

EXTINCTION PROBABILITY IN A BIRTH–DEATH PROCESS WITH KILLING

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

ALEXANDER I. ZEIFMAN,** *Vologda State Pedagogical University and*

Vologda Scientific Coordinate Centre of CEMI RAS

Abstract

We study birth–death processes on the nonnegative integers, where $\{1, 2, \dots\}$ is an irreducible class and 0 an absorbing state, with the additional feature that a transition to state 0 may occur from any state. We give a condition for absorption (extinction) to be certain and obtain the eventual absorption probabilities when absorption is not certain. We also study the rate of convergence, as $t \rightarrow \infty$, of the probability of absorption at time t , and relate it to the common rate of convergence of the transition probabilities that do not involve state 0. Finally, we derive upper and lower bounds for the probability of absorption at time t by applying a technique that involves the logarithmic norm of an appropriately defined operator.

Keywords: Absorption; decay parameter; extinction time; persistence time; rate of convergence; logarithmic norm

2000 Mathematics Subject Classification: Primary 60J80
Secondary 60J27

1. Introduction

We are concerned with a temporally homogeneous continuous-time Markov chain $\mathcal{X} := \{X(t), t \geq 0\}$ taking values in the set $S := \{0\} \cup C$, where $C := \{1, 2, \dots\}$ is an irreducible class and 0 an absorbing state. The q -matrix $Q := (q_{ij}, i, j \in S)$ of the chain is given by

$$\begin{aligned} q_{i,i+1} &= \lambda_i, & q_{i+1,i} &= \mu_{i+1}, & q_{i0} &= \gamma_i, & q_{ii} &= -(\lambda_i + \mu_i + \gamma_i), & i > 0; \\ q_{ij} &= 0, & |i - j| &> 1; & \text{and} & & q_{0j} &= 0, & j \geq 0, \end{aligned} \quad (1)$$

where $\lambda_i > 0$, $\mu_{i+1} > 0$, and $\gamma_i \geq 0$ for $i > 0$, and $\mu_1 = 0$. Following, for example, [17], we will refer to a process of this type as a *birth–death process with killing*. The parameters λ_i and μ_i are the *birth rate* and *death rate*, respectively, in state $i \in C$, while γ_i is the rate of absorption, or *killing rate*, from i into the absorbing state 0. Since, in state 1, ‘death’ and ‘killing’ have the same effect, the assumption that $\mu_1 = 0$ is no restriction of generality. Note that Q will be conservative over C if and only if $\gamma_i = 0$ for all $i \in C$. However, we will assume in what follows that $\gamma_i > 0$ for at least one state $i \in C$, so that 0 is accessible from C . We write $P_i(\cdot) := P\{\cdot \mid X(0) = i\}$.

Received 29 April 2004; revision received 1 September 2004.

* Postal address: Department of Applied Mathematics, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands. Email address: e.a.vandoorn@utwente.nl

** Postal address: Vologda State Pedagogical University, S. Orlova 6, Vologda, Russia.
Email address: zai@uni-vologda.ac.ru

We will assume that the process \mathcal{X} is nonexplosive (Q is *regular*) or, equivalently (see [4, Theorem 7]), that

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n \pi_n} \sum_{i=1}^n (1 + \gamma_i) \pi_i = \infty, \tag{2}$$

where

$$\pi_1 := 1, \quad \pi_i := \frac{\lambda_1 \lambda_2 \cdots \lambda_{i-1}}{\mu_2 \mu_3 \cdots \mu_i}, \quad i > 1. \tag{3}$$

Hence, the transition function $\mathbf{P}(\cdot) := \{p_{ij}(\cdot), i, j \in S\}$, where

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

is the unique Q -function (transition function with q -matrix Q), is honest and satisfies the system

$$\mathbf{P}'(t) = \mathbf{Q}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{Q}, \quad t \geq 0,$$

of backward *and* forward equations (see, for example, [1]). Here, a prime denotes elementwise differentiation.

By T we denote the *killing time*, that is, the (possibly defective) random variable representing the time at which absorption in state 0 occurs. In the terminology of population modelling, T is the *extinction time* or *persistence time*. In what follows we will be mainly interested in the functions

$$\tau_i(t) := P_i(T \leq t), \quad i \in C, t \geq 0,$$

and their limits

$$\tau_i := \lim_{t \rightarrow \infty} \tau_i(t), \quad i \in C.$$

We will refer to $\tau_i(t)$ and τ_i as the *extinction probability at time t* and the *eventual extinction probability*, respectively, when the initial state is i . Note that $\tau_i(t) = p_{i0}(t)$.

After collecting some preliminary results in the next section, we will obtain a necessary and sufficient condition for certain extinction and an explicit expression for the eventual extinction probability in Section 3. In Section 4, we address the problem of obtaining the rate of convergence of $\tau_i(t)$ to its limit. In a pure birth–death process ($\gamma_i = 0$ for $i > 1$) this rate equals the common rate of convergence of the transition probabilities $p_{ij}(t)$, $i, j \in C$, but this is not true, in general, in the present setting. We give a sufficient condition for equality of the rates of convergence. We also indicate how, if the rates are equal, results for pure birth–death processes may be invoked in the present setting. In Section 5, we derive bounds for the extinction probability $\tau_i(t)$ by applying the method developed by the second author in [27]–[29] to the model at hand, and indicate how the results may be generalized to inhomogeneous processes. We conclude with an example, in Section 6.

Apart from their intrinsic interest, our results are instructive because they are indicative of the phenomena that occur once one has wandered off the beaten track of the pure birth–death process.

2. Preliminaries

It is well known (see, for example, [1, Theorem 5.1.9]) that, under our assumptions regarding the Markov chain \mathcal{X} , there exist strictly positive constants c_{ij} (with $c_{ii} = 1$) and a parameter $\alpha \geq 0$ such that

$$p_{ij}(t) \leq c_{ij} e^{-\alpha t}, \quad i, j \in C, t \geq 0, \tag{4}$$

and

$$\alpha = - \lim_{t \rightarrow \infty} \frac{1}{t} \log p_{ij}(t), \quad i, j \in C. \tag{5}$$

The parameter α is known as the *decay parameter* of \mathcal{X} in C . It follows easily from (4) and (5) that α is also the rate of convergence to 0 of the transition probabilities $p_{ij}(t)$, in the sense that

$$\alpha = \inf \left\{ x \geq 0 : \int_0^\infty e^{xt} p_{ij}(t) dt = \infty \right\}, \quad i, j \in C. \tag{6}$$

The rate of convergence of the extinction probabilities $\tau_i(t)$ to their limits τ_i will be denoted by α_0 , that is,

$$\alpha_0 := \inf \left\{ x \geq 0 : \int_0^\infty e^{xt} (\tau_i - \tau_i(t)) dt = \infty \right\}, \quad i \in C.$$

It can be easily shown, by an irreducibility argument, that α_0 is independent of i .

The transition rates of \mathcal{X} determine polynomials R_n through the recurrence relation

$$\begin{aligned} \lambda_n R_{n+1}(x) &= (\lambda_n + \mu_n + \gamma_n - x)R_n(x) - \mu_n R_{n-1}(x), & n > 1, \\ \lambda_1 R_2(x) &= \lambda_1 + \gamma_1 - x, & R_1(x) = 1. \end{aligned} \tag{7}$$

Generalizing Karlin and McGregor’s [16] classic result, it is shown in [26] that the transition probabilities $p_{ij}(t)$, $i, j \in C$, may be represented in the form

$$p_{ij}(t) = \pi_j \int_0^\infty e^{-xt} R_i(x)R_j(x)\psi(dx), \quad t \geq 0, \tag{8}$$

where ψ is a Borel measure of total mass 1 on $[0, \infty)$ with respect to which the polynomials R_n are orthogonal. (The crux of the argument in [26] is that to each q -matrix of type (1), we can associate a unique q -matrix of type (1) that is conservative over C and such that the corresponding transition functions are *similar* in the sense of [21].) It can easily be shown, from [26, Theorem 4] and our Lemma 1 below, that, under our assumption (2), the orthogonalizing measure for $\{R_n\}$ is in fact unique. Since the transition probabilities $p_{ij}(t)$, $i, j \in C$, tend to 0 as t tends to infinity (recall our assumption that $\gamma_i > 0$ for at least one state i), the integral representation (8) tells us that the measure ψ cannot have a point mass at 0. It now follows readily from (6) and (8) that

$$\alpha = \min \text{supp}(\psi), \tag{9}$$

which generalizes an earlier result for birth–death processes (see, for example, [25, Theorem 3.1]).

Since orthogonal polynomials have no zeros outside the support of their orthogonalizing measure, while the smallest point of the support is a limit point of zeros (see, for example, [7, Section II.4]), (9) implies that

$$R_n(x) > 0 \text{ for all } n \geq 1 \Leftrightarrow x \leq \alpha. \tag{10}$$

It will also be useful to observe that

$$\lambda_n \pi_n (R_{n+1}(x) - R_n(x)) = \sum_{j=1}^n (\gamma_j - x) \pi_j R_j(x), \quad n \geq 1,$$

whence

$$R_n(x) = 1 + \sum_{k=1}^{n-1} \frac{1}{\lambda_k \pi_k} \sum_{j=1}^k (\gamma_j - x) \pi_j R_j(x), \quad n > 1.$$

It follows, in particular, that the quantities $r_n := R_n(0)$ satisfy

$$r_1 = 1 \quad \text{and} \quad r_n = 1 + \sum_{k=1}^{n-1} \frac{1}{\lambda_k \pi_k} \sum_{j=1}^k \gamma_j \pi_j r_j, \quad n > 1. \tag{11}$$

We let

$$r_\infty := \lim_{n \rightarrow \infty} r_n = 1 + \sum_{k=1}^{\infty} \frac{1}{\lambda_k \pi_k} \sum_{j=1}^k \gamma_j \pi_j r_j, \tag{12}$$

and note the following lemma.

Lemma 1. *We have $r_\infty = \infty$ if and only if*

$$\sum_{k=1}^{\infty} \frac{1}{\lambda_k \pi_k} \sum_{j=1}^k \gamma_j \pi_j = \infty. \tag{13}$$

Proof. The sufficiency is obvious because $r_n \geq 1$. So, let us define

$$\beta_k := \frac{1}{\lambda_k \pi_k} \sum_{j=1}^k \gamma_j \pi_j, \quad k \geq 1,$$

and assume that $\sum_{k \geq 1} \beta_k$ converges. Since r_n is increasing in n , we have

$$r_{n+1} = r_n + \frac{1}{\lambda_n \pi_n} \sum_{j=1}^n \gamma_j \pi_j r_j \leq r_n (1 + \beta_n), \quad n \geq 1,$$

so that

$$r_{n+1} \leq \prod_{k=1}^n (1 + \beta_k), \quad n \geq 1.$$

However, $\prod_{k \geq 1} (1 + \beta_k)$ and $\sum_{k \geq 1} \beta_k$ converge together, so we must have $r_\infty < \infty$, as required.

We conclude this section with representations for the extinction and eventual extinction probabilities. Indeed, the forward equations tell us that

$$p'_{i0}(t) = \sum_{j \in C} \gamma_j p_{ij}(t), \quad i \in C, t \geq 0.$$

It follows that

$$\tau_i(t) = p_{i0}(t) = \sum_{j \in C} \gamma_j \int_0^t p_{ij}(u) \, du, \quad i \in C, t \geq 0, \tag{14}$$

which, upon the substitution of (8) and the interchange of integrals, leads to

$$\tau_i(t) = \sum_{j \in C} \gamma_j \pi_j \int_0^\infty (1 - e^{-xt}) R_i(x) R_j(x) \frac{\psi(dx)}{x}, \quad i \in C, t \geq 0.$$

Letting $t \rightarrow \infty$ subsequently yields, by monotone convergence,

$$\tau_i = \sum_{j \in C} \gamma_j \pi_j \int_0^\infty R_i(x) R_j(x) \frac{\psi(dx)}{x}, \quad i \in C, \tag{15}$$

so that

$$\tau_i(t) = \tau_i - \sum_{j \in C} \gamma_j \pi_j \int_0^\infty e^{-xt} R_i(x) R_j(x) \frac{\psi(dx)}{x}, \quad i \in C, t \geq 0. \tag{16}$$

Expression (15) will be evaluated in the next section, and $\tau_i(t)$ will be studied in Sections 4 and 5.

3. Eventual extinction probability

We note that, by conditioning on the first event in \mathcal{X} (or using the recurrence relation (7) in (15)), the eventual extinction probabilities τ_i are readily seen to satisfy the recurrence

$$\begin{aligned} (\lambda_i + \mu_i + \gamma_i)\tau_i &= \lambda_i \tau_{i+1} + \mu_i \tau_{i-1} + \gamma_i, \quad i > 1, \\ (\lambda_1 + \gamma_1)\tau_1 &= \lambda_1 \tau_2 + \gamma_1. \end{aligned}$$

In view of (7), it follows that τ_i may be expressed in terms of τ_1 and $r_i := R_i(0)$ as

$$1 - \tau_i = (1 - \tau_1)r_i, \quad i \in C. \tag{17}$$

Since $\{\tau_i, i \in C\}$ constitutes the smallest nonnegative solution of (17) (cf. [10, p. 403]) we must have $\tau_i = 1 - r_i/r_\infty$, with the interpretation that $\tau_i = 1$ whenever $r_\infty = \infty$. This result may also be obtained from Lemma 3.1 of Brockwell [3], who studied eventual extinction probabilities in a more general setting (see also [1, Section 9.2]). Considering Lemma 1, we have at our disposal, in the present setting, a simpler criterion for certain extinction. We summarize our conclusions in the following theorem.

Theorem 1. *If (13) is satisfied then $\tau_i = 1$ for all $i \in C$. Otherwise, the eventual extinction probabilities satisfy*

$$\tau_i = 1 - \frac{r_i}{r_\infty} < 1, \quad i \in C,$$

with r_i and r_∞ given by (11) and (12), respectively.

In view of this result, the condition (2) for nonexplosiveness may be rephrased as follows. A necessary and sufficient condition for nonexplosiveness of \mathcal{X} is that either eventual extinction is certain or

$$\sum_{n=1}^\infty \frac{1}{\lambda_n \pi_n} \sum_{i=1}^n \pi_i = \infty.$$

As might be expected, the latter is precisely the condition for nonexplosiveness of $\mathcal{X}^* := [\mathcal{X} \mid T = \infty]$, the (pure birth–death) process one gets by setting $\gamma_i = 0$ for all $i \in C$ (see [1, Section 8.1]).

4. Rate of convergence

In addition to the accessibility of state 0, we will assume in this section that absorption at 0 is certain; that is, eventual extinction is certain and, hence, (13) is satisfied. Pakes [22, p. 122] has observed (see also [9]) that the latter assumption is no restriction because, if $\tau_i < 1$, we can work with the (Markov) process $\tilde{\mathcal{X}} := [\mathcal{X} \mid T < \infty]$, which has transition rates $\tilde{q}_{ij} = q_{ij}\tau_j/\tau_i$ and transition probabilities $\tilde{p}_{ij}(t) = p_{ij}(t)\tau_j/\tau_i$. Here, $\tau_0 := 1$ and $\tau_i > 0$ because of our accessibility assumption. It follows that

$$\bar{\tau}_i(t) := \bar{p}_{i0}(t) = p_{i0}(t)/\tau_i = \tau_i(t)/\tau_i \rightarrow 1 \quad \text{as } t \rightarrow \infty, \text{ for } i \in C.$$

We note from (16) that $\xi_i(t) := 1 - \tau_i(t) = P_i(T > t)$, the *survival probability* at time t , can be represented in the form

$$\xi_i(t) = \sum_{j \in C} \gamma_j \pi_j \int_0^\infty e^{-xt} R_i(x) R_j(x) \frac{\psi(dx)}{x}, \quad i \in C, t \geq 0. \tag{18}$$

In view of (9) (recall that ψ does not have a point mass at 0), it is therefore tempting to believe that $\alpha_0 = \alpha$, but this is not true in general. Since $1 \geq \xi_i(t) \geq p_{ii}(t)$ we do know, however, that

$$0 \leq \alpha_0 \leq \alpha. \tag{19}$$

This was observed by Kingman [20, Theorem 8] and, more recently, by Jacka and Roberts [15, Equation (3.1.4)], whose example with strict inequalities in (19) is encompassed in the setting that is described next.

Suppose that the killing rates satisfy $\gamma_i \geq \gamma > 0$ for all $i \in C$. Then we may look upon the process \mathcal{X} as a birth–death process with killing, $\tilde{\mathcal{X}}$ say, with rates $\tilde{\lambda}_i := \lambda_i$, $\tilde{\mu}_i := \mu_i$, and $\tilde{\gamma}_i := \gamma_i - \gamma$, which is subject to an additional killing event taking place at rate γ . Evidently, absorption of \mathcal{X} at 0 is certain. By conditioning on the time of the additional killing event, we have $p_{ij}(t) = e^{-\gamma t} \tilde{p}_{ij}(t)$, $i, j \in C$, and, hence,

$$\alpha(\mathcal{X}) = \gamma + \alpha(\tilde{\mathcal{X}}),$$

where the argument of α indicates the relevant process. By conditioning again, we also obtain

$$\xi_i(t) = e^{-\gamma t} (1 - \tilde{\tau}_i(t)) = e^{-\gamma t} (1 - \tilde{\tau}_i) + e^{-\gamma t} (\tilde{\tau}_i - \tilde{\tau}_i(t)), \quad i \in C, t \geq 0,$$

where $\tilde{\tau}_i(t)$ is the extinction probability at time t of the process $\tilde{\mathcal{X}}$ and $\tilde{\tau}_i$ is its limit as $t \rightarrow \infty$. Hence,

$$\alpha_0(\mathcal{X}) = \begin{cases} \gamma & \text{if } \tilde{\tau}_1 < 1, \\ \gamma + \alpha_0(\tilde{\mathcal{X}}) & \text{if } \tilde{\tau}_1 = 1. \end{cases}$$

It follows that strict inequalities apply in (19) when $\tilde{\tau}_1 < 1$ and $\alpha(\tilde{\mathcal{X}}) > 0$. We note, in addition, that the calculation of $\alpha_0(\mathcal{X})$ is reduced to the calculation of $\alpha_0(\tilde{\mathcal{X}})$ if $\tilde{\tau}_1 = 1$.

It has been shown in [15] (in a more general setting and implicitly assuming certain absorption) that we have $\alpha_0 = \alpha$ if only finitely many of the γ_i are positive, which is obvious from the representation (18). The following theorem is a more general result.

Theorem 2. *If $\alpha > 0$ and eventual extinction is certain, then we have*

$$\sum_{j \in C} \gamma_j \pi_j R_j(\alpha) = \alpha \sum_{j \in C} \pi_j R_j(\alpha), \tag{20}$$

and $\alpha_0 = \alpha$ whenever either sum in (20) converges.

Proof. Recalling that $R_j(\alpha) > 0$ and using an argument similar to that in the proof of [24, Theorem 4.1], it is not difficult to show, using (8), that, if $\alpha > 0$,

$$q_j := \lim_{t \rightarrow \infty} \frac{p_{ij}(t)}{\sum_{k \in C} p_{ik}(t)} = \frac{\pi_j R_j(\alpha)}{\sum_{k \in C} \pi_k R_k(\alpha)}, \quad j \in C, \tag{21}$$

which is to be interpreted as 0 if the sum diverges. On the other hand, since extinction is certain, we have $\sum_{j \in C} p_{ij}(t) = \xi_i(t)$ and, hence, we may use the representation (18) to calculate q_j in a similar fashion, yielding

$$q_j = \lim_{t \rightarrow \infty} \frac{p_{ij}(t)}{\xi_i(t)} = \frac{\alpha \pi_j R_j(\alpha)}{\sum_{k \in C} \gamma_k \pi_k R_k(\alpha)}, \quad j \in C, \tag{22}$$

which is again to be interpreted as 0 if the sum diverges. Since the two limits must be equal, (20) must hold. Moreover, if either sum in (20) converges then $q_j > 0$ (and (21) tells us that $\{q_j, j \in C\}$ in fact constitutes a proper distribution). Evidently (see also [15, Theorem 3.3.2(ii)]), the latter is a sufficient condition for $\alpha_0 = \alpha$.

Remark 1. Theorem 2 generalizes part of the lemma in [11] (see also [24, Theorem 3.2]), which concerns pure birth–death processes. When $\gamma_i > 0$ for infinitely many states i , the situation differs essentially from the pure birth–death setting in that we may simultaneously have both $\alpha > 0$ and divergence of the series in (20). If either series in (20) converges then the quantities q_j of (21) (or (22)) constitute a *quasi-stationary distribution* (see, for example, [22]). In this case, we also have

$$\alpha_0 = \alpha = - \lim_{t \rightarrow \infty} \frac{1}{t} \log P_i(T > t)$$

(see [22, Lemma 2.1]).

If $\alpha_0 = \alpha$ then the problem of determining α_0 can be reduced to that of finding the decay parameter in a pure birth–death process, for which many results are available (see [5], [6], [12], [18], [19], [23], [25], and [27]–[29]). Indeed, define $\tilde{\mathcal{X}} := \{\tilde{X}(t), t \geq 0\}$ to be the birth–death process on C with birth and death rates

$$\tilde{\lambda}_i := \lambda_i \frac{r_{i+1}}{r_i} \quad \text{and} \quad \tilde{\mu}_{i+1} := \mu_{i+1} \frac{r_i}{r_{i+1}}, \quad i \in C, \tag{23}$$

respectively, where $r_i := R_i(0)$ as before. Letting $\tilde{\mu}_1 = \mu_1 = 0$, it is easy to see, from (23) and (7), that

$$\tilde{\lambda}_i \tilde{\mu}_{i+1} = \lambda_i \mu_{i+1} \quad \text{and} \quad \tilde{\lambda}_i + \tilde{\mu}_i = \lambda_i + \mu_i + \gamma_i, \quad i \in C.$$

By [26, Theorem 1], this implies that there are constants $\sigma_{ij} > 0$ such that

$$p_{ij}(t) = \sigma_{ij} \tilde{p}_{ij}(t), \quad i, j \in C, t \geq 0,$$

with $\tilde{p}_{ij}(t)$ denoting the transition probabilities of $\tilde{\mathcal{X}}$. (In the terminology of [21], the processes \mathcal{X} and $\tilde{\mathcal{X}}$ are *similar*). Consequently, \mathcal{X} and $\tilde{\mathcal{X}}$ have the same decay parameter.

5. Bounds for the survival probability

To obtain bounds for $\xi_i(t)$, we choose the approach used in [27]–[29] for pure birth–death processes (see also [13] or [14] for an exposition of this method). Application of the technique requires the elements of the q -matrix \mathbf{Q} to be bounded so, in what follows, we assume that

$$\sup_i \{\lambda_i + \mu_i + \gamma_i\} < \infty.$$

We let $\mathbf{A} := (q_{ij}, i, j \in C)$, the matrix that remains after removing the first row and column from \mathbf{Q} , and define

$$\mathbf{x}_i(t) := (p_{i1}(t), p_{i2}(t), \dots)^\top, \quad i \in C, t \geq 0,$$

where superscript ‘ \top ’ denotes transpose. Further, let $\mathbf{D} := \text{diag}(d_1, d_2, \dots)$, with d_1, d_2, \dots denoting positive parameters, and let $\mathbf{z}_i(t) := \mathbf{D}\mathbf{x}_i(t)$. The forward equations for $\mathbf{P}(\cdot)$ then tell us that

$$\mathbf{z}'_i(t) = \mathbf{D}\mathbf{A}\mathbf{D}^{-1}\mathbf{z}_i(t), \quad i \in C, t \geq 0.$$

If the parameters d_i are such that $\mathbf{D}\mathbf{A}\mathbf{D}^{-1}$ can be interpreted as a bounded linear operator on a normed space, then the theory expounded, for example, in [29] and [13] reveals that, for all $i \in C$ and $t \geq 0$,

$$\exp\{-t\theta^*(\mathbf{d})\}\|\mathbf{z}_i(0)\| \leq \|\mathbf{z}_i(t)\| \leq \exp\{tg(\mathbf{D}\mathbf{A}\mathbf{D}^{-1})\}\|\mathbf{z}_i(0)\|. \tag{24}$$

Here,

$$\theta^*(\mathbf{d}) := \sup_{i \in C} \left\{ \lambda_i + \mu_i + \gamma_i - \lambda_i \frac{d_{i+1}}{d_i} - \mu_i \frac{d_{i-1}}{d_i} \right\}, \tag{25}$$

with $\mathbf{d} := (d_1, d_2, \dots)$ and $d_0 := 0$, and

$$g(\mathbf{D}\mathbf{A}\mathbf{D}^{-1}) := \lim_{h \downarrow 0} \frac{\|\mathbf{I} + h\mathbf{D}\mathbf{A}\mathbf{D}^{-1}\| - 1}{h},$$

where \mathbf{I} is the identity matrix, is the *logarithmic norm* of the operator $\mathbf{D}\mathbf{A}\mathbf{D}^{-1}$. Moreover, choosing $\|\cdot\| = \|\cdot\|_1$, the ℓ_1 -norm, we have

$$-g(\mathbf{D}\mathbf{A}\mathbf{D}^{-1}) = \theta(\mathbf{d}) := \inf_{i \in C} \left\{ \lambda_i + \mu_i + \gamma_i - \lambda_i \frac{d_{i+1}}{d_i} - \mu_i \frac{d_{i-1}}{d_i} \right\}. \tag{26}$$

Hence, (24) translates into

$$d_i e^{-\theta^* t} \leq \sum_{j \in C} d_j p_{ij}(t) \leq d_i e^{-\theta t}, \quad i \in C, t \geq 0, \tag{27}$$

where $\theta := \theta(\mathbf{d})$ and $\theta^* := \theta^*(\mathbf{d})$. As an aside, we note that $\theta(\mathbf{d}) = \theta^*(\mathbf{d}) = x$ if and only if $d_i = cR_i(x)$ for some constant c , as can easily be seen from the recurrence relation (7). It follows, in particular, that

$$\sum_{j \in C} R_j(x) p_{ij}(t) = R_i(x) e^{-xt}, \quad i \in C, t \geq 0,$$

from which the representation (8) may be derived (cf. [16, Section I.2]).

Since

$$\xi_i(t) := P_i(T > t) = \sum_{j \in C} p_{ij}(t),$$

the inequalities (27) immediately give us the following bounds for the survival probability $\xi_i(t)$.

Theorem 3. (i) Let $d_j \geq 1$ for all $j \in C$ and $\theta := \theta(\mathbf{d})$ as in (26). Then,

$$\xi_i(t) \leq d_i e^{-\theta t}, \quad i \in C, t \geq 0. \tag{28}$$

(ii) Let $d_j \leq 1$ for all $j \in C$ and $\theta^* := \theta^*(\mathbf{d})$ as in (25). Then

$$\xi_i(t) \geq d_i e^{-\theta^* t}, \quad i \in C, t \geq 0. \tag{29}$$

Note that eventual extinction must be certain when $d_j \geq 1$ for all $j \geq 1$ and $\theta(\mathbf{d}) > 0$.

Corollary 1. If the constants $\mu \geq 0$ and $a \geq 0$ are such that

$$\mu < \mu_{j+1} \quad \text{and} \quad a \leq \mu + \gamma_j - \frac{\lambda_j \mu}{\mu_{j+1} - \mu}, \quad j = 1, 2, \dots,$$

then

$$\xi_i(t) \leq e^{-at} \prod_{j=1}^i \frac{\mu_{j+1}}{\mu_{j+1} - \mu}, \quad i \in C, t \geq 0.$$

Proof. Choosing $d_1 = 1$ and $d_{j+1}/d_j = \mu_{j+1}/(\mu_{j+1} - \mu)$ for $j \geq 1$, we have $d_j \geq 1$ and

$$\theta(\mathbf{d}) = \inf_{j \in C} \left\{ \mu + \gamma_j - \frac{\lambda_j \mu}{\mu_{j+1} - \mu} \right\},$$

so that the conditions of Theorem 3(i) are satisfied. Substitution into (28) gives the result.

Taking $\mu = 0$, it follows, in particular, that $\xi_i(t) \leq e^{-at}$ if $a \leq \inf\{\gamma_j\}$, as we observed by a different argument in the previous section.

If α , the decay parameter of \mathcal{X} in C , is known, then the following corollary might be useful. Recall that $R_j(\alpha) > 0$, by (10).

Corollary 2. If $0 \leq R_{\min} < R_j(\alpha) < R_{\max} \leq \infty$ for all j , then

$$\frac{R_i(\alpha)}{R_{\max}} e^{-\alpha t} < \xi_i(t) < \frac{R_i(\alpha)}{R_{\min}} e^{-\alpha t}, \quad i \in C, t \geq 0,$$

where the left-hand side should be interpreted as 0 if $R_{\max} = \infty$ and the right-hand side should be interpreted as infinity if $R_{\min} = 0$.

Proof. We have already noted that letting $d_j = cR_j(x)$ for some constant c gives us $\theta(\mathbf{d}) = \theta^*(\mathbf{d}) = x$. Hence, if $R_j(\alpha) > R_{\min} > 0$ for all j , then the conditions of Theorem 3(i) are satisfied if we choose $a = \alpha$ and $d_j = R_j(\alpha)/R_{\min}$, and substitution in (28) gives the upper bound. On the other hand, if $R_j(\alpha) < R_{\max} < \infty$ for all j , then the conditions of Theorem 3(ii) are satisfied if we choose $a = \alpha$ and $d_j = R_j(\alpha)/R_{\max}$, and substitution in (29) gives the lower bound.

Under certain circumstances, (27) may lead to other bounds for $\xi_i(t)$. For example, suppose that $\gamma_i > 0$ for all $i \in C$ and choose $d_i = \gamma_i$ in (26) and (27), so that

$$\gamma_i e^{-\theta^* t} \leq \sum_{j \in C} \gamma_j p_{ij}(t) \leq \gamma_i e^{-\theta t}, \quad i \in C, t \geq 0,$$

where $\theta := \theta(\boldsymbol{\gamma})$, $\theta^* := \theta^*(\boldsymbol{\gamma})$, and $\boldsymbol{\gamma} := (\gamma_1, \gamma_2, \dots)$. If $\theta > 0$ we obtain, in view of (14),

$$1 - \frac{\gamma_i}{\theta} (1 - e^{-\theta t}) \leq \xi_i(t) \leq 1 - \frac{\gamma_i}{\theta^*} (1 - e^{-\theta^* t}), \quad i \in C, t \geq 0.$$

At the other extreme, suppose that $\gamma_i = 0$ for $i > 1$; that is, we are dealing with a pure birth–death process. Now choose $d_i \leq d_1$ (for all i) in (26) and (27), and suppose that $\theta := \theta(\mathbf{d}) > 0$. Then we have

$$\gamma_1 p_{i1}(t) \leq \frac{\gamma_1}{d_1} \sum_{j \in C} d_j p_{ij}(t) \leq \gamma_1 \frac{d_i}{d_1} e^{-\theta t}, \quad i \in C, t \geq 0,$$

by (27) and, hence, by (14),

$$\xi_i(t) \geq 1 - \frac{\gamma_1}{\theta} \frac{d_i}{d_1} (1 - e^{-\theta t}), \quad i \in C, t \geq 0.$$

We conclude this section by noting that the result (27) can easily be generalized to inhomogeneous processes. Specifically, let \mathcal{X} be a birth–death process with killing with time-dependent birth rates $\lambda_i(t)$, death rates $\mu_i(t)$, and killing rates $\gamma_i(t)$. Then, under appropriate boundedness conditions and for all $i \in C$ and $t \geq 0$,

$$d_i \exp \left\{ - \int_0^t \theta^*(\mathbf{d}, u) du \right\} \leq \sum_{j \in C} d_j p_{ij}(t) \leq d_i \exp \left\{ - \int_0^t \theta(\mathbf{d}, u) du \right\},$$

where

$$\theta(\mathbf{d}, t) := \inf_{i \in C} \left\{ \lambda_i(t) + \mu_i(t) + \gamma_i(t) - \lambda_i(t) \frac{d_{i+1}}{d_i} - \mu_i(t) \frac{d_{i-1}}{d_i} \right\}, \quad t \geq 0,$$

and

$$\theta^*(\mathbf{d}, t) := \sup_{i \in C} \left\{ \lambda_i(t) + \mu_i(t) + \gamma_i(t) - \lambda_i(t) \frac{d_{i+1}}{d_i} - \mu_i(t) \frac{d_{i-1}}{d_i} \right\}, \quad t \geq 0.$$

The corresponding generalizations of Theorem 4 and Corollary 1 are straightforward.

6. Example

Interesting cases arise if $\gamma_i > 0$ for infinitely many states i , while γ_i is not constant for all i . We will analyse a simple example satisfying these conditions, namely the process with transition rates

$$\lambda_i = \lambda, \quad \mu_i = \mu \mathbf{1}_{\{i>1\}}, \quad \text{and} \quad \gamma_i = \gamma \mathbf{1}_{\{i>1\}}, \quad i \in C, \quad (30)$$

for some constants $\lambda > 0$, $\mu > 0$, and $\gamma > 0$, where $\mathbf{1}_E$ denotes the indicator function of an event E . It is easily seen that (13) is satisfied, so that extinction is certain. The polynomials R_n of (7) satisfy the recurrence relation

$$\begin{aligned} \lambda R_{n+1}(x) &= (\lambda + \mu + \gamma - x)R_n(x) - \mu R_{n-1}(x), & n > 1, \\ \lambda R_2(x) &= \lambda - x, & R_1(x) = 1, \end{aligned}$$

which, by the transformation

$$S_n(x) := (-1)^n \left(\frac{\lambda}{\mu}\right)^{n/2} R_{n+1}(\lambda + \mu + \gamma + 2x\sqrt{\lambda\mu}), \quad n \geq 0, \tag{31}$$

reduces to

$$\begin{aligned} S_n(x) &= 2xS_{n-1}(x) - S_{n-2}(x), & n > 1, \\ S_1(x) &= 2x + \eta, & S_0(x) = 1, \end{aligned}$$

where

$$\eta := \frac{\mu + \gamma}{\sqrt{\lambda\mu}}.$$

The polynomials S_n can be represented as

$$S_n(x) = U_n(x) + \eta U_{n-1}(x), \quad n \geq 1, \tag{32}$$

where $U_n(x)$ denote the Chebyshev polynomials of the second kind. The latter satisfy the recurrence

$$\begin{aligned} U_n(x) &= 2xU_{n-1}(x) - U_{n-2}(x), & n > 1, \\ U_1(x) &= 2x, & U_0(x) = 1, \end{aligned}$$

and may be represented as

$$U_n(x) = \frac{z^{n+1} - z^{-(n+1)}}{z - z^{-1}}, \quad n \geq 0, \quad \text{where } x = \frac{1}{2}(z + z^{-1}). \tag{33}$$

It will be useful to observe that

$$U_n(x) = (-1)^n U_n(-x) \quad \text{and} \quad U_n(1) = n + 1. \tag{34}$$

By appropriately transforming the orthogonalizing measure for $\{S_n(x)\}$ given in [7, p. 205], we can conclude that the polynomials R_n are orthogonal with respect to a measure consisting of a positive density on the interval

$$(\lambda + \mu + \gamma - 2\sqrt{\lambda\mu}, \lambda + \mu + \gamma + 2\sqrt{\lambda\mu})$$

and, if $\mu + \gamma > (\lambda\mu)^{1/2}$, a point mass at $\lambda\gamma/(\mu + \gamma)$. Thus, since

$$\frac{\lambda\gamma}{\mu + \gamma} = \lambda + \mu + \gamma - \sqrt{\lambda\mu}(\eta + \eta^{-1}),$$

it follows from (9) that

$$\alpha = \lambda + \mu + \gamma - \begin{cases} 2\sqrt{\lambda\mu} & \text{if } \mu + \gamma \leq \sqrt{\lambda\mu}, \\ \sqrt{\lambda\mu}(\eta + \eta^{-1}) & \text{if } \mu + \gamma \geq \sqrt{\lambda\mu}. \end{cases}$$

We next wish to determine the value of α_0 . To this end, we will not try to employ (18), but rather argue as follows. Let E_a denote an exponentially distributed random variable with mean a^{-1} , and B a random variable representing the busy period in an M/M/1 queueing system with arrival rate λ and service rate μ . (If $\lambda > \mu$ the distribution of B is defective.) A little reflection then shows that, if the initial state is 1, the extinction time T may be represented as

$$T = E_\lambda + E_\gamma \mathbf{1}_{\{E_\gamma \leq B\}} + (B + T^*) \mathbf{1}_{\{E_\gamma > B\}},$$

where T^* is a random variable that is independent of T but has the same distribution. It follows that

$$\begin{aligned} \tilde{\tau}(s) &:= E[e^{-sT}] \\ &= E[e^{-sT} \mathbf{1}_{\{E_\gamma \leq B\}} + e^{-sT} \mathbf{1}_{\{E_\gamma > B\}}] \\ &= E[e^{-sE_\lambda} (e^{-sE_\gamma} \mathbf{1}_{\{E_\gamma \leq B\}} + e^{-s(B+T^*)} \mathbf{1}_{\{E_\gamma > B\}})] \\ &= \frac{\lambda}{\lambda + s} (E[e^{-sE_\gamma} \mathbf{1}_{\{E_\gamma \leq B\}}] + \tilde{\tau}(s) E[e^{-sB} \mathbf{1}_{\{E_\gamma > B\}}]), \end{aligned}$$

so that

$$(\lambda + s - \lambda E[e^{-sB} \mathbf{1}_{\{E_\gamma > B\}}]) \tilde{\tau}(s) = \lambda E[e^{-sE_\gamma} \mathbf{1}_{\{E_\gamma \leq B\}}]. \tag{35}$$

A little algebra reveals that

$$E[e^{-sE_\gamma} \mathbf{1}_{\{E_\gamma \leq B\}}] = \frac{\gamma}{\gamma + s} (1 - \tilde{B}(\gamma + s))$$

and

$$E[e^{-sB} \mathbf{1}_{\{E_\gamma > B\}}] = \tilde{B}(\gamma + s),$$

where $\tilde{B}(s) := E[e^{-sB}]$. Substitution of these results into (35) gives us

$$\tilde{\tau}(s) = \frac{\gamma(\lambda - \lambda \tilde{B}(\gamma + s))}{(\gamma + s)(\lambda + s - \lambda \tilde{B}(\gamma + s))}. \tag{36}$$

It is well known (see, for instance, [8, Equation (II.2.31)]) that

$$\tilde{B}(s) = \frac{1}{2\lambda} (\lambda + \mu + s - \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu}),$$

which, upon substitution into (36) and some algebra, leads to

$$\tilde{\tau}(s) = \frac{\gamma(s^2 + (\lambda + \mu + \gamma)s + 2\lambda\gamma - s\sqrt{(\lambda + \mu + \gamma + s)^2 - 4\lambda\mu})}{2(\gamma + s)(\lambda\gamma + (\mu + \gamma)s)}.$$

By inverting this expression, we can obtain an explicit formula for $\tau_1(t)$, the extinction time distribution when the initial state is 1. At this point, however, we are interested only in α_0 , the rate of convergence of $\tau_1(t)$, which, up to a minus sign, equals the singularity of $\tilde{\tau}(s)$ that is closest to the imaginary axis. Since the largest branch point at $-\gamma - (\lambda^{1/2} - \mu^{1/2})^2$ is always smaller than the pole at $-\gamma$, it follows that $\alpha_0 = \gamma$ or $\alpha_0 = \lambda\gamma/(\mu + \gamma)$ if $\lambda \geq \mu + \gamma$ or $\lambda \leq \mu + \gamma$, respectively.

We collect all our results in the following theorem.

Theorem 4. *The process with transition rates (30) has rates of convergence α_0 and α given by*

$$\begin{aligned} \alpha_0 = \alpha &= \frac{\lambda\gamma}{\mu + \gamma} \quad \text{if } \lambda \leq \mu + \gamma, \\ \alpha_0 = \gamma < \alpha &= \frac{\lambda\gamma}{\mu + \gamma} \quad \text{if } \sqrt{\lambda\mu} \leq \mu + \gamma < \lambda, \\ \alpha_0 = \gamma < \alpha &= \gamma + (\sqrt{\lambda} - \sqrt{\mu})^2 \quad \text{if } \mu + \gamma < \sqrt{\lambda\mu}. \end{aligned}$$

Observe that our findings are in accordance with the intuitive result that α_0 must tend to 0 as γ tends to 0.

It is interesting to establish how much of the information in Theorem 4 may be obtained from Theorem 2. To this end, we note that, by (3) and (30), $\pi_{n+1} = (\lambda/\mu)^n, n \geq 0$, so that, by (31),

$$\pi_{n+1} R_{n+1}(x) = (-1)^n \left(\frac{\lambda}{\mu}\right)^{n/2} S_n\left(\frac{x - \lambda - \mu - \gamma}{2\sqrt{\lambda\mu}}\right), \quad n \geq 0.$$

Hence, after some algebra, it follows from (32), (33), and (34) that, for $n \geq 0$,

$$\pi_{n+1} R_{n+1}(\alpha) = \begin{cases} (1 + (1 - \eta)n) \left(\frac{\lambda}{\mu}\right)^{n/2} & \text{if } \mu + \gamma \leq \sqrt{\lambda\mu}, \\ \left(\frac{\lambda}{\mu + \gamma}\right)^n & \text{if } \mu + \gamma \geq \sqrt{\lambda\mu}. \end{cases} \tag{37}$$

Since $(\lambda\mu)^{1/2} < \mu + \gamma$ if $\lambda < \mu + \gamma$, while $\lambda > \mu$ if $\lambda \geq \mu + \gamma$, we conclude that the series in (20) converge if and only if $\lambda < \mu + \gamma$. Hence, Theorem 2 tells us that $\alpha_0 = \alpha$ if $\lambda < \mu + \gamma$. In the opposite case, Theorem 2 does not help us.

By extending the method by which we have calculated $\tilde{\tau}(s)$, we can obtain the Laplace–Stieltjes transform of the extinction time distribution when the initial state is any state $i \in C$ other than 1. By inversion, we can therefore calculate $\tau_i(t)$ and, hence, $\xi_i(t)$, at least in principle. However, the procedure is cumbersome, so it is of interest to apply the methodology of Section 5 to the present example. For instance, choosing $d_1 = 1$ and

$$d_{j+1} = \left(\frac{\mu}{\mu + \gamma}\right)^j, \quad j \geq 1,$$

in (25) gives us $\theta^* = \lambda\gamma/(\mu + \gamma)$ and, hence, by Theorem 3(ii),

$$\xi_i(t) \geq \left(\frac{\mu}{\mu + \gamma}\right)^{i-1} \exp\left\{\frac{-\lambda\gamma t}{\mu + \gamma}\right\}, \quad i \in C, t \geq 0. \tag{38}$$

This is also the bound produced by Corollary 2 when $\mu + \gamma \geq (\lambda\mu)^{1/2}$. In the case that $\mu + \gamma < (\lambda\mu)^{1/2}$, Corollary 2 yields a lower bound that improves upon (38) for t sufficiently large.

As an aside, we finally note that ours is yet another example, along with those of [22] and [2], showing that *asymptotic remoteness*, that is,

$$\lim_{i \rightarrow \infty} p_{i0}(t) = 0, \quad t \geq 0, \tag{39}$$

is not necessary for the existence of a quasi-stationary distribution. Indeed, it is obvious that (39) is not satisfied in the present setting, while, in view of (37) and Remark 1, a quasi-stationary distribution does exist when $\lambda < \mu + \gamma$.

References

- [1] ANDERSON, W. J. (1991). *Continuous-Time Markov Chains*. Springer, New York.
- [2] BOBROWSKI, A. (2004). Quasi-stationary distributions of a pair of Markov chains related to time evolution of a DNA locus. *Adv. Appl. Prob.* **36**, 56–77.
- [3] BROCKWELL, P. J. (1986). The extinction time of a general birth and death process with catastrophes. *J. Appl. Prob.* **23**, 851–858.
- [4] CHEN, A., POLLETT, P., ZHANG, H. AND CAIRNS, B. (2004). Uniqueness criteria for continuous-time Markov chains with general transition structure. Preprint, University of Queensland.
- [5] CHEN, M. F. (1991). Exponential L^2 -convergence and L^2 -spectral gap for Markov processes. *Acta Math. Sinica (N.S.)* **7**, 19–37.
- [6] CHEN, M. F. (1996). Estimation of the spectral gap for Markov chains. *Acta Math. Sinica (N.S.)* **12**, 337–360.
- [7] CHIHARA, T. S. (1978). *An Introduction to Orthogonal Polynomials*. Gordon and Breach, New York.
- [8] COHEN, J. W. (1982). *The Single Server Queue* (North-Holland Ser. Appl. Math. Mech. **8**), 2nd edn. North-Holland, Amsterdam.
- [9] ELMES, S., POLLETT, P. AND WALKER, D. (2000). Further results on the relationship between μ -invariant measures and quasi-stationary distributions for absorbing continuous-time Markov chains. *Math. Comput. Modelling* **31**, 107–113.
- [10] FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd edn. John Wiley, New York.
- [11] GOOD, P. (1968). The limiting behavior of transient birth and death processes conditioned on survival. *J. Austral. Math. Soc.* **8**, 716–722.
- [12] GRANOVSKY, B. L. AND ZEIFMAN, A. I. (1997). The decay function of nonhomogeneous birth–death processes, with applications to mean-field models. *Stoch. Process. Appl.* **72**, 105–120.
- [13] GRANOVSKY, B. L. AND ZEIFMAN, A. I. (2000). The N -limit of spectral gap of a class of birth–death Markov chains. *Appl. Stoch. Models Business Industry* **16**, 235–248.
- [14] GRANOVSKY, B. L. AND ZEIFMAN, A. I. (2004). Nonstationary queues: estimation of the rate of convergence. *Queueing Systems* **46**, 363–388.
- [15] JACKA, S. D. AND ROBERTS, G. O. (1995). Weak convergence of conditioned processes on a countable state space. *J. Appl. Prob.* **32**, 902–916.
- [16] 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**, 589–646.
- [17] KARLIN, S. AND TAVARÉ, S. (1982). Linear birth and death processes with killing. *J. Appl. Prob.* **19**, 477–487.
- [18] KARTASHOV, N. V. (1998). Calculation of the exponential ergodicity exponent for birth–death processes. *Theory Prob. Math. Statist.* **57**, 53–60.
- [19] KIJIMA, M. (1992). Evaluation of the decay parameter for some specialized birth–death processes. *J. Appl. Prob.* **29**, 781–791.
- [20] KINGMAN, J. F. C. (1963). The exponential decay of Markov transition probabilities. *Proc. London Math. Soc.* **13**, 337–358.
- [21] LENIN, R. B., PARTHASARATHY, P. R., SCHEINHARDT, W. R. W. AND VAN DOORN, E. A. (2000). Families of birth–death processes with similar time-dependent behaviour. *J. Appl. Prob.* **37**, 835–849.
- [22] PAKES, A. G. (1995). Quasi-stationary laws for Markov processes: examples of an always proximate absorbing state. *Adv. Appl. Prob.* **27**, 120–145.
- [23] 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.
- [24] VAN DOORN, E. A. (1991). Quasi-stationary distributions and convergence to quasi-stationarity of birth–death processes. *Adv. Appl. Prob.* **23**, 683–700.
- [25] VAN DOORN, E. A. (2002). Representations for the rate of convergence of birth–death processes. *Theory Prob. Math. Statist.* **65**, 37–43.
- [26] VAN DOORN, E. A. AND ZEIFMAN, A. I. (2005). Birth–death processes with killing. To appear in *Statist. Prob. Lett.*
- [27] ZEIFMAN, A. I. (1991). Some estimates of the rate of convergence for birth and death processes. *J. Appl. Prob.* **28**, 268–277.
- [28] ZEIFMAN, A. I. (1995). On the estimation of probabilities for birth and death processes. *J. Appl. Prob.* **32**, 623–634.
- [29] ZEIFMAN, A. I. (1995). Upper and lower bounds on the rate of convergence for nonhomogeneous birth and death processes. *Stoch. Process. Appl.* **59**, 157–173.