Intertwining of birth-and-death processes

Jan M. Swart

April 30, 2010

Abstract

It has been known for a long time that for birth-and-death processes started in zero the first passage time of a given level is distributed as a sum of independent exponentially distributed random variables, the parameters of which are the negatives of the eigenvalues of the stopped process. Recently, Diaconis and Miclo have given a probabilistic proof of this fact by constructing a coupling between a general birth-and-death process and a process whose birth rates are the negatives of the eigenvalues ordered from high to low and whose death rates are zero in such a way that the latter process is always ahead of the former, and both arrive at the same time at the given level. In this note, we extend their methods by constructing a third process, whose birth rates are the negatives of the eigenvalues ordered from low to high and whose death rates are zero, which always lags behind the original process and also arrives at the same time.

MSC 2000. Primary: 60J27; Secondary: 15A18, 37A30, 60G40, 60J35, 60J80
Keywords. Intertwining of Markov processes; birth and death process; averaged Markov process; first passage time; coupling; eigenvalues.

Contents

1 Introduction

1.1 First passage times of birth-and-death processes

Let $X = (X_t)_{t\geq 0}$ be the continuous-time Markov process in $\mathbb{N} = \{0, 1, \ldots\}$, started from $X_0 = 0$, that jumps from $x-1$ to x with *birth rate* $b_x > 0$ and from x to $x-1$ with *death rate* $d_x > 0$ ($x \geq 1$). Let

$$\tau_N := \inf\{t \geq 0 : B_t = N\} \qquad (N \geq 1) \tag{1.1}$$

denote the first passage time of N. The following result has been known at least since [KM59, Prop. 1].

Proposition 1.1 (Law of first passage times) *The first passage time τ_N is distributed as a sum of independent exponentially distributed random variables whose parameters $\lambda_1 < \cdots < \lambda_N$ are the negatives of the nonzero eigenvalues of the generator of the process stopped in N.*

Older proofs of this fact are based on a calculation of the Laplace transform of τ_N by purely algebraic methods, see [DM09] for a historical overview. In the latter paper, Diaconis and Miclo gave for the first time a probabilistic proof of Proposition 1.1, by coupling the process X to another birth-and-death process X^+ with birth rates $b_1^+ = \lambda_N, \ldots, b_N^+ = \lambda_1$ and zero death rates, in such a way that $X_{t\wedge\tau_N} \leq X_t^+$ for all $t \geq 0$ and X and X^+ arrive in N at the same time. In the present paper, we will extend their methods by showing that X and X^+ can in addition be coupled to a process X^- with birth rates $b_1^+ = \lambda_1, \ldots, b_N^+ = \lambda_N$ and zero death rates, in such a way that $X_t^- \leq X_{t\wedge\tau_N} \leq X_t^+$ for all $t \geq 0$ and all three processes arrive in N at the same time.

1.2 Intertwining of Markov processes

The coupling technique used by Diaconis and Miclo in [DM09] is of a special kind, which is sometimes called *intertwining of Markov processes*. Let X and X' be continuous-time Markov processes with finite state spaces S and S' and generators G and G', respectively, and let K be a probability kernel from S to S'. Then K defines a linear operator from $\mathbb{R}^{S'}$ to $\mathbb{R}^S$, also denoted by K, by the formula

$$Kf(x) := \sum_{y\in S'} K(x,y)f(y). \tag{1.2}$$

The following result, which is based on an observation by Rogers and Pitman [RP81], was proved by Fill in [Fil92, Thm. 2]. (An independent proof can be found in [AS10, Prop. 4]).

Proposition 1.2 (Intertwining of Markov processes) *Assume that*

$$GK = KG'. \tag{1.3}$$

Then there exists a generator $\hat{G}$ of an $S \times S'$-valued Markov process with the property that if (X, X') evolves according to $\hat{G}$ and satisfies

$$\mathbb{P}[X_0' = y \,|\, X_0] = K(X_0, y) \qquad (y \in S'), \tag{1.4}$$

then

$$\mathbb{P}[X_t' = y \,|\, (X_s)_{0\leq s\leq t}] = K(X_t, y) \qquad (t \geq 0,\ y \in S'), \tag{1.5}$$

and the processes X and X', on their own, are Markov processes evolving according to the generators G and G', respectively.

Algebraic relations of the type (1.3) are called intertwining relations, hence the name *intertwining of Markov processes*. We note that the operator K needs in general not have an inverse, and even if it does, this inverse will in general not be associated to a probability kernel from S' to S. In view of this, an intertwining of Markov processes is not a symmetric relation. To express this, following terminology introduced in [AS10], we will also say that in the set-up of Proposition 1.2, X' is an *averaged Markov process on X*.

1.3 Intertwining of birth-and-death processes

We are now ready to formulate our main result. Deviating slightly from our notation in Section 1.1, we let $X = (X_t)_{t\geq 0}$ be a continuous-time Markov process with state space $\{0, \ldots, N\}$,

started from $X_0 = 0$, that jumps from $x-1$ to x with birth rate b_x and from x to $x-1$ with death rate d_x, where $b_1, \ldots, b_N > 0$, $d_1, \ldots, d_{N-1} > 0$, but $d_N = 0$, i.e., X is the stopped process from Section 1.1. We let G denote the generator of X, i.e.,

$$Gf(x) := b_{x+1}\big(f(x+1) - f(x)\big) + d_x\big(f(x-1) - f(x)\big) \qquad (0 \leq x \leq N), \tag{1.6}$$

where $f : \{0, \ldots, N\} \to \mathbb{R}$ is a real function and we adopt the convention that $d_0 = 0$ and $b_{N+1} = 0$ so that the corresponding terms in (1.6) are zero, regardless of the (fictive) values of f in -1 and $N+1$. The following theorem is our main result.

Theorem 1.3 (Intertwining of birth-and-death processes) *The operator G has $N+1$ distinct eigenvalues $0 = -\lambda_0 > -\lambda_1 > \cdots > -\lambda_N$. Let X^- and X^+ be the pure birth processes in $\{0, \ldots, N\}$, started from $X_0^- = X_0^+ = 0$, with birth rates $b_1^- := \lambda_1, \ldots, b_N^- := \lambda_N$ and $b_1^+ := \lambda_N, \ldots, b_N^+ := \lambda_1$, respectively, and let G^- and G^+ be their generators. Then there exist probability kernels K^- and K^+ on $\{0, \ldots, N\}$ satisfying*

$$\begin{aligned} &K^-(x, \{0, \ldots, x\}) = 1, \qquad &&K^+(x, \{0, \ldots, x\}) = 1, \quad (0 \leq x \leq N) \\ &K^-(N, N) = 0, &&K^+(N, N) = 0, \end{aligned} \tag{1.7}$$

and

$$\text{(i)} \quad K^+G = G^+K^+ \quad \textit{and} \quad \text{(ii)} \quad GK^- = K^-G^-. \tag{1.8}$$

Moreover, the processes X^-, X, and X^+ can be coupled in such a way that

$$\begin{aligned} &\text{(i)} \qquad \mathbb{P}[X_t = y \,|\, (X_s^+)_{0 \leq s \leq t}] = K^+(X_t^+, y) \qquad (t \geq 0,\ 0 \leq y \leq N), \\ &\text{(ii)} \quad \mathbb{P}[X_t^- = y \,|\, (X_s^+, X_s)_{0 \leq s \leq t}] = K^-(X_t, y) \qquad (t \geq 0,\ 0 \leq y \leq N). \end{aligned} \tag{1.9}$$

The existence of a kernel K^+ such that (1.8) (i) and (1.9) (i) hold has been proved before in [DM09, Prop. 10]. Our new contribution is the construction of the kernel K^- such that moreover (1.8) (ii) and (1.9) (ii) hold. It is easy to see that formulas (1.7) and (1.9) imply that

$$\begin{aligned} &\text{(i)} \quad X_t^- \leq X_t \leq X_t^+ \qquad (t \geq 0), \\ &\text{(ii)} \quad \tau_N^- = \tau_N = \tau_N^+, \end{aligned} \tag{1.10}$$

where $\tau_N := \inf\{t \geq 0 : X_t = N\}$ and τ_N^- and τ_N^+ are defined similarly for X^- and X^+, respectively. We note that X^- and X^+ move, in a sense, in the slowest resp. fastest possible way from 0 to N, given that they have to arrive at exactly the same time as X. Note that, using terminology introduced at the end of Section 1.2, X is an averaged Markov process on X^+ and X^- is an averaged Markov process on X.

1.4 Discussion

In comparison to the paper by Diaconis and Miclo [DM09], the present paper does not add too much that is new. In particular the construction of the kernel K^- in Theorem 1.3 is very similar to the construction of the kernel K^+, which was already carried out in [DM09]. However, we believe that the observation that both constructions are possible, with an interesting symmetry between them, is of some interest.

The methods of [DM09] can also be used to study birth-and-death processes on $\{0, \ldots, N\}$ whose death rate d_N is not zero and which, therefore, converge in law to a unique equilibrium.

In particular, Diaconis and Miclo use their intertwining relation (1.8) (i) to contruct a fastest strong stationary time for such processes (we refer to [DM09] for the definition). In contrast, it seems that the interwining relation (1.8) (ii) does not generalize to such a setting.

On a more general level, one may ask what the advantage is of a 'probabilistic' proof of Proposition 1.1 as opposed to older, more algebraic proofs. Since most of the work behind Theorem 1.3 goes into proving the intertwining relations (1.8), one might even argue that the present proof is still rather algebraic in nature, although with a strong probabilistic flavour. In this context, it is interesting to note that the fact that G is diagonalizable with real, distinct eigenvalues follows as a result of our proofs (in particular, this follows from a repeated application of the Perron-Frobenius theorem) and does not have to be provided by some extra argument (based on, for example, reversibility).

In general, diagonalizing a generator of a Markov process gives very strong information about the process, but in practice, if the state space is large, it is hard to get good information about the position of eigenvalues etc. The idea of interwining generators with transition kernels may in some cases be a good way to transform generators of complicated processes into generators of more simple processes and thus provide a more probabilistic alternative to diagonalization.

The methods of this paper can certainly be extended to one-dimensional processes with two traps, to dicrete-time processes, and to one-dimensional diffusions. Miclo [Mic10] has proved a generalization of Proposition 1.1 for reversible Markov chains. In [AS10], intertwining relations were used to estimate the time to extinction for large hierarchical contact processes. The present work was partly motivated by an open problem from that paper. (To be precise, Question 1° from Section 3.3.)

2 Proofs

2.1 Leading eigenvectors

Let X be the birth-and-death process in $S_N := \{0, \dots, N\}$ from Section 1.3 and let $G : \mathbb{R}^{S_N} \to \mathbb{R}^{S_N}$ be its generator, defined in (1.6). We equip $\mathbb{R}^{S_N}$ with the usual inner product $\langle \pi | f \rangle := \sum_{x=0}^{N} \pi(x) f(x)$ and let $G^\dagger$ denote the adjoint of G with respect to this inner product. Then

$$G^\dagger \pi(x) = b_x \pi(x-1) - b_{x+1} \pi(x) + d_{x+1} \pi(x+1) - d_x \pi(x), \tag{2.1}$$

where as in (1.6) we use the convention that $d_0 = 0$ and $b_{N+1} = 0$ so that the corresponding terms in (2.1) are zero, regardless of the (fictive) values of π in -1 and $N+1$.

Since δ_N (the delta mass in N) is the unique invariant law of X, the eigenvalue 0 of the generator G has multiplicity one and its unique left and right eigenvectors are δ_N and the constant function 1, respectively. We will need the following result on the next largest eigenvalue and its left and right eigenvectors.

Lemma 2.1 (Leading eigenvectors) *There exists a $\lambda > 0$ and $f, \pi \in \mathbb{R}^{S_N}$ such that*

(i) *f is strictly decreasing on $\{0, \dots, N\}$ and satisfies $f(0) = 1$, $f(N) = 0$*

(ii) *π is strictly positive on $\{0, \dots, N-1\}$ and satisfies $\sum_{x=0}^{N-1} \pi(x) = 1 = -\pi(N)$.* (2.2)

(ii) *$Gf = -\lambda f$ and $G^\dagger \pi = -\lambda \pi$.*

Proof Set

$$\begin{array}{lll} e(x) := \delta_x & (0 \le x \le N-1) & \text{and} \quad e(N) := 1, \\ \xi(x) := \delta_x - \delta_N & (0 \le x \le N-1) & \text{and} \quad \xi(N) := \delta_N. \end{array} \tag{2.3}$$

Then $\{e(0), \dots, e(N)\}$ is a basis for $\mathbb{R}^{S_N}$ and $\{\xi(0), \dots, \xi(N)\}$ is its associated dual basis, i.e., $\langle e(x)|\xi(y)\rangle = 1_{\{x=y\}}$. Set

$$\begin{aligned} \mathcal{E} &:= \text{span}\{e(0), \dots, e(N-1)\} = \{f \in \mathbb{R}^{S_N} : f(N) = 0\}, \\ \mathcal{F} &:= \text{span}\{\xi(0), \dots, \xi(N-1)\} = \{\pi \in \mathbb{R}^{S_N} : \textstyle\sum_{x=0}^{N} \pi(x) = 0\}. \end{aligned} \tag{2.4}$$

Since N is a trap for the process X, it is easy to see that the operator G maps the space $\mathcal{E}$ and into itself. It follows that with respect to the basis $\{e(0), \dots, e(N)\}$, the matrix $[G]$ of G has the form

$$[G] = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \tag{2.5}$$

where A is the matrix of the restriction of G to $\mathcal{E}$. The restriction of the process X to the space $\{0, \dots, N-1\}$ is irreducible in the sense that there is a positive probability of going from any state to any other state. Therefore, a standard application of the Perron-Frobenius theorem shows that A has a real eigenvalue $-\lambda$ of multiplicity one, which is larger in absolute value than all other eigenvalues, and associated left and right eigenvectors $\pi \in \mathcal{F}$ and $f \in \mathcal{E}$ that are strictly positive with respect to the bases $\{\xi(0), \dots, \xi(N-1)\}$ and $\{e(0), \dots, e(N-1)\}$, respectively. Since Markov semigroups are contractive we have $-\lambda \le 0$ and since the eigenvalue zero of G has multiplicity one we conclude that $-\lambda < 0$. Since we can always normalize our eigenvectors such that $\sum_{x=0}^{N-1} \pi(x) = 1$ and $\max_{x=0}^{N-1} f(x) = 1$, this proves all statements of the lemma except for the fact that f is strictly decreasing.

To prove this latter fact, we observe that by the facts that $Gf = -\lambda f$ and $f > 0$ on $\{0, \dots, N-1\}$,

$$b_1\big(f(1) - f(0)\big) = -\lambda f(0) < 0, \tag{2.6}$$

which show that $f(0) > f(1)$. By the same argument,

$$b_{x+1}\big(f(x+1) - f(x)\big) = -\lambda f(x) - d_x\big(f(x-1) - f(x)\big) < 0 \qquad (1 \le x \le N-1), \tag{2.7}$$

from which we see by induction that $f(x) > f(x+1)$ for all $0 \le x \le N-1$. ∎

2.2 Intertwining the fast process

In this section, we prove the existence of a kernel K^+ satisfying (1.7) and (1.8). Our proof is basically the same as the proof given in [DM09], but as a preparation for the next section it will be convenient to review their proof and shorten it somewhat. The proof in [DM09] is written in such a way as to make clear how the authors arrived at their argument and uses discrete derivatives that are presumably also useful if one wants to generalize the theory to one-dimensional diffusions. If our only aim is Theorem 1.3, however, we can summarize their arguments quite a bit.

The kernel K^+ will be constructed as the concatenation of an inductively defined sequence of kernels $K^{(N-1)+}, \dots, K^{(1)+}$. Associated with these kernels is a sequence of generators $G^{(N-1),+}, \dots, G^{(0)+}$ of birth-and-death processes in $\{0, \dots, N\}$ satisfying the intertwining relations

$$K^{(M)+}G^{(M)+} = G^{(M-1)+}K^{(M)+} \qquad (1 \le M \le N-1), \tag{2.8}$$

where the process with generator $G^{(M)}$ has birth rates $b_1^{(M)}, \dots, b_N^{(M)} > 0$ and death rates $d_1^{(M)}, \dots, d_M^{(M)} > 0$, $d_{M+1}^{(M)} = \cdots = d_N^{(M)} = 0$; see Figure 1 for a picture. In particular, we will choose $G^{(N-1)+} := G$ and setting $G^+ := G^{(0)+}$ will yield the desired pure birth process with birth rates $b_1^+ = \lambda_N, \dots, b_N^+ = \lambda_1$.

The core the proof is the following proposition, which corresponds to the inductive step in the argument.

Proposition 2.2 (Inductive step) *Let $1 \le M \le N-1$ and let G be the generator of a birth-and-death process in $\{0, \dots, N\}$ with birth rates $b_1, \dots, b_N > 0$ and death rates $d_1, \dots, d_M > 0$, $d_{M+1} = \cdots = d_N = 0$. Then there exists a probability kernel K on $\{0, \dots, N\}$ satisfying*

$$K(x, \{x, \dots, x\}) = 1 \quad (0 \le x \le N) \quad \text{and} \quad K(x, x) = 1 \quad (M+1 \le x \le N), \tag{2.9}$$

and a generator G' of a birth-and-death process in $\{0, \dots, N\}$ with birth rates $b'_1, \dots, b'_N > 0$ and death rates $d'_1, \dots, d'_{M-1} > 0$, $d'_M = \cdots = d'_N = 0$, such that $KG = G'K$.

Proof It follows from Lemma 2.1 applied to the process stopped at $M+1$ that there exists a function $\rho : \{0, \dots, N\} \to \mathbb{R}$ such that $\rho > 0$ on $\{0, \dots, M\}$, $\rho = 0$ on $\{M+1, \dots, N\}$, $\sum_{x=0}^N \rho(x) = 1$, and

$$G^\dagger \rho(x) = -\lambda \rho(x) + \lambda \delta_{M+1}(x) \qquad (0 \le x \le N), \tag{2.10}$$

where

$$\lambda = b_{M+1} \rho(M) > 0. \tag{2.11}$$

The law ρ is sometimes called a *quasi-stationary law*. Using ρ, we define the kernel K on $\{0, \dots, N\}$ by

$$K(x, y) := \begin{cases} 1_{\{y \le x\}} \dfrac{\rho(y)}{H(x)} & \text{if } x \le M, \\ 1_{\{y = x\}} & \text{if } M+1 \le x, \end{cases} \tag{2.12}$$

where

$$H(x) := \sum_{y=0}^{x} \rho(y) \qquad (0 \le x \le M). \tag{2.13}$$

Since K is a lower triangular matrix, it is invertible, so there exists a unique linear operator G' satisfying $KG = G'K$ and G' is in fact given by $G' = KGK^{-1}$. Since $G'1 = G'K1 = KG1 = 0$ we see that

$$G'(x, x) = -\sum_{y \ne x} G'(x, y). \tag{2.14}$$

In view of this, to prove our claim, it suffices to check that the off-diagonal entries of G' coincide with those of a birth-and-death process in $\{0, \dots, N\}$ with birth rates $b'_1, \dots, b'_N > 0$ and death rates $d'_1, \dots, d'_{M-1} > 0$, $d'_M = \cdots = d'_N = 0$.

To determine the off-diagonal entries of G', we calculate, using (2.1) and (2.10),

$$\begin{aligned} &(KG)(x, y) = G^\dagger K(x, \cdot)(y) \\ &= \begin{cases} -\lambda 1_{\{y \le x\}} \dfrac{\rho(y)}{H(x)} - d_{x+1} \dfrac{\rho(x+1)}{H(x)} \delta_x(y) + b_{x+1} \dfrac{\rho(x)}{H(x)} \delta_{x+1}(y) & \text{if } x < M, \\ -\lambda \rho(y) + \lambda \delta_{M+1}(y) & \text{if } x = M, \\ -b_{x+1} \delta_x(y) + b_{x+1} \delta_{x+1}(y) & \text{if } x > M. \end{cases} \end{aligned} \tag{2.15}$$

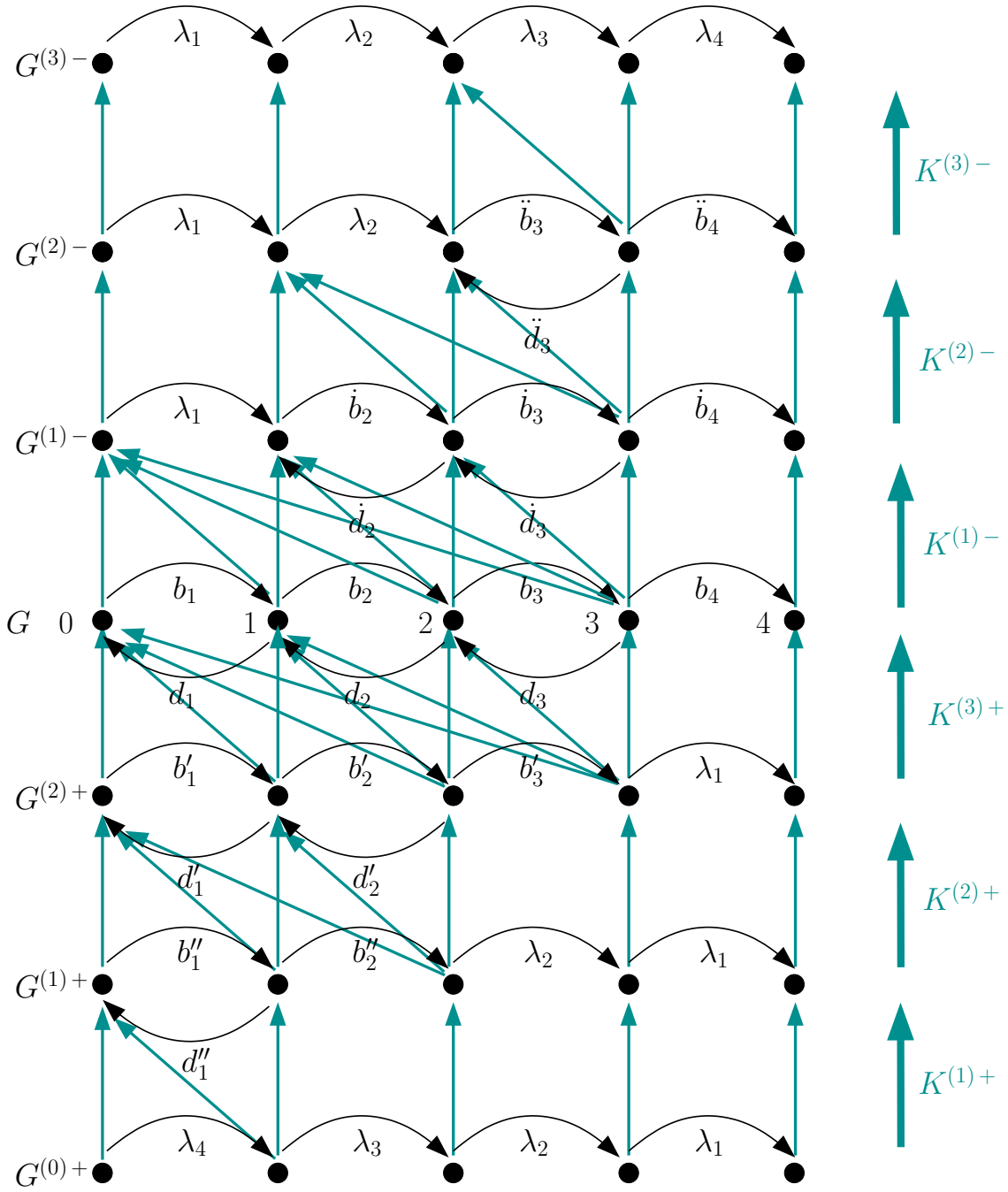

Figure 1: Intertwining of birth and death processes. In this picture $N = 4$. All nonzero transition rates and probabilities have been indicated with arrows.

In order to find G', we need to express these formulas, as functions of y, as linear combinations of the basis vectors $(K(x,\,\cdot\,))_{0\le x\le N}$. To that aim, we observe that

$$\delta_x = K(x,\,\cdot\,) \qquad (M+1 \le x \le N), \tag{2.16}$$

while for $1 \le x \le M$, we have

$$\begin{aligned}\delta_x(y) &= \big(1_{\{y\le x\}} - 1_{\{y\le x-1\}}\big)\frac{\rho(y)}{\rho(x)} \\ &= \frac{H(x)}{\rho(x)}1_{\{y\le x\}}\frac{\rho(y)}{H(x)} - \frac{H(x-1)}{\rho(x)}1_{\{y\le x-1\}}\frac{\rho(y)}{H(x-1)} \\ &= \frac{H(x)}{\rho(x)}K(x,y) - \frac{H(x-1)}{\rho(x)}K(x-1,y).\end{aligned} \tag{2.17}$$

Inserting this into (2.15), we find that

$$\begin{aligned}&\sum_{x'=0}^{N} G'(x,x')K(x',y) := (KG)(x,y) \\ &= \begin{cases} -\lambda K(0,y) - d_1\dfrac{\rho(1)}{H(0)}K(0,y) & \\ \quad +b_1\dfrac{\rho(0)}{H(0)}\Big(\dfrac{H(1)}{\rho(1)}K(1,y) - \dfrac{H(0)}{\rho(1)}K(0,y)\Big) & \text{if } x=0, \\ -\lambda K(x,y) - d_{x+1}\dfrac{\rho(x+1)}{H(x)}\Big(\dfrac{H(x)}{\rho(x)}K(x,y) - \dfrac{H(x-1)}{\rho(x)}K(x-1,y)\Big) & \\ \quad +b_{x+1}\dfrac{\rho(x)}{H(x)}\Big(\dfrac{H(x+1)}{\rho(x+1)}K(x+1,y) - \dfrac{H(x)}{\rho(x+1)}K(x,y)\Big) & \text{if } 0<x<M, \\ -\lambda K(M,y) + \lambda K(M+1,y) & \text{if } x=M, \\ -b_{x+1}K(x,y) + b_{x+1}K(x+1,y) & \text{if } x>M. \end{cases}\end{aligned} \tag{2.18}$$

From this, we can read off the off-diagonal entries of G'. Indeed,

$$\begin{aligned} b'_{x+1} = G'(x,x+1) &= \begin{cases} b_{x+1}\dfrac{\rho(x)H(x+1)}{H(x)\rho(x+1)} & \text{if } x<M, \\ \lambda & \text{if } x=M, \\ b_{x+1} & \text{if } x>M, \end{cases} \\ d'_x = G'(x,x-1) &= \begin{cases} d_{x+1}\dfrac{\rho(x+1)H(x-1)}{H(x)\rho(x)} & \text{if } 0<x<M, \\ 0 & \text{if } x\ge M, \end{cases}\end{aligned} \tag{2.19}$$

and all other off-diagonal entries are zero. ∎

Remark The proof of Proposition 2.2 is straightforward except for the clever choice of K in (2.12)–(2.13). For some motivation of this choice and the way the authors arrived at it we refer to [DM09].

Using Proposition 2.2 we can construct a sequence of kernels $K^{(N-1)+},\dots,K^{(1)+}$ and generators $G^{(N-1),+},\dots,G^{(0)+}$ satisfying the intertwining relations (2.8), such that $G^+ := G^{(0)+}$ is a pure birth process with birth rates $b_1^+,\dots,b_N^+ > 0$, say. It is now easy to see that the composed kernel

$$K^+ := K^{(1)+}\cdots K^{(N-1)+} \tag{2.20}$$

satisfies $K^+(x,\{0,\dots,x\})=1$ $(0\le x\le N)$, $K^+(N,N)=1$ and $K^+G=G^+K^+$. It is straightforward to check that the eigenvalues of G' are $-b_1^+,\dots,-b_N^+,0$. Since $G=(K^+)^{-1}G^+K^+$, the operators G and G^+ have the same spectrum.

We claim that $b_1^+>\cdots>b_N^+>0$. To see this, recall from the proofs of Lemma 2.1 and Proposition 2.2 that $-b_M^+$ is the Perron-Frobenius eigenvalue of the process with generator $G^{(M)+}$ stopped at $M+1$. It follows from the intertwining relation (2.8) that $-b_{M-1}^+$ is also an eigenvalue of this process, corresponding to a different eigenvector, hence by the Perron-Frobenius theorem, $b_{M-1}>b_M$.

2.3 Intertwining the slow process

In the previous section, we have constructed a kernel K^+ and generator of a pure birth process G^+ such that (1.8) (i) holds. In this section, we construct a kernel K^- and generator of a pure birth process G^- satisfying (1.8) (ii). The proof will be very similar to the previous case, except that some things will 'go he other way around'. In particular, using terminology introduced at the end of Section 1.2, G^- will be the generator of an avaraged Markov process X^- on X while in the previous section we constructed a pure birth process X^+ such that X is an averaged Markov process on X^+.

As in the previous section, the kernel K^- will be constructed as the concatention of an inductively defined sequence of kernels $K^{(1)-},\dots,K^{(N-1)-}$. Associated with these kernels is a sequence of generators $G^{(1),-},\dots,G^{(N-1)-}$ of birth-and-death processes in $\{0,\dots,N\}$ satisfying the intertwining relations

$$G^{(M-1)-}K^{(M)-}=K^{(M)-}G^{(M)-}\qquad(1\le M\le N-1),\tag{2.21}$$

where the process with generator $G^{(M)}$ has birth rates $b_1^{(M)},\dots,b_N^{(M)}>0$ and death rates $d_1^{(M)}=\cdots=d_M^{(M)}=0$, $d_{M+1}^{(M)},\dots,d_{N-1}^{(M)}>0$, and $d_N^{(M)}=0$. We again refer to Figure 1 for an illustration.

The core of the argument is the following proposition.

Proposition 2.3 (Inductive step) *Let $0\le M\le N-2$ and let G be the generator of a birth-and-death process in $\{0,\dots,N\}$ with birth rates $b_1,\dots,b_N>0$ and death rates $d_1=\cdots=d_M=0$, $d_{M+1},\dots,d_{N-1}>0$, and $d_N=0$. Then there exists a probability kernel K on $\{0,\dots,N\}$ satisfying*

$$K(x,\{x,\dots,x\})=1\quad(0\le x\le N)\quad\text{and}\quad K(x,x)=1\quad(x\notin\{M,\dots,N-1\}),\tag{2.22}$$

and a generator $\dot G$ of a birth-and-death process in $\{0,\dots,N\}$ with birth rates $\dot b_1,\dots,\dot b_N>0$ and death rates $\dot d_1=\cdots=\dot d_{M+1}=0$, $\dot d_{M+2},\dots,\dot d_{N-1}>0$, and $\dot d_N=0$, such that $GK=K\dot G$.

Proof It follows from Lemma 2.1 applied to the process restricted to $\{M,\dots,N\}$ that there exists a function $f:\{0,\dots,N\}\to\mathbb{R}$ such that $f=0$ on $\{0,\dots,M-1\}$, f is strictly decreasing on $\{M,\dots,N\}$, $f(M)=1$, $f(N)=0$, and

$$Gf(x)=-\lambda f(x)+b_M\delta_{M-1}(x)\qquad(0\le x\le N),\tag{2.23}$$

where

$$\lambda=b_{M+1}\big(1-f(M+1)\big)>0.\tag{2.24}$$

We set

$$K(x,y) := 1_{\{x=y\}} \qquad (y \notin \{M,\dots,N-1\}). \tag{2.25}$$

For $y = M,\dots,N-1$, we claim that we can inductively define the kernel $K(x,y)$ and contants $C_y > 0$ in such a way that

$$\left.\begin{aligned} &\text{(i)} \quad K(x,y) := C_y 1_{\{y\le x\}} f(x), \\ &\text{(ii)} \quad \sum_{y'=M}^{y} K(y,y') = 1, \end{aligned}\right\} \quad (M \le y \le N-1). \tag{2.26}$$

To see that this is all right, note that for $y = M$ (2.26) (i) and (ii) are satisfied by choosing $C_M := 1$, while for $M+1 \le y \le N-1$ (2.26) (i) and (ii) imply that we must choose

$$C_y := \frac{1}{f(y)}\Big(1 - \sum_{y'=M}^{y-1} K(y,y')\Big). \tag{2.27}$$

Since f is strictly decreasing on $\{M,\dots,N\}$, one has, by induction,

$$\sum_{y'=M}^{y-1} K(y,y') = \sum_{y'=M}^{y-1} C_{y'} f(y) < \sum_{y'=M}^{y-1} C_{y'} f(y-1) = \sum_{y'=M}^{y-1} K(y-1,y') = 1, \tag{2.28}$$

which shows that $C_y > 0$. We now calculate

$$\begin{aligned} (GK)(x,y) &= GK(\,\cdot\,,y)(x) \\ &= \begin{cases} b_y\delta_{y-1}(x) - b_{y+1}\delta_y(x) & \text{if } 0 \le y \le M-1, \\ -\lambda f(x) + b_M\delta_{M-1}(x) & \text{if } y = M, \\ -\lambda C_y 1_{\{y\le x\}} f(x) + b_y C_y \delta_{y-1}(x) & \\ \quad -d_y C_y\big(f(y-1) - f(y)\big)\delta_y(x) & \text{if } M+1 \le y \le N-1, \\ b_N\delta_{N-1}(x) & \text{if } y = N. \end{cases} \end{aligned} \tag{2.29}$$

By the same arguments as those in the previous section, there exists a unique linear operator $\dot{G}$ such that $GK = K\dot{G}$. In order to check that $\dot{G}$ is the generator of a birth-and-death process in $\{0,\dots,N\}$ with birth rates $\dot{b}_1,\dots,\dot{b}_N > 0$ and death rates $\dot{d}_1 = \cdots = \dot{d}_{M+1} = 0$, $\dot{d}_{M+2},\dots,\dot{d}_{N-1} > 0$, and $\dot{d}_N = 0$, it suffices to check that the off-diagonal entries $\dot{G}(x,y)$ have the desired form. In order to do this, we must express the formulas in (2.29), as functions of x, as linear combinations of the basis vectors $(K(\,\cdot\,,y))_{0\le y\le N}$. We observe that

$$\delta_y(x) = K(\,\cdot\,,y) \qquad \big(y \notin \{M,\dots,N-1\}\big), \tag{2.30}$$

while for $M \le y \le N-2$, we have

$$\begin{aligned} \delta_y(x) &= \big(1_{\{y\le x\}} - 1_{\{y+1\le x\}}\big)\frac{f(x)}{f(y)} \\ &= \frac{1}{f(y)C_y} C_y 1_{\{y\le x\}} f(x) - \frac{1}{f(y)C_{y+1}} C_{y+1} 1_{\{y+1\le x\}} f(x) \\ &= \frac{1}{f(y)C_y} K(x,y) - \frac{1}{f(y)C_{y+1}} K(x,y+1), \end{aligned} \tag{2.31}$$

and

$$\delta_{N-1}(x) = \frac{1}{f(N-1)C_{N-1}} K(x, N-1). \tag{2.32}$$

Inserting this into (2.29), we obtain

$$\begin{aligned}
&\sum_{y'} K(x,y')\dot{G}(y',y) = (GK)(x,y) \\
&= \begin{cases}
b_y K(x,y-1) - b_{y+1}K(x,y) & \text{if } 0 \le y \le M-1, \\
-\lambda K(x,M) + b_M K(x,M-1) & \text{if } y = M, \\
-\lambda K(x,y) + b_y C_y \Big(\dfrac{K(x,y-1)}{f(y-1)C_{y-1}} - \dfrac{K(x,y)}{f(y+1)C_y} \Big) & \\
\quad -d_y C_y \big(f(y-1) - f(y) \big) \Big(\dfrac{K(x,y)}{f(y)C_y} - \dfrac{K(x,y+1)}{f(y)C_{y+1}} \Big) & \text{if } M+1 \le y \le N-2, \\
-\lambda K(x,y) + b_y C_y \Big(\dfrac{K(x,y-1)}{f(y-1)C_{y-1}} - \dfrac{K(x,y)}{f(y+1)C_y} \Big) & \\
\quad -d_y C_y \big(f(y-1) - f(y) \big) \dfrac{K(x,y)}{f(y)C_y} & \text{if } y = N-1, \\
b_N \dfrac{K(x,N-1)}{f(N-1)C_{N-1}} & \text{if } y = N,
\end{cases}
\end{aligned} \tag{2.33}$$

where we use the convention that $b_0 = 0$ and hence $b_0 K(x,-1) = 0$, regardless of the (fictive) value of $K(x,-1)$. From (2.33) we can read off the off-diagonal entries of $\dot{G}$. Indeed,

$$\begin{aligned}
\dot{b}_y = \dot{G}(y-1,y) &= \begin{cases} b_y & \text{if } 1 \le y \le M, \\ b_y \dfrac{C_y}{f(y-1)C_{y-1}} & \text{if } M+1 \le y \le N, \end{cases} \\
\dot{d}_{y+1} = \dot{G}(y+1,y) &= \begin{cases} 0 & \text{if } y \notin \{M+1, \dots, N-2\}, \\ d_y \dfrac{C_y(f(y-1) - f(y))}{C_{y+1} f(y)} & \text{if } M+1 \le y \le N-2, \end{cases}
\end{aligned} \tag{2.34}$$

and all other off-diagonal entries are zero. We note that in particular, by (2.24) and the definition of the C_y's,

$$\dot{b}_{M+1} = \frac{b_{M+1}C_{M+1}}{f(M)C_M} = b_{M+1}C_{M+1} = b_{M+1}\big(1 - f(M+1)\big) = \lambda. \tag{2.35}$$

∎

Remark As in the case of Proposition 2.2, the proof of Proposition 2.3 is straightforward except for the choice of the kernel K. We have guessed formula (2.26) by analogy with formula (2.12), which is due to [DM09].

With the help of Proposition 2.3, we can inductively define kernels $K^{(1)-}, \dots, K^{(N-1)-}$ and operators $G^{(1),-}, \dots, G^{(N-1)-}$. Setting $G^- := G^{(N-1)-}$ and

$$K^- = K^{(1)-} \cdots K^{(N-1)-} \tag{2.36}$$

now yields a generator of a pure birth process with birth rates $b_1^-, \dots, b_N^-$ and a kernel K^- with the properties described in (1.7)–(1.8).

In the same way as in the previous section, we see that $0, -b_1^-, \ldots, -b_N^-$ are the eigenvalues of G. To see that $0 < b_1 < \cdots < b_N$ we observe from (2.35) that $-b_M^-$ is the Perron-Frobenius eigenvalue of the process with generator $G^{(M)-}$ restricted to $\{M, \ldots, N\}$. It follows from the intertwining relation (2.21) that $-b_{M+1}^+$ is also an eigenvalue of this process, corresponding to a different eigenvector, hence by the Perron-Frobenius theorem, $b_M < b_{M+1}$.

2.4 Proof of the main theorem

Proof of Theorem 1.3 The existence of generators G^-, G^+ and kernels K^-, K^+ satisfying (1.7)–(1.8) has been proved in the previous sections. By Proposition 1.2, it follows that X^+ and X can be coupled such that (1.9) (i) holds. By applying Proposition 1.2 to the kernel L from $\{0, \ldots, N\}^2$ to $\{0, \ldots, N\}$ given by

$$L\big((x,y),z\big) := K^+(x,y)K^-(y,z) \qquad (0 \leq x, y, z \leq N), \tag{2.37}$$

we see that X^+, X, and X^- can be coupled in such a way that both (1.9) (i) and (ii) hold. ∎

References

[AS10] S.R. Athreya and J.M. Swart. Survival of contact processes on the hierarchical group. *Probab. Theory Relat. Fields* 147(3), 529–563, 2010.

[DM09] P. Diaconis and L. Miclo. On times to quasi-stationarity for birth and death processes. *J. Theor. Probab.* 22, 558–586, 2009.

[Fil92] J.A. Fill. Strong stationary duality for continuous-time Markov chains. I. Theory. *J. Theor. Probab.* 5(1), 45–70, 1992.

[KM59] S. Karlin and J. McGregor. Coincidence properties of birth and death processes. *Pac. J. Math.* 9, 1109–1140, 1959.

[Mic10] L. Miclo. On absorbing times and Dirichlet eigenvalues. Preprint, 2010.

[RP81] L.C.G. Rogers and J.W. Pitman. Markov functions. *Ann. Probab.* 9(4), 573–582, 1981.