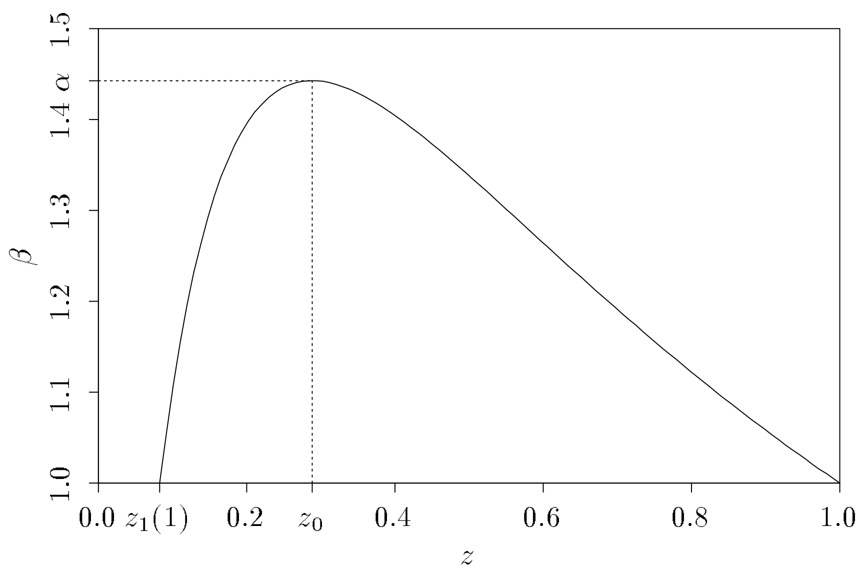

Figure 1: Graph of the one-dimensional manifold of solutions to equation (3.1)

For $\beta < \alpha$, there are always two points of solution and we label these as $z_1(\beta)$ and $z_2(\beta)$ with the understanding that $z_1(\cdot) < z_2(\cdot)$. Note that $z_1(\alpha) = z_2(\alpha) = z_0$.

For any value of $z \in [z_1(1), 1]$ it is possible to define a unique value of β, say $\beta(z)$, for which equation (3.1) is obeyed. Thus $\beta(z)$ is defined by the function

$$\beta(z) = \frac{z}{\chi(z)}. \tag{3.2}$$

Having defined $\beta(z)$ we can now define $R(\beta(z))$, via equation (2.3), for all $z \in [z_1(1), z_0]$ and we henceforth write this as $R_1(z)$. Note that $R_1(z)$ is the minimal nonnegative solution to equation (2.4) for $\beta = \beta(z)$ and has maximal eigenvalue z. In the next section we show that for $y \in (z_0, 1]$ there is a nonnegative solution $R_2(y)$ to equation (2.4) for $\beta = \beta(y)$ with maximal eigenvalue y.

It is very important to note that $R_1(z)$ is the matrix that arises probabilistically as defined in equation (2.3), due to the minimality requirement. In contrast, $R_2(z)$ has no such interpretation.

4 The Quasistationary Distributions

In this section we show that the distribution $\boldsymbol{m}(\beta) = (\boldsymbol{m}_1(\beta), \ldots)$ given by

$$\boldsymbol{m}_j(\beta) = c\boldsymbol{x}\left[R_2(z_2(\beta))^j - R_1(z_1(\beta))^j\right] , \tag{4.1}$$

with c a normalising constant, is a β-invariant measure, and hence quasistationary distribution, for (X_n) when $\beta < \alpha$. We also show that $\boldsymbol{m}(\alpha) = (\boldsymbol{m}_1(\alpha), \ldots)$ given by

$$\boldsymbol{m}_j(\alpha) = c\boldsymbol{x}\frac{d}{dz}\left[R_1(z)^j\right]_{z=z_0} , \tag{4.2}$$

with c a normalising constant, is the α-invariant measure, and hence quasistationary and limiting-conditional distribution, for (X_n).

The major contribution of this paper is that this is the first presentation of the complete set of β-invariant measures. Further, these distributions are the obvious matrix analogues of the well-known scalar results for the ordinary birth-and-death process. This is in contrast to the form of the α-invariant measure presented in Makimoto (1993) and Bean *et al.* (1995).

4.1 Calculation of the matrix $R_2(y)$

First define $y(x) = z_2(\beta(x))$ for $x \in [z_1(1), z_0]$ and $z(x) = z_1(\beta(x))$ for $x \in [z_0, 1]$. For simplicity, we usually write y instead of $y(x)$ and z instead of $z(x)$, where there can be no confusion. Thus z and y are $z_1(\beta)$ and $z_2(\beta)$, respectively, where $\beta = \beta(z) = \beta(y)$ and we will freely interchange the notation when convenient.

Throughout this section, we assume that $z \in [z_1(1), z_0)$ and hence $y = y(z) \in (z_0, 1]$.

Lemma 3 *The matrix*

$$S(y) = R_1(z) + \boldsymbol{w}(z)\boldsymbol{u}(y)\left[yI - R_1(z)\right], \tag{4.3}$$

where $\boldsymbol{u}(y)$ is normalised so that $\boldsymbol{u}(y)\boldsymbol{w}(z) = 1$, has the following spectral properties:

1. *The maximal eigenvalue of $S(y)$ is y.*
2. *The left and right eigenvectors of $S(y)$ associated with the eigenvalue y are given by $\boldsymbol{u}(y)$ and $\boldsymbol{w}(z)$, respectively.*
3. *Otherwise the eigenvalues of $S(y)$ and $R_1(z)$ are identical.*
4. *For eigenvalues of $S(y)$ not equal to y, the associated left eigenvectors of $S(y)$ and $R_1(z)$ are identical. The right eigenvectors are, in general, different from those of $R_1(z)$.*

Proof: All these properties are simple consequences of the definition of $\boldsymbol{u}(y)$, $\boldsymbol{w}(z)$ and $S(y)$ on recalling the normalization condition for $\boldsymbol{u}(y)$. ∎

Lemma 4 *For all $j \geq 1$,*

$$S^j(y) = R_1^j(z) + \boldsymbol{w}(z)\boldsymbol{u}(y)\left[y^jI - R_1^j(z)\right]. \tag{4.4}$$

Proof: The proof follows a simple mathematical induction argument using the spectral properties of $S(y)$ given in Lemma 3. ∎

Theorem 5 *A nonnegative solution to equation (2.4) for $\beta = \beta(y) = \beta(z)$ with maximal eigenvalue y is given by*

$$R_2(y) = S(y) = R_1(z) + \boldsymbol{w}(z)\boldsymbol{u}(y)\left[yI - R_1(z)\right]. \tag{4.5}$$

Proof: First we shall prove that $S(y)$ obeys equation (2.4). Consider the right-hand side of equation (2.4).

$$\begin{aligned}
&\beta(z)\left[A_0 + S(y)A_1 + S^2(y)A_2\right] \\
&\quad = \beta(z)\left[A_0 + R_1(z)A_1 + R_1^2(z)A_2\right] \\
&\qquad + \beta(z)\boldsymbol{w}(z)\boldsymbol{u}(y)\left[\left(yI - R_1(z)\right)A_1 + \left(y^2 I - R_1^2(z)\right)A_2\right], \\
&\quad = R_1(z) + \boldsymbol{w}(z)\boldsymbol{u}(y)\beta(z)\left[\left(A_0 + yA_1 + y^2A_2\right) - \left(A_0 + R_1(z)A_1 + R_1^2(z)A_2\right)\right], \\
&\quad = R_1(z) + \boldsymbol{w}(z)\boldsymbol{u}(y)\left[yI - R_1(z)\right], \\
&\quad = S(y),
\end{aligned}$$

since $\beta(z) = \beta(y)$, $\chi(y) = y/\beta(y)$ and $R_1(z)$ obeys equation (2.4) when $\beta = \beta(z)$.

Next, we shall prove that $\boldsymbol{u}(y)(yI - R_1(z))$ is a nonnegative vector. To prove this result we follow a very similar argument to that presented in the proof of Theorem 5 in Bean *et al.* (1995), as follows. Let R_0 be the zero matrix and define

$$R_{n+1} = \beta(z)\left[A_0 + R_nA_1 + R_n^2A_2\right].$$

It is easy to show by induction that the sequence $\{R_n\}$ is nondecreasing. It is also possible to show, again by induction, that

$$\boldsymbol{u}(y)R_n \leq y\boldsymbol{u}(y), \tag{4.6}$$

because $\beta(z) = \beta(y)$ and $\boldsymbol{u}(y)$ is the strictly positive left eigenvector of $A(y)$ with eigenvalue $y/\beta(y)$. Since Theorem 5 of Bean *et al.* (1995) shows that the sequence $\{R_n\}$ converges monotonically to $R_1(z)$, we can conclude that

$$\boldsymbol{u}(y)R_1(z) \leq y\boldsymbol{u}(y), \tag{4.7}$$

and hence that $S(y) \geq R_1(z) \geq 0$ (where all inequalities are treated elementwise).

Therefore, $S(y)$ is a nonnegative solution to equation (2.4) with maximal eigenvalue y. We take $R_2(y) = S(y)$. ■

4.2 The quasistationary distribution for each $\beta < \alpha$

Theorem 6 *If $\beta < \alpha$, then for any $\boldsymbol{x}$ such that $\boldsymbol{x}\boldsymbol{w}(z_1(\beta)) \neq 0$, the distribution $\boldsymbol{m}(\beta) = (\boldsymbol{m}_1(\beta), \ldots)$ given by*

$$\boldsymbol{m}_j(\beta) = c\boldsymbol{x}\left[R_2(z_2(\beta))^j - R_1(z_1(\beta))^j\right], \tag{4.8}$$

with c a normalising constant, is a β-invariant measure, and hence quasistationary distribution, for (X_n). Moreover, $\boldsymbol{m}_j(\beta)$ has the more compact form

$$\boldsymbol{m}_j(\beta) = k\boldsymbol{u}(z_2(\beta))\left[z_2(\beta)^j I - R_1^j(z_1(\beta))\right], \tag{4.9}$$

with $k = c\boldsymbol{x}\boldsymbol{w}(z_1(\beta))$.

Proof: We first show that $\boldsymbol{m}(\beta)$ obeys equations (2.1) and (2.2). Consider the right-hand side of equation (2.2) for $k \geq 2$.

$$\begin{aligned}
&\beta\left[\boldsymbol{m}_{k-1}(\beta)A_0 + \boldsymbol{m}_k(\beta)A_1 + \boldsymbol{m}_{k+1}(\beta)A_2\right] \\
&\quad = \boldsymbol{x}R_2(z_2(\beta))^{k-1}\beta\left[A_0 + R_2(z_2(\beta))A_1 + R_2^2(z_2(\beta))A_2\right] \\
&\qquad -\boldsymbol{x}R_1(z_1(\beta))^{k-1}\beta\left[A_0 + R_1(z_1(\beta))A_1 + R_1^2(z_1(\beta))A_2\right], \\
&\quad = \boldsymbol{x}\left(R_2(z_2(\beta))^k - R_1(z_1(\beta))^k\right), \\
&\quad = \boldsymbol{m}_k(\beta),
\end{aligned}$$

since $R_2(z_2)$ and $R_1(z_1)$ obey equation (2.4). Consider now the right-hand side of equation (2.1).

$$\begin{aligned}
&\beta\left[\boldsymbol{m}_1(\beta)A_1 + \boldsymbol{m}_2(\beta)A_2\right] \\
&\quad = \boldsymbol{x}\beta\left[R_2(z_2(\beta))A_1 + R_2^2(z_2(\beta))A_2\right] - \boldsymbol{x}\beta\left[R_1(z_1(\beta))A_1 + R_1^2(z_1(\beta))A_2\right], \\
&\quad = \boldsymbol{x}\beta\left[A_0 + R_2(z_2(\beta))A_1 + R_2^2(z_2(\beta))A_2\right] \\
&\qquad -\boldsymbol{x}\beta\left[A_0 + R_1(z_1(\beta))A_1 + R_1^2(z_1(\beta))A_2\right], \\
&\quad = \boldsymbol{x}\left(R_2(z_2(\beta)) - R_1(z_1(\beta))\right), \\
&\quad = \boldsymbol{m}_1(\beta),
\end{aligned}$$

since $R_2(z_2)$ and $R_1(z_1)$ obey equation (2.4).

The only requirement on the choice of the vector $\boldsymbol{x}$ is that $\boldsymbol{m}_j(\beta)$ should be nonnegative for all $j \geq 1$. This raises the question of how many distinct solutions to equations (2.1) and (2.2) of the form in given in equation (4.8) there are for each value of β?

It follows from equation (4.7) that $R_2(z_2)^j - R_1(z_1)^j = \boldsymbol{w}(z_1)\boldsymbol{u}(z_2)\left(z_2^j I - R_1^j(z_1)\right)$ is nonnegative for all $j \geq 1$. It is also a rank one matrix. Therefore, there is only one such solution for each value of β. Any choice of vector $\boldsymbol{x}$ such that $\boldsymbol{x}\boldsymbol{w}(z_1(\beta)) \neq 0$ will realise this β-invariant measure on appropriate normalisation. Equation (4.8) reduces to equation (4.9) by letting $k = c\boldsymbol{x}\boldsymbol{w}(z_1(\beta))$. ■

4.3 Calculation of the derivative of the matrix $R_1(z)$

When $\beta = \alpha$, we cannot use the form of the β-invariant measures that we found earlier because $z_1(\alpha) = z_2(\alpha)$ and so $R_2(z_2)$ is identical to $R_1(z_1)$. In order to determine the α-invariant measure, proposed in equation (4.2), we need to calculate the derivative of $R_1(z)$. In fact we do not necessarily calculate the derivative, instead we find a solution to the equation that the derivative must obey, which is sufficient for our purposes. Therefore, we shall abuse the notation and label this solution as the derivative.

Lemma 7 *The derivative of $R_1(z)$ must obey*

$$\frac{\beta'(z)}{\beta(z)} R_1(z) + \beta(z)\Big[TA_1 + (TR_1(z) + R_1(z)T)A_2\Big] - T = 0\ , \tag{4.10}$$

where $\beta'(z) = \dfrac{\chi(z) - z\chi'(z)}{\chi^2(z)}$.

Proof: First, $\beta(z)$ is defined in equation (3.2) as $\beta(z) = \frac{z}{\chi(z)}$. Therefore, the form of $\beta'(z)$ follows trivially. Also, $R_1(z)$ is defined to be the minimal nonnegative solution to

$$S = \beta\left[A_0 + SA_1 + S^2A_2\right]. \tag{4.11}$$

Differentiating this equation with respect to z, while remembering the functional dependencies, completes the proof of the lemma. ■

Throughout the remainder of this section we assume that $\beta = \alpha$. Recall that $z_0 = z_1(\alpha) = z_2(\alpha)$ and that $\beta(z_0) = \alpha$. Note that α is defined by the fact that $\chi(z_0) = z_0\chi'(z_0)$ and hence that $\beta'(z_0) = 0$.

Lemma 8 *A solution to equation (4.10) when $z = z_0$ is given by*

$$T = \boldsymbol{w}(z_0)\Big(\boldsymbol{u}(z_0) + \boldsymbol{u}'(z_0)\Big(z_0 I - R_1(z_0)\Big)\Big)\ , \tag{4.12}$$

where $\boldsymbol{u}'(z_0)$ is the solution to

$$\boldsymbol{b}\left[z_0^2A_2 + z_0(A_1 - \frac{1}{\alpha}I) + A_0\right] = -\boldsymbol{u}(z_0)\left[2z_0A_2 + (A_1 - \frac{1}{\alpha}I)\right] \tag{4.13}$$

subject to $\boldsymbol{be} = 0$.

Proof: The existence of the vector $\boldsymbol{u}'(z_0)$ is shown in Lemma 9 of Bean *et al.* (1995). It is a simple matter to substitute the expression for T into equation (4.10) to complete the proof. ■

Henceforth, we shall label T as $R_1'(z_0)$, whilst being aware that this may be an abuse of notation.

4.4 The quasistationary distribution for $\beta = \alpha$

Theorem 9 *For any $\boldsymbol{x}$ such that $\boldsymbol{x}R_1'(z_0) \neq \boldsymbol{0}$, the distribution $\boldsymbol{m}(\alpha) = (\boldsymbol{m}_1(\alpha), \ldots)$ given by*

$$\boldsymbol{m}_j(\alpha) = c\boldsymbol{x}\frac{d}{dz}\left[R_1(z)^j\right]_{z=z_0}\ , \tag{4.14}$$

with c a normalising constant, is the α-invariant measure, and hence quasistationary and limiting-conditional distribution, for (X_n). Moreover, $\boldsymbol{m}_j(\alpha)$ has the more explicit form

$$\boldsymbol{m}_j(\alpha) = k\left(z_0^j \boldsymbol{u}'(z_0) + j z_0^{j-1}\boldsymbol{u}(z_0) - \boldsymbol{u}'(z_0)R_1(\alpha)^j\right) , \tag{4.15}$$

with $k = c\boldsymbol{x}\boldsymbol{w}(z_0)$.

Proof: Note that

$$\frac{d}{dz}\left[R_1(z)^j\right]_{z=z_0} = \sum_{\ell=1}^{j} R_1^{\ell-1}(z_0)R_1'(z_0)R_1^{j-\ell}(z_0). \tag{4.16}$$

To prove that $\boldsymbol{m}(\alpha)$ is an α-invariant measure, we simply need to show that it obeys equation (2.1) and (2.2) when $\beta = \alpha$. Consider the right-hand side of equation (2.2) for $k \geq 2$.

$$\begin{aligned}
&\alpha\left[\boldsymbol{m}_{k-1}(\alpha)A_0 + \boldsymbol{m}_k(\alpha)A_1 + \boldsymbol{m}_{k+1}(\alpha)A_2\right] \\
&\quad = \boldsymbol{x}\sum_{\ell=1}^{k-1} R_1^{\ell-1}(z_0)R_1'(z_0)R_1^{k-1-\ell}(z_0)\alpha\left[A_0 + R_1(z_0)A_1 + R_1^2(z_0)A_2\right] \\
&\quad\quad -\boldsymbol{x}R_1(z_0)^{k-1}\alpha\left[R_1'(z_0)A_1 + R_1'(z_0)R_1(z_0)A_2 + R_1(z_0)R_1'(z_0)A_2\right], \\
&\quad = \boldsymbol{x}\sum_{\ell=1}^{k-1} R_1^{\ell-1}(z_0)R_1'(z_0)R_1^{k-\ell}(z_0) + \boldsymbol{x}R_1(z_0)^{k-1}R_1'(z_0), \\
&\quad = \boldsymbol{x}\sum_{\ell=1}^{k} R_1^{\ell-1}(z_0)R_1'(z_0)R_1^{k-\ell}(z_0), \\
&\quad = \boldsymbol{m}_k(\alpha),
\end{aligned}$$

by the definition of $R_1(z_0)$ and $R_1'(z_0)$. Consider now the right-hand side of equation (2.1).

$$\begin{aligned}
\beta\left[\boldsymbol{m}_1(\alpha)A_1 + \boldsymbol{m}_2(\alpha)A_2\right] &= \boldsymbol{x}\beta\left[R_1'(z_0)A_1 + R_1'(z_0)R_1(z_0)A_2 + R_1(z_0)R_1'(z_0)A_2\right], \\
&= \boldsymbol{x}R_1'(z_0), \\
&= \boldsymbol{m}_1(\alpha),
\end{aligned}$$

again by the definition of $R_1'(z_0)$.

It is easy to show that $R_1'(z_0)$ is a rank one matrix. Therefore, there is a unique solution to equations (2.1) and (2.2) of the form in given in equation (4.14). This solution will be realised for all $\boldsymbol{x}$ such that $\boldsymbol{x}R_1'(z_0) \neq \boldsymbol{0}$.

We now show that equation (4.14) reduces to equation (4.15). Simple substitution is sufficient to complete the proof on recalling that $\boldsymbol{u}(z_0)$ and $\boldsymbol{w}(z_0)$ are the left and right eigenvectors of $R_1(z_0)$, respectively, associated with the eigenvalue z_0.

$$\boldsymbol{m}_j(\alpha) = c\boldsymbol{x}\frac{d}{dz}\left[R_1(z)^j\right]_{z=z_0},$$

$$
\begin{aligned}
&= \boldsymbol{c}\boldsymbol{x} \sum_{\ell=1}^{j} R_1^{\ell-1}(z_0) R_1'(z_0) R_1^{j-\ell}(z_0), \\
&= \boldsymbol{c}\boldsymbol{x} \sum_{\ell=1}^{j} R_1^{\ell-1}(z_0)\boldsymbol{w}(z_0)\left(\boldsymbol{u}(z_0) + \boldsymbol{u}'(z_0)\left(z_0 I - R_1(z_0)\right)\right) R_1^{j-\ell}(z_0), \\
&= \boldsymbol{c}\boldsymbol{x} \sum_{\ell=1}^{j} z_0^{\ell-1}\boldsymbol{w}(z_0)\left(\boldsymbol{u}(z_0) + \boldsymbol{u}'(z_0)\left(z_0 I - R_1(z_0)\right)\right) R_1^{j-\ell}(z_0), \\
&= \boldsymbol{c}\boldsymbol{x} \sum_{\ell=1}^{j} z_0^{\ell-1}\boldsymbol{w}(z_0)\left(\boldsymbol{u}(z_0)R_1^{j-\ell}(z_0) + \boldsymbol{u}'(z_0) z_0 R_1^{j-\ell}(z_0) - \boldsymbol{u}'(z_0) R_1^{j-\ell+1}(z_0)\right), \\
&= \boldsymbol{c}\boldsymbol{x}\boldsymbol{w}(z_0)\left(j\boldsymbol{u}(z_0)z_0^{j-1} + \boldsymbol{u}'(z_0)\sum_{\ell=1}^{j}\left[z_0^{\ell} R_1^{j-\ell}(z_0) - z_0^{\ell-1} R_1^{j-\ell+1}(z_0)\right]\right), \\
&= k\left(j\boldsymbol{u}(z_0)z_0^{j-1} + \boldsymbol{u}'(z_0)z_0^{j} - \boldsymbol{u}'(z_0)R_1^{j}(z_0)\right),
\end{aligned}
$$

by letting $k = \boldsymbol{c}\boldsymbol{x}\boldsymbol{w}(z_0)$.

Theorem 10 of Bean *et al.* (1995) shows that the expression for $\boldsymbol{m}_j(\alpha)$ given in equation (4.15) is nonnegative for all $j \geq 1$. Hence, the expression for $\boldsymbol{m}_j(\alpha)$ given in equation (4.14) must also be nonnegative for all $j \geq 1$.

Therefore, the solution to equations (2.1) and (2.2) proposed in equation (4.14) is the unique α-invariant measure, and hence quasistationary and limiting-conditional distribution, for (X_n). ■

5 Computation of the Distributions

In this section we briefly indicate the steps involved in the computation of the quasistationary distributions. We assume that α and z_0 have already been computed according to Theorem 2, for more computational details see Section 6(i) of Bean *et al.* (1995).

5.1 If $\beta < \alpha$

In this situation we need to calculate $\boldsymbol{m}_j(\beta)$ according to equation (4.8). That is, we need to find $R_1(z_1(\beta))$ and $R_2(z_2(\beta))$. First we need to determine both $z_1(\beta)$ and $z_2(\beta)$. These can be evaluated by performing bisection searches on the intervals $[z_1(1), z_0]$ and $[z_0, 1]$, respectively, to determine the two solutions to $\chi(z) = \frac{z}{\beta(z)}$ (this is a similar procedure to that of finding z_0 and α). At the same time it is most efficient also to generate $\boldsymbol{u}(z_2(\beta))$. Then we must find $R_1(z_1(\beta))$. This can be evaluated using the algorithm explained in Theorem 11 of Bean *et al.* (1995). It is then easy to calculate $\boldsymbol{w}(z_1(\beta))$ by elementary methods and $R_2(z_2(\beta))$ using Theorem 5. Finally, Theorem 6 can be applied to generate the required quasistationary distribution, where it is computationally easier to use equation (4.9) rather than equation (4.8).

5.2 If $\beta = \alpha$

In this situation we need to calculate $\boldsymbol{m}_j(\alpha)$ according to equation (4.14). That is, we need to find $R_1(z_0)$ and $R_1'(z_0)$. We assume that z_0 has already been evaluated at the same time as α. Again, $R_1(z_0)$ can be evaluated using the algorithm explained in Theorem 11 of Bean *et al.* (1995). It is then easy to calculate $\boldsymbol{u}(z_0)$ by elementary methods, $\boldsymbol{u}'(z_0)$ as in equation (4.13) (more details are given in Section 6(ii) of Bean *et al.* (1995)) and $R_1'(z_0)$ using Lemma 8. Finally, Theorem 9 can be applied to generate the required quasistationary (and in fact limiting-conditional) distribution. In fact this is not the best numerical method for evaluating the limiting-conditional distribution. From a computational point of view, it is better to use the explicit representation given in equation (4.15) instead of that given in equation (4.14).

Acknowledgements

The financial support of the Australian Research Council through grants A69532348 and A69022762 is gratefully acknowledged.

References

N.G. Bean, L. Bright, G. Latouche, C.E.M. Pearce, P.K. Pollett, and P.G. Taylor (1995). The quasistaionary behaviour of quasi-birth-and-death processes. To appear in Annals of Applied Probability and available at http://www.maths.uq.oz.au/~pkp/papers/qbds.ps and in html form at http://www.maths.uq.oz.au/~pkp/papers/qbds/qbds.html

M. Kijima (1993). Quasi-stationary distributions of phase-type queues. *Math. Oper. Res.*, **18**(2):423–437.

M. Kijima and N. Makimoto (1995). Quasi-stationary distributions of Markov chains arising from queueing processes: A survey. Research reports on Mathematical and Computing Sciences, Series B: Operations Research B-304, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan.

E. K. Kyprianou (1972a). On the quasi-stationary distributions of the GI/M/1 queues. *J. Appl. Prob.*, **9**:117–128.

E. K. Kyprianou (1972b). The quasi-stationary distributions of queues in heavy traffic. *J. Appl. Prob.*, **9**:821–831.

N. Makimoto (1993). Quasi-stationary distributions in a PH/PH/c queue. *Stochastic Models*, **9**(2):195–212.

M. F. Neuts (1981). *Matrix Geometric Solutions in Stochastic Models.* The John Hopkins University Press, Baltimore.

E. Seneta (1981). *Non-negative Matrices and Markov Chains.* Springer-Verlag, New York.