SOME RESULTS FOR QUASI-STATIONARY DISTRIBUTIONS OF BIRTH–DEATH PROCESSES

MASAAKI KIJIMA,* The University of Tsukuba, Tokyo
E. SENETA,** University of Sydney

Abstract

Quasi-stationary distributions are considered in their own right, and from the standpoint of finite approximations, for absorbing birth–death processes. Results on convergence of finite quasi-stationary distributions and a stochastic bound for an infinite quasi-stationary distribution are obtained. These results are akin to those of Keilson and Ramaswamy (1984). The methodology is a synthesis of Good (1968) and Cavender (1978).

INVARIANT MEASURES; ORTHOGONAL POLYNOMIALS; STOCHASTIC BOUNDS; TRUNCATIONS

1. Introduction

Consider a birth–death process on the non-negative integers \( \{0, 1, 2, \cdots \} \) with birth rates \( \lambda_i \) and death rates \( \mu_i \), strictly positive for \( i \geq 1 \), and \( \mu_0 = 0 \). Let us denote the process where \( \lambda_0 > 0 \) by \( \{N(t)\}, t \geq 0 \). Under conditions on the rates which make the process well defined, with a stochastic transition matrix, ergodic and the event \( \{T(i) < \infty\} \) certain, where \( T(i) = \inf\{t \mid N(t) = 0; N(0) = i\} \) is the first-passage time from \( i \geq 1 \) to zero, \( T(i) \) may still be sufficiently large to allow \( N(t) \) to settle down to a statistical equilibrium within this period. This leads to the questions of existence of a limiting distribution

\[
(1.1) \quad \lim_{t \to \infty} \Pr[ N(t) = j \mid T(i) > t ]
\]

called a quasi-stationary distribution, its independence of initial state, and its characterization by quasi-stationary equations, in a manner akin to these questions for the ergodic distribution vector.

The question may be reformulated by considering the absorbing birth-and-death process \( \{X(t)\}, t \geq 0 \), where \( \lambda_0 = 0 \), under conditions which make the process well-
defined, with stochastic transition matrix \( P(t) = \{ p_{ij}(t) \}_{i,j \geq 0} \) at each time point, and absorption from any initial state certain. Then, (1.1) takes on the form

\[
\lim_{t \to \infty} \Pr[X(t) = j \mid X(t) > 0] = \lim_{t \to \infty} \frac{p_{ij}(t)}{1 - p_{i0}(t)}.
\]

The ideas of conditional limit distributions of this kind have been extensively discussed for finite Markov chains (e.g. Darroch and Seneta (1965)) and countably infinite Markov chains (Seneta and Vere-Jones (1966)) and finite (Mandl (1960), Darroch and Seneta (1967)) and countably infinite Markov processes (Vere-Jones (1969), Pollett (1988), (1989)). Whereas in the finite case the theory is relatively complete, in the countably infinite case in continuous time (which is our setting here), the general theory is heavily dependent on the assumption of \( \lambda \)-positivity, which establishes a parallelism with the finite case.

For countably infinite continuous-time Markov processes with special structure such as birth-and-death processes, one might expect more general results. For the simple case where \( \lambda_i = \lambda, \mu_i = \mu, i \geq 1, \lambda_0 = 0 \), it was shown that, when \( \mu > \lambda \), (1.2) defines a probability distribution independent of initial state in Seneta (1966). Subsequently, in a little-known but important paper, Good (1968) obtained conditions under which this result holds in general. There is an error in the proof of his main theorem, which, however, can be corrected (Kesten (1970)). Unaware of Good’s work, Cavender (1978) produced a study of the quasi-stationary equations which might be associated with such a limiting distribution. This study focused on the intensity matrix only, under virtually no assumptions about the \( \lambda_i \)'s and \( \mu_i \)'s in regard to the resultant Markov matrix \( P(t) \). Its results were therefore also somewhat unspecific, although its approach is interesting; we summarize it in our Section 2.

In the present paper, we synthesize the approach and results of Good (1968) and Cavender (1978), using results given in van Doorn (1985), (1987), to generalize the conclusions of both papers, and to produce the analogue for our setting of the main result of Keilson and Ramaswamy (1984) about convergence of quasi-stationary distributions.

In what follows we assume \( \lambda_0 = 0 \) unless stated otherwise, and denote by \( A \) the intensity matrix confined to the non-absorbing states \( \{1, 2, \ldots\} \), and denote by \( A_N \) the \((N \times N)\) northwest corner truncation (section) of the matrix \( A \), for \( N \geq 1 \).

### 2. Quasi-stationary measures

Consider a row vector \( \{ q_i \}, i \geq 1 \), satisfying \( q_i \geq 0 \) for each \( i \geq 1 \), and satisfying the system for \( i \geq 1 \)

\[
\lambda_{i-1}q_{i-1} - (\lambda_i + \mu_i)q_i + \mu_{i+1}q_{i+1} + \mu_{i+1}q_i q_i = 0
\]

with \( q_0 = 0 \). We shall call such vectors \textit{quasi-stationary measures}. Cavender (1978), Theorem 2, shows that there is a finite number \( a \geq 0 \) such that all quasi-stationary
measures are given by the vector \( \{ f_i(x) \} \), \( i \geq 1 \), for \( 0 \leq x \leq a \), where \( (f_i(x)) \), \( i \geq 1 \), is a sequence of polynomials with \( f_i(x) = x \). Thus all quasi-stationary measures are parameterized by the finite entry \( q_i \), \( 0 \leq q_i \leq a \).

Cavender (1978), Section 3, also considers the related ‘approximative’ birth-and-death process on the state space \( \{0, 1, \cdots, N\} \) with \( \mu_i \), \( i = 0, \cdots, N \), and \( \lambda_i \) for \( i = 0, \cdots, N - 1 \) as before, and \( \lambda_N = 0 \). Then, there is a unique non-zero row vector \( \{ q^{(N)}_i \} \), \( i = 1, \cdots, N \), satisfying (2.1) with \( \mu_{N+1} = q^{(N)}_{N+1} = 0 \). This vector satisfies \( q^{(N)}_i > 0 \), \( i = 1, \cdots, N \). \( \Sigma^{(N)}_i q^{(N)}_i = 1 \), \( \mu_i q^{(N)}_i \) is the smallest eigenvalue of \( -B_N \), where \( B_N \) is the intensity matrix confined to the non-absorbing states \( \{1, \cdots, N\} \) of this finite birth-and-death process, and \( q^{(N)}_i = f_i(q^{(N)}) \), \( i = 1, \cdots, N \), where the \( f_i(x) \) are the same as for the original infinite-state process.

Returning to the infinite-state process, it follows from the above that if \( q^{(N)} \leq a \) then \( \{ f_i(q^{(N)}) \} \), \( i \geq 1 \), would be a non-zero quasi-stationary measure \( \{ q_i \} \) satisfying \( \Sigma^{(N)}_i q_i > 1 \) (since, from (2.1), not all \( f_i(q^{(N)}) = 0 \), \( i > N \)).

3. Main results

Define \( \pi_1 = 1 \), \( \pi_n = \pi_{n-1}(\lambda_{n-1}/\mu_n) \) \( (n \geq 2) \). We shall make use of some or all of the conditions:

(i) \( \Sigma^{\infty}_{n=1}(1/\lambda_n \pi_n) = \infty \);
(ii) \( \Sigma^{\infty}_{n=1}(1/\lambda_n \pi_n) \Sigma^{n-1}_{i=1} \pi_i = \infty \);
(iii) \( \Sigma^{\infty}_{i=1}(1/\lambda_n \pi_n) \Sigma^{\infty}_{i=n+1} \pi_i = \infty \).

According to Karlin and McGregor (1957a,b), (i) ensures that there is a well-defined Markov process corresponding to the intensity matrix and also that absorption from each of the states \( i \), \( i \geq 1 \), into 0 is certain, and while (ii) ensures that the transition matrix of the process is stochastic at each time point. In relation to the polynomial system \( \{ Q_i(x) \} \), \( i \geq 0 \), generated by \( Q_0(x) = 0 \), \( Q_1(x) = 1 \), and for \( i \geq 1 \),

\[
\mu_i Q_{i-1}(x) - (\lambda_i + \mu_i) Q_i(x) + \lambda_i Q_{i+1}(x) + x Q_i(x) = 0,
\]

condition (iii) assures that there exists a unique regular measure \( \psi \) on \( 0 \leq x < \infty \) such that

\[
\pi_j \int_0^\infty Q_i(x)Q_j(x)d\psi(x) = \delta_{ij}, \quad i, j \geq 1.
\]

It is readily seen from (3.1) that for \( i \geq 1 \)

\[
Q_{i+1}(0) = 1 + \mu_i \sum_{k=1}^i \frac{1}{\lambda_k \pi_k}
\]

so that \( Q_{i+1}(0) > 1 \) for \( i \geq 1 \). If \( \gamma \) is the first point of increase of \( \psi \), it is well known from the theory of orthogonal polynomials that all roots of \( Q_{i+1}(x) = 0 \), \( i \geq 1 \), exceed \( \gamma \) (Szegö (1959), Section 3.3), whence it follows that \( Q_i(x) > 0 \), \( 0 \leq x \leq \gamma \), \( i \geq 1 \).
Definition 1. For an \( \alpha \geq 0 \), we call the column vector \( y = \{y_i\}, \ i \geq 1 \), an \( \alpha \)-invariant vector of the matrix \( A \) if \( y \geq 0, \ y \neq 0 \) and \( Ay = -\alpha y \). Similarly, we call the row vector \( x' = \{x_i\}, \ i \geq 1 \), an \( \alpha \)-invariant measure if \( x \geq 0, \ x \neq 0 \) and \( x'A = -\alpha x' \).

Lemma 1. Under assumption (iii), \( y = \{y_i\} \) is an \( \alpha \)-invariant vector of \( A \) if and only if \( x' = \{x_i,y_i\} \) is an \( \alpha \)-invariant measure of \( A \). The column vector \( \{Q_i(x)\}, \ i \geq 1 \), is an \( \alpha \)-invariant vector of \( A \), and the row vector \( \{\pi_i Q_i(x)\}, \ i \geq 1 \), is an \( \alpha \)-invariant measure, \( 0 \leq x \leq \gamma \).

Proof. Suppose \( y = \{y_i\} \) is an \( \alpha \)-invariant vector; thus, defining \( y_0 = 0 \), we have for \( i \geq 1 \)

\[
(3.3) \quad \mu_i y_{i-1} - (\lambda_i + \mu_i)y_i + \lambda_i y_{i+1} = -\alpha y_i.
\]

Note that \( \mu_i \pi_i = \lambda_i - \mu_i \pi_i, \ i \geq 2 \); and in fact for \( i \geq 1 \) if we define \( \lambda_0 \pi_0 = \mu_i \pi_i \). Multiplying (3.3) by \( \pi_i, \ i \geq 1 \), and putting \( x_i = \pi_i y_i, \ i \geq 1 \), we obtain for \( i \geq 1 \)

\[
(3.4) \quad \lambda_i - \mu_i \pi_i = \lambda_i x_i - \mu_i + \mu_i x_i = \lambda_i x_i = -\alpha x_i,
\]

if we define \( x_0 = 0 \), so \( \{x_i\}, \ i \geq 1 \), is an \( \alpha \)-invariant measure. Since the argument can be reversed, the first conclusion follows. Since we know from the preceding \( Q_0(x) = 0 \) and \( Q_i(x) > 0 \), from (3.1) we see that (3.3) is satisfied with \( y_i = Q_i(x), \ i \geq 0 \), whence the second conclusion follows.

Theorem 1. Under conditions (i) and (iii), if \( x' = \{x_i\} \) is any \( \alpha \)-invariant measure, \( \alpha \geq 0 \), then

\[
\sum_{i=1}^{\infty} x_i = \frac{\mu_i x_i}{\alpha}
\]

(the right-hand side to be understood as infinity if \( \alpha = 0 \)).

Proof. Note that \( x_1 > 0 \), otherwise, from (3.4), all \( x_i = 0 \), which contradicts the definition of \( \alpha \)-invariant measure. Putting \( y_i = x_i/(\pi_i x_i) \), \( i \geq 1 \), from Lemma 1 we see that \( y = \{y_i\} \) is an \( \alpha \)-invariant vector with \( y_1 = 1 \). Defining \( y_0 = 0 \), and following precisely the proof of the Lemma of Good (1968), p. 719, replacing his \( a \) by \( \alpha \), and his \( Q_n(a) \) by \( y_n \) (the strict inequality \( Q_n(a) > 0 \), \( n \geq 2 \), is not needed, so \( y_n \geq 0 \) suffices), we find that his conclusion translates to \( \sum_{i=1}^{\infty} \pi_i y_i = \mu_i/\alpha \), as required.

Corollary (Good (1968)).

\[
\sum_{i=1}^{\infty} \pi_i Q_i(\gamma) = \frac{\mu_1}{\gamma}.
\]

Theorem 2. Under conditions (i) and (iii), if \( \alpha > 0 \), any \( \alpha \)-invariant measure normed to sum to unity is a quasi-stationary measure. Conversely, any non-zero quasi-stationary measure \( q' = \{q_i\} \) is an \( \alpha \)-invariant measure with \( \alpha > 0 \) and sums to unity, and \( \alpha = q_i/\mu_i \).
Proof. The first part follows by dividing (3.4) by $\Sigma_{i=1}^{\infty} x_i$, and using Theorem 1. For the converse, from (2.1), $q^* = \{q_i\}$ is an $\alpha$-invariant measure with $\alpha = q_i \mu_i > 0$ ($q_i$ cannot be zero, otherwise from (2.1) $q_i = 0$). Thus by Theorem 1,

$$\sum_{i=1}^{\infty} q_i = \frac{\mu_i q_i}{q_i \mu_i} = 1. \tag{3.5}$$

Corollary. In the notation of Section 2, $q_i^{(N)} > a$. (Otherwise, we would have a non-zero quasi-stationary measure satisfying $\Sigma_{i=1}^{\infty} q_i > 1$, as indicated in Section 2.)

The 'converse' part of Theorem 2 strengthens (under the conditions stated) Lemma 8 of Cavender (1978). The fact that under conditions (i) and (iii) $q_i(x) = \pi_i x Q_i(x)/\mu_i$, $i \geq 1$, forms a unit-sum quasi-stationary measure for any $x$, $0 < x \leq \gamma$ when $\gamma > 0$, has been noted by van Doorn (1990), Theorem 3.2.

Theorem 3. Under the conditions (i) and (iii), as in Theorem 1, if $\gamma > 0$, then $\mu_i a = \gamma$ and $q_i^{(N)} \to q_i^*$, $i \geq 1$, where $q_i^* = \pi_i Q_i(\gamma)/\Sigma_{j=1}^{\infty} \pi_j Q_j(\gamma)$.

Proof. From Lemma 1, Theorem 1 and Theorem 2, there is a quasi-stationary measure

$$q_i = \frac{\pi_i Q_i(\gamma)}{\sum_{j=1}^{\infty} \pi_j Q_j(\gamma)} = \frac{1}{\sum_{j=1}^{\infty} \pi_j Q_j(\gamma)} = \frac{\gamma}{\mu_i}.$$  

Thus, by the definition of $a$, $\gamma \leq \mu_i a$.

Now, according to van Doorn (1985), Theorems 3.1 and 3.3, $x_i^{(N)} \sim \gamma$ as $N \to \infty$ where $x_i^{(N)} < x_i^{(N)} < \ldots < x_i^{(N)}$ are the roots of the equation $Q_{N+1}(x) = 0$. Further, according to van Doorn (1987),

$$f_N(x) = \det(-A - x I_N) = Q_{N+1}(x) \prod_{i=1}^{N} \lambda_i, \quad N \geq 1,$$

so that the $x_i^{(N)}$, $i = 1, \ldots, N$, are the roots of $f_N(x) = 0$; and according to Ledermann and Reuter (1954), Theorem 3, $\mu_i q_i^{(N)} < x_i^{(N)}$, since $\mu_i q_i^{(N)}$ is the smallest eigenvalue of $(-B_N)$. We know from the corollary to Theorem 2 that $q_i^{(N)} > a$, so that $0 < \gamma \leq \mu_i a < \mu_i q_i^{(N)} < x_i^{(N)}$. Thus, letting $N \to \infty$, the first conclusion follows, and also that

$$q_i^* = \lim_{N \to \infty} q_i^{(N)} = a = \frac{\gamma}{\mu_i} > 0.$$  

According to Cavender (1978), Proposition 7, $q_i^* = \lim_{N \to \infty} q_i^{(N)}$ exists for each $i \geq 1$, and $(q_i^*)$ is a quasi-stationary measure. According to the corollary to Theorem 1, $q_i^* = \gamma/\mu_i = \{\Sigma_{j=1}^{\infty} \pi_j Q_j(\gamma)\}^{-1}$, and the one-parameter character of quasi-stationary measures, the conclusion follows.

The above theorem tells us how, when $\gamma > 0$, the quasi-stationary measure $\{q_i^*\}$, $i \geq 1$, a probability distribution and hence a quasi-stationary distribution is approximable in the sense of convergence in distribution by the finite probability distributions $q_i^{(N)} =$
\(\{q_i^{(N)}\}, i = 1, \ldots, N\), corresponding to \(B_N\). Of more interest are the finite quasi-stationary distributions corresponding to the absorbing finite birth-and-death process where the intensity matrix corresponding to the non-absorbing states is \(A_N\), defined at the end of Section 1.

If \(\{q_i^{(N)}\}, i = 1, \ldots, N\), is the (unique, strictly positive) left eigenvector normed to sum to unity, of the matrix \((-A_N)\) corresponding to the smallest eigenvalue \(x_i^{(N)}\), then it is readily checked that \(X_i^{(N)} = \mu_i q_i^{(N)}\), and that (2.1) is satisfied for \(i = 1, \ldots, N\), with \(q_i = q_i^{(N)}\), \(i = 1, \ldots, N\), providing we put \(0 = q_0^{(N)} = q_{N+1}^{(N)} = \mu_{N+1}\). Since \(q_i^{(N)}\) is determined by \(q_i^{(N)}\), \(q_i^{(N)}\), and \(q_i^{(N)}\), and we know that

\[
\lim_{N \to \infty} q_i^{(N)} = \lim_{N \to \infty} \frac{x_i^{(N)}}{\mu_1} = \frac{\gamma}{\mu_1} = q_i^*,
\]

it follows that, when \(\gamma > 0\), \(q_i = \lim_{N \to \infty} q_i^{(N)}\) exists for each \(i\), and forms a non-zero quasi-stationary measure, with \(q_i^* = q_i = \gamma / \mu_1 = \{\Sigma_{i=1}^N \pi_i \xi_i(\gamma)\}^{-1}\), and hence that \(q_i = q_i^*\). This is the analogue of the conclusion (Theorem 2) of Keilson and Ramaswamy (1984) for such convergence in the non-absorbing situation where \(\mu_i = 0\) under conditions (i) and condition \(\Sigma_{i=1}^N \pi_i < \infty\) (these two conditions together clearly imply (iii)). They pursue this convergence through stochastic bounds for the finite quasi-stationary distribution \(\{q_i^{(N)}\}, i = 1, \ldots, N\), each of which is shown to converge to the quasi-stationary distribution for the process \(\{X(t)\}\) on non-negative integers. The upper stochastic bound in this case \((\mu_i = 0)\) is in fact \(\{q_i^{(N)}\}\). We take up the idea of stochastic bounding for the infinite quasi-stationary distribution \(\{q_i^*\}, i \geq 1\), in our setting \(\mu_i > 0\) and \(\gamma > 0\), and obtain a result (Theorem 4 below) which relates it to the ergodic stationary distribution in the setting \(\mu_i = 0\), under the Keilson–Ramaswamy assumptions.

**Definition 2.** Let \(a = \{a_n\}\) and \(b = \{b_n\}\), \(n \geq 1\), be positive vectors. Then \(a\) is said to be strictly smaller than \(b\) in the likelihood ratio (write \(a <_l b\)) if \(a_n b_{n+1} > a_{n+1} b_n\) for all \(n \geq 1\).

If the strict inequality ‘\(\succ\)’ is replaced by ‘\(\succeq\)’ then we say \(a\) is smaller than \(b\) in the likelihood ratio, and write \(a \preceq_l b\). If \(a \preceq_l b\), then \(a_n b_m \geq a_n a_m\) for \(m \geq n\) so \(\Sigma_{j=1}^n b_j a_m\). whence \(\Sigma_{j=1}^n a_j(b_{n+1} + \cdots + b_m) \geq \Sigma_{j=1}^n b_j(a_{n+1} + \cdots + a_m)\), \(m > n\), which in turn implies \(\Sigma_{j=1}^n a_j \Sigma_{i=1}^m b_j \geq \Sigma_{i=1}^m a_i \Sigma_{j=1}^m b_j\), \(m > n\), which, by letting \(m \to \infty\) in the case where \(a\) and \(b\) are probability distributions, implies \(a\) is stochastically smaller than \(b\) (write \(a < b\)) in the usual sense (see Keilson and Ramaswamy (1984), Appendix for example). Cavender (1978), Corollary 3, in effect shows that if \(q' = \{q_i\}\) and \(r' = \{r_i\}\) are non-zero quasi-stationary measures with \(q_i > r_i\), then \(q <_l r\); under the conditions (i) and (iii) both \(q\) and \(r\) are probability distributions (our Theorem 2). This can in theory be used to provide a stochastic lower bound for our \(\{q_i^*\}\), since any non-zero quasi-stationary measure not equal to \(\{q_i^*\}\) has \(0 < r_i < q_i^* = \gamma / \mu_i = a\). A more practically useful and theoretically interesting result is the following.

**Theorem 4.** Under conditions (i) and (iii) and assuming \(\gamma > 0\), it is true that
Some results for quasi-stationary distributions of birth–death processes

\[ \sum_{i=1}^{\infty} \pi_i < \infty, \quad Q_{i+1}(\gamma) > Q_i(\gamma) \quad \text{for} \ i \geq 1, \ \text{and} \ \pi^* <_\text{a} q^*, \]

where \( \pi^* = \pi_i / \Sigma_{k=1}^{\infty} \pi_k \).

**Proof.** From the proof of Theorem 3,

\[ Q_i(\gamma) = \frac{f_i(\gamma)}{\prod_{k=1}^{i-1} \lambda_k} = \frac{\det(-A_{i-1} - \gamma I_{i-1})}{\prod_{k=1}^{i-1} \lambda_k}, \quad i \geq 2. \tag{3.6} \]

Expanding \( \det(-A_{i-1} - \gamma I_{i+1}) \) in the last row yields

\[ \det(-A_{i-1} - \gamma I_{i+1}) = (\lambda_i + \mu_i - \gamma) \det(-A_i - \gamma I_i) - \lambda_i \mu_i \det(-A_{i-1} - \gamma I_{i-1}). \]

Also expanding \( \det(-B_{i-1} - \gamma I_{i+1}) \) in the last row yields

\[ \det(-B_{i-1} - \gamma I_{i+1}) = (\mu_i + \gamma) \det(-A_i - \gamma I_i) - \lambda_i \mu_i \det(-A_{i-1} - \gamma I_{i-1}), \]

whence it follows that

\[ \det(-A_{i-1} - \gamma I_{i+1}) = \lambda_i + \mu_i \det(-A_i - \gamma I_i) + \det(-B_{i-1} - \gamma I_{i+1}). \]

These relations remain valid for \( i = 1 \) if we define \( \det(-A_0 - \gamma I_0) = 1 \). But the smallest eigenvalue of \( (-B_{i-1}) \) is \( \mu_i q^{i+1} \), so that \( \det(-B_{i-1} - \gamma I_{i+1}) > 0 \). Hence

\[ \det(-A_{i+1} - \gamma I_{i+1}) > \lambda_i + \mu_i \det(-A_i - \gamma I_i), \quad i \geq 1. \tag{3.7} \]

Now from (3.6) and (3.7)

\[ Q_{i+2}(\gamma) \prod_{k=1}^{i+1} \lambda_k > \lambda_{i+1} Q_{i+1}(\gamma) \prod_{k=1}^{i} \lambda_k, \]

i.e.

\[ Q_{i+2}(\gamma) > Q_{i+1}(\gamma), \quad i \geq 1. \]

Now \( Q_2(\gamma) = (\mu_1 - \gamma) / \lambda_1 + 1 = Q_1(\gamma) \), since \( \gamma = \mu_i q^*_i < \mu_1 \) by Theorem 3 and Corollary to Theorem 1. Hence it follows that for \( i \geq 1 \)

\[ \frac{(Q_i(\gamma) \pi_i) \pi_{i+1}}{(Q_{i+1}(\gamma) \pi_{i+1}) \pi_i} < 1, \tag{3.8} \]

whence \( \{ \pi_i \} <_\text{a} \{ \pi_i Q_i(\gamma) \} \) with \( \pi_i Q_i(\gamma) = \pi_i \), so \( \Sigma_{k=1}^{\infty} \pi_i < \Sigma_{k=1}^{\infty} \pi_i Q_i(\gamma) < \infty \) by the corollary to Theorem 1, whence the conclusion follows by dividing numerator and denominator of (3.8) by \( (\Sigma_{k=1}^{\infty} \pi_i Q_i(\gamma)) (\Sigma_{k=1}^{\infty} \pi_i) \).

As is well known, \( \pi_i / \Sigma_{k=1}^{\infty} \pi_k \) is the ergodic distribution of the process with \( \mu_1 = 0 \) (ignoring state 0). For the finite absorbing process governed by the section \( A_N \), the divergence in condition (i) and hence the fact that \( Q_i(0) \to \infty \) as \( i \to \infty \) from (3.2) is equivalent to \( E(T|N) \to \infty \) as \( N \to \infty \), where \( T|N \) is the time to absorption from state 1, under the conditions of Theorem 4. A parallel to the strict monotonicity of (3.2) is the strict monotonicity of the \( \gamma \)-invariant vector \( Q_i(\gamma) \) in Theorem 4.
Of more interest is the relation of \( \{q_i^*\} \) to the ergodic distribution of the process with \( \mu_i > 0 \) and \( \lambda_i > 0 \), which in Section 1 we called \( \{N(t)\} \), \( t \geq 0 \), on \( \{0, 1, 2, \ldots\} \). As is well known, if we define \( \pi_0 = 1 \), \( \pi_i = \pi_{i-1} \lambda_{i-1} / \mu_i \), \( i \geq 1 \), and if \( \sum_{i=0}^{\infty} \pi_i < \infty \) then the process is ergodic and the limiting stationary distribution is \( \pi_i / \sum_{i=0}^{\infty} \pi_i \), \( i \geq 0 \). Now,

\[
\pi_i = \frac{\lambda_0}{\mu_1} \cdots \frac{\lambda_{i-1}}{\mu_i} \pi_i, \quad i \geq 1.
\]

Hence, under the conditions in Theorem 4, \( \sum_{i=0}^{\infty} \pi_i < \infty \) and the ergodic distribution of this process, restricted to (conditioned on) the states \( \{1, 2, \ldots\} \) provides a stochastic lower bound to the quasi-stationary distribution \( \{q_i^*\}, i \geq 1 \).

Finally, we state the limit theorem foreshadowed in Section 1 which provides the physical best for the term ‘quasi-stationary distribution’.

**Theorem 5** (Good (1968); Kesten (1970)). Under conditions (i), (ii) and (iii) and \( \gamma > 0 \), then for \( i, j \geq 1 \)

\[
\lim_{t \to \infty} \Pr[X(t) = j | X(t) \in \{1, 2, \ldots\}, X(0) = i] = \lim_{t \to \infty} \Pr[N(t) = j | T(i) > t] = q_i^* = \frac{\pi_j Q_i(\gamma)}{\sum_{k=1}^{\infty} \pi_k Q_k(\gamma)}.
\]

Further investigation along these lines of asymptotic conditional limits occurs in van Doorn (1990).

**Acknowledgement**

One of the authors (M.K.) wishes to thank E. van Doorn for helpful comments on an earlier version of this paper.

**References**

Some results for quasi-stationary distributions of birth–death processes


