REPRESENTATIONS FOR THE DECAY PARAMETER OF A BIRTH–DEATH PROCESS BASED ON THE COURANT–FISCHER THEOREM

ERIK A. VAN DOORN,* *University of Twente*

Abstract

We study the *decay parameter* (the rate of convergence of the transition probabilities) of a birth–death process on $\{0, 1, \ldots\}$, which we allow to evanesce by escape, via state 0, to an absorbing state -1. Our main results are representations for the decay parameter under four different scenarios, derived from a unified perspective involving the orthogonal polynomials appearing in Karlin and McGregor's representation for the transition probabilities of a birth–death process, and the Courant–Fischer theorem on eigenvalues of a symmetric matrix. We also show how the representations readily yield some upper and lower bounds that have appeared in the literature.

Keywords: Birth–death process; exponential decay; rate of convergence; orthogonal polynomials

2010 Mathematics Subject Classification: Primary 60J80
Secondary 42C05

1. Introduction

A *birth–death process* is a continuous-time Markov chain $\mathcal{X} := \{X(t), \ t \geq 0\}$ taking values in $S := \{0, 1, 2, \ldots\}$ with q-matrix $Q := (q_{ij}, \ i, j \in S)$ given by

$$q_{i,i+1} = \lambda_i, \qquad q_{i+1,i} = \mu_{i+1}, \qquad q_{ii} = -(\lambda_i + \mu_i),$$
$$q_{ij} = 0, \qquad |i - j| > 1,$$

where $\lambda_i > 0$ for $i \geq 0$, $\mu_i > 0$ for $i \geq 1$, and $\mu_0 \geq 0$. Positivity of μ_0 entails that the process may evanesce by escaping from S, via state 0, to an absorbing state -1. Throughout this paper we will assume that the *birth rates* λ_i and *death rates* μ_i uniquely determine the process $\mathcal{X}$. Karlin and McGregor [9] have shown that this is equivalent to assuming that

$$\sum_{n=0}^{\infty} \left(\pi_n + \frac{1}{\lambda_n \pi_n} \right) = \infty, \tag{1}$$

where π_n are constants given by

$$\pi_0 := 1 \quad \text{and} \quad \pi_n := \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \mu_2 \ldots \mu_n}, \qquad n > 0.$$

We note that condition (1) does not exclude the possibility of *explosion*, escape from S, via all states larger than the initial state, to an absorbing state ∞.

It is well known that the transition probabilities

$$p_{ij}(t) := \mathbb{P}\{X(t) = j \mid X(0) = i\}, \quad t \geq 0, \ i, j \in S,$$

have limits

$$p_j := \lim_{t \to \infty} p_{ij}(t) = \begin{cases} \pi_j \left(\sum_{n=0}^{\infty} \pi_n \right)^{-1} & \text{if } \mu_0 = 0 \text{ and } \sum_{n=0}^{\infty} \pi_n < \infty, \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

which are independent of the initial state i. If $\mu_0 > 0$ and the initial state is i then a_i, the probability of eventual absorption at -1, is given by

$$a_i = \frac{\mu_0 \sum_{n=i}^{\infty} (\lambda_n \pi_n)^{-1}}{1 + \mu_0 \sum_{n=0}^{\infty} (\lambda_n \pi_n)^{-1}}, \quad i \in S, \tag{3}$$

where the right-hand side of (3) should be interpreted as 1 if $\sum_n (\lambda_n \pi_n)^{-1}$ diverges (see Karlin and McGregor [10, Theorem 10]).

The exponential rate of convergence of $p_{ij}(t)$ to its limit p_j will be denoted by α_{ij}, that is,

$$\alpha_{ij} := -\lim_{t \to \infty} \frac{1}{t} \log |p_{ij}(t) - p_j| \geq 0, \quad i, j \in S.$$

From Callaert [1] we know that these limits exist, and that

$$\alpha := \alpha_{00} \leq \alpha_{ij}, \quad i, j \in S,$$

with equality whenever $\mu_0 > 0$, and inequality prevailing for at most one value of i or j when $\mu_0 = 0$. We will refer to α as the *decay parameter* of $\mathcal{X}$.

In this paper our interest focuses on representations and bounds for α. Our main goal is to provide new proofs for a number of results that have appeared in the literature, notably in the work of Chen [2]–[5], but see also Sirl *et al.* [13]. Our approach involves the orthogonal polynomials appearing in Karlin and McGregor's spectral representation for the transition probabilities of a birth–death process, and the Courant–Fischer theorem on eigenvalues of a symmetric matrix.

2. Results

We discern four different scenarios depending on whether $\mu_0 = 0$ or $\mu_0 > 0$, and the series $\sum_n \pi_n$ (if $\mu_0 = 0$) or $\sum_n (\lambda_n \pi_n)^{-1}$ (if $\mu_0 > 0$) converges or diverges. We will prove the representations and bounds for α that are given in Theorems 1 to 4 below. These results readily yield a number of known bounds for α, which are displayed in Corollaries 1 to 4.

In what follows $\mathbf{0}$ denotes a sequence consisting entirely of zeros and U is the set of sequences of real numbers $\boldsymbol{u} := (u_0, u_1, \dots) \neq \mathbf{0}$ that are *eventually vanishing*, that is, $U = \cup_{n \geq 0} U_n$, where

$$U_n := \{\boldsymbol{u} = (u_0, u_1, \dots) \neq \mathbf{0} : u_i = 0 \text{ for } i > n\}. \tag{4}$$

Theorem 1. *Let $\mu_0 > 0$ and $\sum_n (\lambda_n \pi_n)^{-1} = \infty$. Then*

$$\alpha = \inf_{\boldsymbol{u} \in U} \left\{ \frac{\sum_{i=0}^{\infty} \mu_i \pi_i u_i^2}{\sum_{i=0}^{\infty} \pi_i (\sum_{j=0}^{i} u_j)^2} \right\}. \tag{5}$$

Corollary 1. ([5], [13].) *Let $\mu_0 > 0$ and $\sum_n (\lambda_n \pi_n)^{-1} = \infty$. If*

$$R_0 := \sup_{n \geq 0} \left\{ \sum_{i=0}^{n} \frac{1}{\mu_i \pi_i} \sum_{i=n}^{\infty} \pi_i \right\} = \infty$$

then $\alpha = 0$, while

$$R_0 < \infty \implies \frac{1}{4R_0} < \alpha < \frac{1}{R_0}.$$

Theorem 2. *Let $\mu_0 > 0$ and $\sum_n (\lambda_n \pi_n)^{-1} < \infty$. Then*

$$\alpha = \inf_{\boldsymbol{u} \in U} \left\{ \frac{\sum_{k=0}^{\infty} (\mu_k \pi_k)^{-1} \sum_{i=0}^{\infty} u_i^2 / \pi_i}{\sum_{k=0}^{\infty} (\mu_k \pi_k)^{-1} \sum_{i=1}^{\infty} (\mu_i \pi_i)^{-1} \left(\sum_{j=0}^{i-1} u_j \right)^2 - \left(\sum_{i=1}^{\infty} (\mu_i \pi_i)^{-1} \sum_{j=0}^{i-1} u_j \right)^2} \right\},$$

whence,

$$\tilde{\alpha}_a \leq \alpha \leq \tilde{\alpha}_a \left(1 + \mu_0 \sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n} \right),$$

where

$$\tilde{\alpha}_a := \inf_{\boldsymbol{u} \in U} \left\{ \frac{\sum_{i=0}^{\infty} u_i^2 / \pi_i}{\sum_{i=0}^{\infty} (\lambda_i \pi_i)^{-1} \left(\sum_{j=0}^{i} u_j \right)^2} \right\}.$$

Corollary 2. ([5].) *Let $\mu_0 > 0$ and $\sum_n (\lambda_n \pi_n)^{-1} < \infty$. If*

$$S := \sup_{n \geq 0} \left\{ \sum_{i=0}^{n} \pi_i \sum_{i=n}^{\infty} \frac{1}{\lambda_i \pi_i} \right\} = \infty \tag{6}$$

then $\alpha = 0$, while

$$S < \infty \implies \frac{1}{4S} < \alpha < \frac{1}{S} \left(1 + \mu_0 \sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n} \right).$$

Theorem 3. *Let $\mu_0 = 0$ and $\sum_n \pi_n = \infty$. Then*

$$\alpha = \inf_{\boldsymbol{u} \in U} \left\{ \frac{\sum_{i=0}^{\infty} u_i^2 / \pi_i}{\sum_{i=0}^{\infty} (\lambda_i \pi_i)^{-1} \left(\sum_{j=0}^{i} u_j \right)^2} \right\}.$$

Corollary 3. ([5].) *Let $\mu_0 = 0$ and $\sum_n \pi_n = \infty$. If (6) holds then $\alpha = 0$, while*

$$S < \infty \implies \frac{1}{4S} < \alpha < \frac{1}{S}.$$

Theorem 4. *Let $\mu_0 = 0$ and $\sum_n \pi_n < \infty$. Then*

$$\alpha = \inf_{\boldsymbol{u} \in U} \left\{ \frac{\sum_{k=0}^{\infty} \pi_k \sum_{i=0}^{\infty} \lambda_i \pi_i u_i^2}{\sum_{k=0}^{\infty} \pi_k \sum_{i=0}^{\infty} \pi_{i+1} \left(\sum_{j=0}^{i} u_j \right)^2 - \left(\sum_{i=0}^{\infty} \pi_{i+1} \sum_{j=0}^{i} u_j \right)^2} \right\}, \tag{7}$$

whence,

$$\tilde{\alpha}_r \leq \alpha \leq \tilde{\alpha}_r \sum_{n=0}^{\infty} \pi_n,$$

where

$$\tilde{\alpha}_r := \inf_{\boldsymbol{u} \in U} \left\{ \frac{\sum_{i=0}^{\infty} \lambda_i \pi_i u_i^2}{\sum_{i=0}^{\infty} \pi_{i+1} \left(\sum_{j=0}^{i} u_j \right)^2} \right\}.$$

Corollary 4. ([3], [4].) *Let $\mu_0 = 0$ and $\sum_n \pi_n < \infty$. If*

$$R_1 := \sup_{n \geq 1} \left\{ \sum_{i=1}^{n} \frac{1}{\mu_i \pi_i} \sum_{i=n}^{\infty} \pi_i \right\} = \infty$$

then $\alpha = 0$, while

$$R_1 < \infty \implies \frac{1}{4R_1} < \alpha < \frac{1}{R_1} \sum_{n=0}^{\infty} \pi_n.$$

Note that the corollaries provide simple criteria for α to be positive. This is particularly relevant in the setting of a birth–death process for which absorption at -1 is certain (that is, in view of (3), the setting of Theorem 1), since positivity of the decay parameter is necessary and sufficient for the existence of a *quasi-stationary distribution* (see [16, Section 5.1] for detailed information).

Before proving the theorems and corollaries in Section 4, we present a number of preliminary results in Section 3. In Section 5 we provide some additional information on related literature.

3. Preliminaries

3.1. Birth–death polynomials

The birth and death rates of the process $\mathcal{X}$ determine a sequence of polynomials $\{Q_n\}$ through the recurrence relation

$$\begin{aligned}
\lambda_n Q_{n+1}(x) &= (\lambda_n + \mu_n - x) Q_n(x) - \mu_n Q_{n-1}(x), \qquad n > 0, \\
\lambda_0 Q_1(x) &= \lambda_0 + \mu_0 - x, \qquad Q_0(x) = 1.
\end{aligned} \tag{8}$$

It is sometimes convenient to renormalize the polynomials Q_n by letting

$$P_0(x) := 1 \quad \text{and} \quad P_n(x) := (-1)^n \lambda_0 \lambda_1 \ldots \lambda_{n-1} Q_n(x), \quad n > 0,$$

so that the recurrence relation (8) translates into

$$\begin{aligned}
P_{n+1}(x) &= (x - \lambda_n - \mu_n) P_n(x) - \lambda_{n-1} \mu_n P_{n-1}(x), \qquad n > 0, \\
P_1(x) &= x - \lambda_0 - \mu_0, \qquad P_0(x) = 1.
\end{aligned} \tag{9}$$

It will also be convenient to set $\lambda_{-1} := 0$.

The sequence $\{Q_n\}$ plays an important role in the analysis of the birth–death process $\mathcal{X}$ since, by a famous result of Karlin and McGregor [9], the transition probabilities of $\mathcal{X}$ can be represented as

$$p_{ij}(t) = \pi_j \int_0^{\infty} e^{-xt} Q_i(x) Q_j(x) \psi(\mathrm{d}x), \qquad t \geq 0, \; i, j \in S, \tag{10}$$

where ψ is a probability measure on the nonnegative real axis, which is uniquely determined by the birth and death rates if (1) is satisfied. Note that as a result of (10) we have $p_j = \pi_j \psi(\{0\})$, so (2) implies that

$$\psi(\{0\}) = \begin{cases} \left(\sum_{n=0}^{\infty} \pi_n \right)^{-1} & \text{if } \mu_0 = 0 \text{ and } \sum_{n=0}^{\infty} \pi_n < \infty, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

The measure ψ has a finite moment of order -1 if $\mu_0 = 0$ and $\sum_n (\lambda_n \pi_n)^{-1} < \infty$, or if $\mu_0 > 0$. Indeed, by [9, Equation (2.4) and Lemma 6] we have

$$\int_0^\infty \frac{\psi(\mathrm{d}x)}{x} = \frac{\sum_{n=0}^\infty (\lambda_n \pi_n)^{-1}}{1 + \mu_0 \sum_{n=0}^\infty (\lambda_n \pi_n)^{-1}}, \tag{12}$$

which should be interpreted, if $\sum_n (\lambda_n \pi_n)^{-1}$ diverges, as ∞ for $\mu_0 = 0$ and as μ_0^{-1} for $\mu_0 > 0$.

Of particular interest to us will be the quantities ξ_i, recurrently defined by

$$\xi_1 := \inf \operatorname{supp}(\psi) \tag{13}$$

and

$$\xi_{i+1} := \inf\{\operatorname{supp}(\psi) \cap (\xi_i, \infty)\}, \qquad i \geq 1, \tag{14}$$

where $\operatorname{supp}(\psi)$ denotes the support of the measure ψ (also referred to as the *spectrum* of the process). Namely, representation (10) implies (see [14, Theorem 3.1 and Lemma 3.2]) that the decay parameter α of $\mathcal{X}$ can be expressed as

$$\alpha = \begin{cases} \xi_2 & \text{if } \xi_2 > \xi_1 = 0, \\ \xi_1 & \text{otherwise.} \end{cases} \tag{15}$$

If $\xi_2 > \xi_1 = 0$ we must have $p_j = \pi_j \psi(\{0\}) > 0$, so (11) tells us

$$\mu_0 > 0 \quad \text{or} \quad \sum_{n=0}^\infty \pi_n = \infty \quad \implies \quad \alpha = \xi_1. \tag{16}$$

We further define

$$\sigma := \lim_{i \to \infty} \xi_i, \tag{17}$$

the first accumulation point of $\operatorname{supp}(\psi)$ if it exists, and ∞ otherwise. It is clear from the definition of ξ_i that, for all $i \geq 1$,

$$\xi_{i+1} \geq \xi_i \geq 0,$$

and

$$\xi_i = \xi_{i+1} \iff \xi_i = \sigma.$$

Note that we must have $\sigma = 0$ if $\xi_1 = 0$ but $\psi(\{0\}) = 0$.

Since $p_{ij}(0) = \delta_{ij}$, where δ_{ij} is Kronecker's delta, (10) implies that

$$\pi_j \int_0^\infty Q_i(x) Q_j(x) \psi(\mathrm{d}x) = \delta_{ij}, \qquad i, j \in S, \tag{18}$$

that is, the polynomials $\{Q_n(x)\}$ are orthogonal with respect to the measure ψ. In the terminology of the theory of moments the *Stieltjes moment problem* associated with $\{Q_n\}$ is said to be *determined* if there is a unique probability measure ψ on the nonnegative real axis satisfying (18), and *indeterminate* otherwise. In the latter case there is, by [6, Theorem 5], a unique orthogonalizing probability measure for which the infimum of its support is maximal. We will refer to this measure (which happens to be discrete) as the *natural* measure for $\{Q_n\}$. Our assumption (1) does not necessarily imply that the Stieltjes moment problem associated with $\{Q_n\}$ is determined, but if it is indeterminate then (10) will be satisfied only by the natural measure. For details and related results we refer the reader to [9] (see also [8] and [15]).

In what follows the measure ψ, if not uniquely determined by (18), should be interpreted as the natural measure. With this convention the quantities ξ_n and σ of (13), (14), and (17) may be defined alternatively in terms of the (simple and positive) zeros of the polynomials $Q_n(x)$ (see [7, Section II.4]). Namely, with $x_{n1} < x_{n2} < \cdots < x_{nn}$ denoting the n zeros of $Q_n(x)$, we have the classical separation result

$$0 < x_{n+1,i} < x_{ni} < x_{n+1,i+1}, \qquad i = 1, 2, \ldots, n, \ n \geq 1, \tag{19}$$

so that the limits as $n \to \infty$ of x_{ni} exist, and

$$\lim_{n \to \infty} x_{ni} = \xi_i, \qquad i = 1, 2, \ldots, n. \tag{20}$$

3.2. Dual birth–death processes

Our point of departure in this subsection is a birth–death process $\mathcal{X}$ with birth rates λ_i and death rates μ_i such that $\mu_0 > 0$. Following Karlin and McGregor [9], [10], we define the process $\mathcal{X}^d$ to be a birth–death process on S with birth rates λ_i^d and death rates μ_i^d given by $\mu_0^d = 0$ and

$$\lambda_i^d := \mu_i, \quad \mu_{i+1}^d := \lambda_i, \qquad i \geq 0. \tag{21}$$

Accordingly, we define

$$\pi_0^d := 1 \quad \text{and} \quad \pi_n^d := \frac{\lambda_0^d \lambda_1^d \ldots \lambda_{n-1}^d}{\mu_1^d \mu_2^d \ldots \mu_n^d} = \frac{\mu_0 \mu_1 \ldots \mu_{n-1}}{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}, \qquad n \geq 1,$$

and note that

$$\pi_{n+1}^d = \mu_0 (\lambda_n \pi_n)^{-1} \quad \text{and} \quad (\lambda_n^d \pi_n^d)^{-1} = \mu_0^{-1} \pi_n, \qquad n \geq 0. \tag{22}$$

So our assumption (1) is equivalent to

$$\sum_{n=0}^{\infty} \left(\pi_n^d + \frac{1}{\lambda_n^d \pi_n^d} \right) = \infty,$$

and, hence, the process $\mathcal{X}^d$ is uniquely determined by its rates. So, within the setting of birth–death processes satisfying (1), (21) establishes a one-to-one correspondence between processes with $\mu_0 = 0$ and those with $\mu_0 > 0$. Therefore, $\mathcal{X}$ and $\mathcal{X}^d$ will be called each other's *dual*.

The transition probabilities of $\mathcal{X}^d$ satisfy a representation formula analogous to (10), involving birth–death polynomials Q_n^d (with corresponding monic polynomials P_n^d) and a unique probability measure ψ^d on the nonnegative real axis with respect to which the polynomials Q_n^d are orthogonal. By [9, Lemma 3] (see also [14]), we have

$$\mu_0 \psi([0, x]) = x \psi^d([0, x]), \qquad x \geq 0.$$

With ξ_i^d and σ^d denoting the quantities defined by (13), (14), and (17) if we replace ψ by ψ^d, we thus have $\sigma^d = \sigma$ and

$$\xi_i = \begin{cases} \xi_{i+1}^d & \text{if } \xi_1^d = 0, \\ \xi_i^d & \text{if } \xi_1^d > 0, \end{cases} \qquad i \geq 1. \tag{23}$$

The relations between the polynomials corresponding to $\mathcal{X}$ and $\mathcal{X}^d$ are most conveniently expressed in terms of the monic polynomials P_n and P_n^d, namely

$$P_{n+1}^d(x) = P_{n+1}(x) + \lambda_n P_n(x), \qquad n \geq 0, \tag{24}$$

and

$$x P_n(x) = P_{n+1}^d(x) + \lambda_n^d P_n^d(x), \qquad n \geq 0. \tag{25}$$

These relations, which are easy to verify, reveal the fact that the zeros of the polynomials corresponding to a birth–death process—which determine the decay parameter of the process through (15) and (20)– may be studied via the polynomials of the dual process. This will prove to be a crucial observation, since the technique that is used in the next subsection to obtain representations for the zeros, although applicable to $P_n(x)$ and $P_n^d(x)$, appears more rewarding when applied to $P_{n+1}(x) + \lambda_n P_n(x)$ and $P_{n+1}^d(x) + \lambda_n^d P_n^d(x)$. We will obtain representations for the smallest zero of $P_{n+1}(x) + \lambda_n P_n(x)$, and, hence, for the smallest zero of $P_{n+1}^d(x)$, and for the second smallest zero of $P_{n+1}^d(x) + \lambda_n^d P_n^d(x)$ (the smallest being 0), and, hence, for the smallest zero of $P_n(x)$.

The superindex d, used in this subsection to identify quantities related to the dual process in one direction only, will from now on be used in two directions, so that, for example, $(\mathcal{X}^d)^d = \mathcal{X}$.

3.3. Representations for the zeros of $P_{n+1}(x) + \lambda_n P_n(x)$

In this subsection we allow $\mu_0 \geq 0$ again, and define $\tilde{P}_0(x) = 1$ and

$$\tilde{P}_{n+1}(x) := P_{n+1}(x) + \lambda_n P_n(x), \qquad n \geq 0.$$

The zeros of $\tilde{P}_n(x)$ will be denoted by $\tilde{x}_{ni}$, $i = 1, 2, \ldots, n$. In view of (19), (24), and (25) we have $\tilde{x}_{n,1} = 0$ for all n if $\mu_0 = 0$ and, for $\mu_0 \geq 0$,

$$0 \leq \tilde{x}_{n+1,i} \leq \tilde{x}_{ni} < \tilde{x}_{n+1,i+1}, \qquad i = 1, 2, \ldots, n, \ n \geq 1, \tag{26}$$

which implies the existence of the limits

$$\tilde{\xi}_i := \lim_{n \to \infty} \tilde{x}_{ni}, \qquad i = 1, 2, \ldots, n. \tag{27}$$

To obtain suitable representations for $\tilde{x}_{n1}$ and $\tilde{\xi}_1$, and, if $\mu_0 = 0$, for $\tilde{x}_{n2}$ and $\tilde{\xi}_2$, we will generalise the approach leading to [17, Theorem 3].

First note that, by the recurrence relation (9),

$$\tilde{P}_{n+1}(x) = (x - \mu_n) P_n(x) - \lambda_{n-1} \mu_n P_{n-1}(x), \qquad n > 0,$$

so that the polynomials $P_0(x), P_1(x), \ldots, P_n(x), \tilde{P}_{n+1}(x)$ satisfy a three-terms recurrence relation similar to (9) except that λ_n is replaced by 0. Next, let the $(n+1) \times (n+1)$ symmetric tridiagonal matrix M_n be defined by $M_0 := (\mu_0)$ and, for $n > 0$,

$$M_n := \begin{pmatrix} \lambda_0 + \mu_0 & -\sqrt{\lambda_0 \mu_1} & 0 & \cdots & 0 & 0 \\ -\sqrt{\lambda_0 \mu_1} & \lambda_1 + \mu_1 & -\sqrt{\lambda_1 \mu_2} & \cdots & 0 & 0 \\ 0 & -\sqrt{\lambda_1 \mu_2} & \lambda_2 + \mu_2 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} + \mu_{n-1} & -\sqrt{\lambda_{n-1} \mu_n} \\ 0 & 0 & 0 & \cdots & -\sqrt{\lambda_{n-1} \mu_n} & \mu_n \end{pmatrix}.$$

Denoting the $n \times n$ identity matrix by I_n, it is now readily verified by expanding $\det(x I_{n+1} - M_n)$ by its last row that

$$\det(x I_{n+1} - M_n) = \tilde{P}_{n+1}(x), \qquad n \geq 0,$$

so that the zeros $\tilde{x}_{n+1,1}, \ldots, \tilde{x}_{n+1,n+1}$ of $\tilde{P}_{n+1}(x)$ are precisely the (real and simple) eigenvalues of M_n. From the *Courant–Fischer theorem* for symmetric matrices (see, for example, Meyer [11, page 550]) it follows that

$$\tilde{x}_{n+1,1} = \min_{y \neq 0} \frac{y M_n y^T}{y y^T}$$

and

$$\tilde{x}_{n+1,2} = \max_{\dim \mathcal{V}=n} \min_{\substack{y \in \mathcal{V} \\ y \neq 0}} \frac{y M_n y^T}{y y^T}, \tag{28}$$

where $y := (y_0, y_1, \ldots, y_n)$. Writing

$$y_i = s_i \sqrt{\pi_i} \quad \text{and} \quad s_i = \sum_{j=0}^{i} u_j, \qquad i \geq 0, \tag{29}$$

we obtain

$$
\begin{aligned}
y M_n y^T &= \sum_{i=0}^{n} \left(y_i^2 (\lambda_i(1 - \delta_{in}) + \mu_i) - 2 y_{i-1} y_i \sqrt{\lambda_{i-1}\mu_i} \right) \\
&= \sum_{i=0}^{n-1} \lambda_i \pi_i s_i^2 + \sum_{i=0}^{n} \mu_i \pi_i s_i^2 - 2 \sum_{i=1}^{n} s_{i-1} s_i \sqrt{\lambda_{i-1}\pi_{i-1}\mu_i\pi_i} \\
&= \sum_{i=1}^{n} \mu_i \pi_i (s_{i-1}^2 + s_i^2 - 2 s_{i-1} s_i) + \mu_0 s_0^2 \\
&= \sum_{i=0}^{n} \mu_i \pi_i u_i^2,
\end{aligned}
$$

where we have exploited the fact that $\lambda_{i-1}\pi_{i-1} = \mu_i \pi_i$. It follows that

$$\tilde{x}_{n+1,1} = \min_{u \neq 0} \left\{ \frac{\sum_{i=0}^{n} \mu_i \pi_i u_i^2}{\sum_{i=0}^{n} \pi_i (\sum_{j=0}^{i} u_j)^2} \right\}, \tag{30}$$

where $u = (u_0, u_1, \ldots, u_n)$ is a sequence of real numbers.

If $\mu_0 = 0$ the expression between the braces is minimised by choosing $u = (1, 0, \ldots, 0)$, yielding $\tilde{x}_{n+1,1} = 0$, which is in complete agreement with (25). In this case, we can use (28) to find a suitable representation for $\tilde{x}_{n+1,2}$. Note that $u = (1, 0, \ldots, 0)$ corresponds to $y = a := (\sqrt{\pi_0}, \sqrt{\pi_1}, \ldots, \sqrt{\pi_n})$, which is readily seen to be a left eigenvector of M_n corresponding to the eigenvalue 0. Hence, choosing $\mathcal{V}$ to be the space orthogonal to a we have

$$\tilde{x}_{n+1,2} \leq \min_{\substack{ya^T=0 \\ y \neq 0}} \frac{y M_n y^T}{y y^T}.$$

But, in fact, equality holds, since we may choose $\mathbf{y}$ to be a left eigenvector of M_n corresponding to the eigenvalue $\tilde{x}_{n+1,2}$. Indeed, since the eigenvalues of M_n are simple, the space of eigenvectors corresponding to a particular eigenvalue is one-dimensional. Using the notation (29) again, it is readily seen that

$$\mathbf{y}\mathbf{a}^T = 0 \quad \Longleftrightarrow \quad \sum_{i=0}^{n} \pi_i \sum_{j=0}^{i} u_j = 0 \quad \Longleftrightarrow \quad u_0 = -\frac{\sum_{i=1}^{n} \pi_i \sum_{j=1}^{i} u_j}{\sum_{i=0}^{n} \pi_i}.$$

Hence, if $\mathbf{y}\mathbf{a}^T = 0$ we have

$$\begin{aligned}
\mathbf{y}\mathbf{y}^T &= \sum_{i=0}^{n} \pi_i \left(\sum_{j=0}^{i} u_j \right)^2 \\
&= \sum_{i=0}^{n} \pi_i \left(u_0 + \sum_{j=1}^{i} u_j \right)^2 \\
&= \sum_{i=1}^{n} \pi_i \left(\sum_{j=1}^{i} u_j \right)^2 + 2u_0 \sum_{i=0}^{n} \pi_i \left(\sum_{j=0}^{i} u_j - u_0 \right) + u_0^2 \sum_{i=0}^{n} \pi_i \\
&= \sum_{i=1}^{n} \pi_i \left(\sum_{j=1}^{i} u_j \right)^2 - u_0^2 \sum_{i=0}^{n} \pi_i,
\end{aligned}$$

so that

$$\mathbf{y}\mathbf{y}^T = \sum_{i=1}^{n} \pi_i \left(\sum_{j=1}^{i} u_j \right)^2 - \frac{(\sum_{i=0}^{n} \pi_i \sum_{j=1}^{i} u_j)^2}{\sum_{i=0}^{n} \pi_i}.$$

The preceding observations can be summarised by stating that, if $\mu_0 = 0$,

$$\tilde{x}_{n+1,2} = \min_{\mathbf{u} \neq \mathbf{0}} \left\{ \frac{\sum_{k=0}^{n} \pi_k \sum_{i=1}^{n} \mu_i \pi_i u_i^2}{\sum_{k=0}^{n} \pi_k \sum_{i=1}^{n} \pi_i (\sum_{j=1}^{i} u_j)^2 - (\sum_{i=1}^{n} \pi_i \sum_{j=1}^{i} u_j)^2} \right\}, \tag{31}$$

where $\mathbf{u} = (u_1, u_2, \ldots, u_n)$. It follows that

$$\min_{\mathbf{u} \neq \mathbf{0}} \left\{ \frac{\sum_{i=1}^{n} \mu_i \pi_i u_i^2}{\sum_{i=1}^{n} \pi_i (\sum_{j=1}^{i} u_j)^2} \right\} \leq \tilde{x}_{n+1,2} \leq \min_{\mathbf{u} \neq \mathbf{0}} \left\{ \frac{\sum_{i=0}^{n} \pi_i \sum_{i=1}^{n} \mu_i \pi_i u_i^2}{\sum_{i=1}^{n} \pi_i (\sum_{j=1}^{i} u_j)^2} \right\}, \tag{32}$$

since, by the Cauchy–Schwarz inequality,

$$\sum_{k=1}^{n} \pi_k \sum_{i=1}^{n} \pi_i \left(\sum_{j=1}^{i} u_j \right)^2 - \left(\sum_{i=1}^{n} \pi_i \sum_{j=1}^{i} u_j \right)^2 \geq 0.$$

4. Proofs

In what follows we allow the birth–death process $\mathcal{X}$ to have $\mu_0 \geq 0$ and will use the superindex d bidirectionally to identify quantities related to the dual process. Note that

$$\begin{aligned}
\mu_0 > 0 &\implies \tilde{\xi}_i = \xi_i^d, \quad i \geq 1, \\
\mu_0 = 0 &\implies \tilde{\xi}_1 = 0, \quad \tilde{\xi}_{i+1} = \xi_i^d, \quad i \geq 1,
\end{aligned} \tag{33}$$

as a consequence of (20), (24), (25), and, (27). Before proving Theorem 1 we observe the following.

Proposition 1. *If $\mu_0 > 0$ and $\sum_n (\lambda_n \pi_n)^{-1} = \infty$, then $\alpha = \tilde{\xi}_1$.*

Proof. By (16) we have $\alpha = \xi_1$. Moreover, $\sum_n (\lambda_n \pi_n)^{-1} = \infty$ is equivalent to $\sum_n \pi_n^d = \infty$ by (22). Since $\mu_0^d = 0$, we conclude from (11) that $\psi^d(\{0\}) = 0$, so that 0 cannot be an isolated point in the support of ψ^d. Hence, either $\xi_1^d > 0$ or $\xi_1^d = \xi_2^d = \sigma^d = 0$, so that, by (23), $\xi_1 = \xi_1^d$. Finally, by (33), $\xi_1^d = \tilde{\xi}_1$, which establishes the result.

Proof of Theorem 1. Theorem 1 follows from the preceding result and representation (30). Indeed, let $C_\infty(\boldsymbol{u})$ and $C_n(\boldsymbol{u})$ denote the expressions between the braces in (5) and (30), respectively. Then, recalling definition (4), we have, for all $n \geq 0$,

$$\inf_{\boldsymbol{u} \in U} C_\infty(\boldsymbol{u}) \leq \inf_{\boldsymbol{u} \in U_n} C_\infty(\boldsymbol{u}) \leq \inf_{\boldsymbol{u} \in U_n} C_n(\boldsymbol{u}) = \tilde{x}_{n+1,1}, \tag{34}$$

so that $\inf_{\boldsymbol{u} \in U} C_\infty(\boldsymbol{u}) \leq \tilde{\xi}_1 = \lim_{n \to \infty} \tilde{x}_{n1}$. Assuming that $\inf_{\boldsymbol{u} \in U} C_\infty(\boldsymbol{u}) < \tilde{\xi}_1$, there must be an eventually vanishing sequence $\tilde{\boldsymbol{u}} \neq \boldsymbol{0}$ such that $C_\infty(\tilde{\boldsymbol{u}}) < \tilde{\xi}_1$, whence, $C_n(\tilde{\boldsymbol{u}}) < \tilde{\xi}_1$ for sufficiently large n. This, however, contradicts (30) and the fact that $\tilde{x}_{n1}$ decreases to $\tilde{\xi}_1$ as n tends to ∞, as a consequence of (26).

The second proposition leads to the proof of Theorem 4.

Proposition 2. *If $\mu_0 = 0$ and $\sum_n \pi_n < \infty$, then $\alpha = \tilde{\xi}_2$.*

Proof. We have $\xi_1 = 0$ in view of (11). Hence, $\alpha = \xi_2$ by (15), and $\xi_2 = \xi_1^d$ by (23). Finally, from (33) it follows that $\xi_1^d = \tilde{\xi}_2$, which proves the statement.

Proof of Theorem 4. The representation for α in Theorem 4 follows from the preceding result and representation (31) for $\tilde{x}_{n+1,2}$. Namely, with $D_\infty(\boldsymbol{u})$ and $D_n(\boldsymbol{u})$ denoting the expressions between the braces in (7) and (31), we have, for all $n \geq 0$,

$$\inf_{\boldsymbol{u} \in U} D_\infty(\boldsymbol{u}) \leq \inf_{\boldsymbol{u} \in U_n} D_\infty(\boldsymbol{u}) \leq \inf_{\boldsymbol{u} \in U_n} D_n(\boldsymbol{u}) = \tilde{x}_{n+1,2},$$

where the second inequality follows straightforwardly but, in contrast to the second inequality of (34), with somewhat cumbersome calculations. The remainder of the proof is similar to the proof of Theorem 1, now using the fact that $\tilde{x}_{n2}$ decreases to $\tilde{\xi}_2$ as n tends to ∞.

The lower bound for α in Theorem 4 is trivial, while the upper bound, as in (32), is implied by the Cauchy–Schwartz inequality.

Theorems 2 and 3 follow from Theorems 1 and 4 by duality.

Proof of Theorem 2. If $\mu_0 > 0$ and $\sum_n \lambda_n \pi_n < \infty$, then $\mu_0^d = 0$ and $\sum_n \pi_n^d < \infty$, so, by (11), $\xi_1^d = 0$. Moreover, by (16), (15), and (23), $\alpha = \xi_1 = \xi_2^d = \alpha^d$. So we can apply Theorem 4 to the dual process and obtain Theorem 2 after translation in terms of the original process.

Proof of Theorem 3. If $\mu_0 = 0$ and $\sum_n \pi_n = \infty$, then, by (11), $\psi(\{0\}) = 0$, implying that either $\xi_1 > 0$ or $\xi_1 = \xi_2 = \sigma = 0$. Moreover, $\mu_0^d > 0$ and $\sum_n \lambda_n^d \pi_n^d = \infty$, so, by (15) and (22), $\alpha = \xi_1 = \xi_1^d = \alpha^d$. Theorem 3 results from applying Theorem 1 to the dual process.

The corollaries can be proven in various ways. The most efficient proof uses the *weighted discrete Hardy's inequalities* given by Miclo [12, Proposition 1.1], which state that when

$\boldsymbol{v} = (v_0, v_1, \dots)$ and $\boldsymbol{w} = (w_0, w_1, \dots)$ are sequences of positive real numbers (weights), the smallest constant $A \le \infty$ such that, for all real sequences $\boldsymbol{u} = (u_0, u_1, \dots)$,

$$\sum_{i=0}^{\infty} v_i \left(\sum_{j=0}^{i} u_j \right)^2 \le A \sum_{i=0}^{\infty} w_i u_i^2,$$

satisfies

$$B \le A \le 4B, \tag{35}$$

where

$$B = \sup_{n \ge 0} \left\{ \sum_{i=0}^{n} \frac{1}{w_i} \sum_{i=n}^{\infty} v_i \right\}.$$

Proofs of Corollaries 1–4. To prove Corollary 1 we first observe that (5) may be reformulated as

$$\frac{1}{\alpha} = \inf \left\{ A \le \infty : \sum_{i=0}^{\infty} \pi_i \left(\sum_{j=0}^{i} u_j \right)^2 \le A \sum_{i=0}^{\infty} \mu_i \pi_i u_i^2 \text{ for all } \boldsymbol{u} \in U \right\}. \tag{36}$$

But it is easy to see that (36) remains valid if we require that the inequality should hold for *all* real sequences $\boldsymbol{u}$ instead of all real sequences $\boldsymbol{u} \ne \boldsymbol{0}$ that are eventually vanishing. Subsequently, using the weighted discrete Hardy's inequalities (35) with suitable interpretations for the weights, yields $R_0 \le \alpha^{-1} \le 4R_0$, establishing the corollary.

In the same way we can apply the weighted discrete Hardy's inequalities to α in the setting of Corollary 3, and to α_a of Theorem 2 and α_r of Theorem 4, establishing Corollaries 2–4.

We finally note that as a consequence of Theorems 1 and 3 we always have $\alpha = 0$ if $\sum_n \pi_n = \sum_n (\lambda_n \pi_n)^{-1} = \infty$. But this is also obvious from the fact that $\sigma = 0$ in this case (by (11), (12), and the fact that $\sigma = \sigma^d$). Thirdly, arguing probabilistically, $\alpha = 0$ is implied (if $\mu_0 = 0$) by $\int_0^{\infty} p_{00}(t) dt = \infty$, divergence of both sums being equivalent to null recurrence of the process.

5. Concluding remarks

In a series of papers published in Chinese journals since the early 1990s, Chen has studied, among related and more general issues, the problem of evaluating, or finding bounds for, the decay parameter of a birth–death process using the theory of *Dirichlet forms*. With the exception of [5] all of his publications involving birth–death processes pertain to ergodic processes (the setting of Theorem 4). The representation for α in Theorem 4 may be obtained already from results in [2], but the bounds of Corollary 4 appear for the first time in [3], together with some more refined (but less explicit) bounds. For a survey of Chen's results up to 2005, see [4]. Since then Chen's approach was adopted by Sirl *et al.* [13] in the setting of Theorem 1, resulting in the bounds in Corollary 1, and also in more refined bounds. Only recently, in the very comprehensive paper [5], Chen himself applied his methods to birth–death processes of all four types, yielding, among many more results, the bounds in Corollaries 2 and 3.

We also note that in [12], where Miclo develops the weighted discrete Hardy inequalities (35), the inequalities are actually applied to obtain bounds on the decay parameter of a birth–death process on the entire set of integers on the basis of a representation for α in terms of a Dirichlet form. Miclo suggests (page 324) that a similar approach may be applied in the setting of a birth–death process on the nonnegative integers, but does not supply explicit results.

Besides Dirichlet forms and the techniques used in this paper, there are many more approaches towards evaluation of the decay parameter of a birth–death process. For an overview of methods and results see [13].

Acknowledgements

References

[1] CALLAERT, H. (1974). On the rate of convergence in birth-and-death processes. *Bull. Soc. Math. Belg.* **26,** 173–184.

[2] CHEN, M.-F. (1991). Exponential L^2-convergence and L^2-spectral gap for Markov processes. *Acta Math. Sinica (N.S.)* **7,** 19–37.

[3] CHEN, M.-F. (2000). Explicit bounds of the first eigenvalue. *Sci. China Ser. A* **43,** 1051–1059.

[4] CHEN, M.-F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory*. Springer, London.

[5] CHEN, M.-F. (2010). Speed of stability for birth–death processes. *Front. Math. China* **5,** 379–515.

[6] CHIHARA, T. S. (1968). On indeterminate Hamburger moment problems. *Pacific J. Math.* **27,** 475–484.

[7] CHIHARA, T. S. (1978). *An Introduction to Orthogonal Polynomials*. Gordon and Breach, New York.

[8] CHIHARA, T. S. (1982). Indeterminate symmetric moment problems. *J. Math. Anal. Appl.* **85,** 331–346.

[9] KARLIN, S. AND MCGREGOR, J. L. (1957). The differential equations of birth-and-death processes, and the Stieltjes moment problem. *Trans. Amer. Math. Soc.* **85,** 489–546.

[10] KARLIN, S. AND MCGREGOR, J. L. (1957). The classification of birth and death processes. *Trans. Amer. Math. Soc.* **86,** 366–400.

[11] MEYER, C. (2000). *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, PA. Updates available at http://www.matrixanalysis.com.

[12] MICLO, L. (1999). An example of application of discrete Hardy's inequalities. *Markov Process. Relat. Fields* **5,** 319–330.

[13] SIRL, D., ZHANG, H. AND POLLETT, P. (2007). Computable bounds for the decay parameter of a birth–death process. *J. Appl. Prob.* **44,** 476–491.

[14] VAN DOORN, E. A. (1985). Conditions for exponential ergodicity and bounds for the decay parameter of a birth–death process. *Adv. Appl. Prob.* **17,** 514–530.

[15] VAN DOORN, E. A. (1987). The indeterminate rate problem for birth–death processes. *Pacific J. Math.* **130,** 379–393.

[16] VAN DOORN, E. A. AND POLLETT, P. K. (2013). Quasi-stationary distributions for discrete-state models. *Europ. J. Operat. Res.* **230,** 1–14.

[17] VAN DOORN, E. A. AND ZEIFMAN, A. I. (2009). On the speed of convergence to stationarity of the Erlang loss system. *Queueing Systems* **63,** 241–252.