ESAIM: PS 24 (2020) 100–112
https://doi.org/10.1051/ps/2020004

ESAIM: Probability and Statistics
www.esaim-ps.org

CONTINUOUS-TIME MARKOV PROCESSES, ORTHOGONAL POLYNOMIALS AND LANCASTER PROBABILITIES

Ramsés H. Mena and Freddy Palma*

Abstract. This work links the conditional probability structure of Lancaster probabilities to a construction of reversible continuous-time Markov processes. Such a task is achieved by using the spectral expansion of the corresponding transition probabilities in order to introduce a continuous time dependence in the orthogonal representation inherent to Lancaster probabilities. This relationship provides a novel methodology to build continuous-time Markov processes *via* Lancaster probabilities. Particular cases of well-known models are seen to fall within this approach. As a byproduct, it also unveils new identities associated to well known orthogonal polynomials.

Mathematics Subject Classification. 60J25, 60J35, 33C45.

Received December 14, 2018. Accepted January 15, 2020.

1. Introduction

The theory of continuous-time Markov processes represents one of the main tools for modeling random phenomena. For instance, diffusion processes have been widely used to model the stock price of derivatives, the evolution of genes, particles in physics (*cf.* [19]). Nonetheless, the procedure for modeling these phenomena is commonly done through a stochastic differential equation (SDE). This implies that, in general, we do not have expressions for transition probabilities driving these kind of processes (*cf.* [7]). An interesting way to deal with this issue is given by using orthogonal polynomials, *i.e.* a spectral expansion for the transition probabilities. Indeed, some of these have been widely used to propose inference and simulation methods, see for example *cf.* [9, 10] where the case of the Wright-Fisher diffusion is treated.

In a related direction, Mena and Walker [22] presented a method to devise tractable expressions for the transition probabilities driving some continuous-time reversible Markov processes. Such a methodology requires the specification of a bivariate distribution with known marginal distributions. Having this in mind, an alternative natural direction is to use Lancaster probabilities, namely bivariate distributions satisfying a bi-orthogonal property which involves orthogonal polynomials (*cf.* [14–17]).

Our proposal is motivated by the work of Diaconis *et al.* [4] and Letac [18], which used the aforementioned idea to build one-step ahed transition probabilities of some discrete-time Markov models, by generalizing it to the continuous-time framework. We find a connection with the spectral expansion of transition probabilities associated to continuous time Markov processes. To be precise, we consider time-dependent Lancaster sequences,

Keywords and phrases: Continuous-time reversible Markov process, Lancaster probabilities, orthogonal polynomials, spectral expansion.

Department of Probability and Statistics, IIMAS, UNAM, Mexico city, Mexico.

* Corresponding author: fred_25@ciencias.unam.mx

Article published by EDP Sciences

and thus explore the relationship between the construction of [22] and the theory of Lancaster probabilities. As a result, one can use Lancaster probabilities to build continuous-time transition probabilities, and viceversa. This reveals, simple expressions related to orthogonal polynomials as well as novel bi-orthogonal properties.

As a consequence of such a connection, we derive expressions for the Lancaster probabilities associated to some classical continuous-time Markov models, *e.g.* models with gamma, negative-binomial, Poisson, Normal, beta stationary distributions. In particular the Poisson and negative-binomial cases, which correspond to a $M/M/\infty$ queue and the birth, death and immigration process, allow to find new identities associated to widely know orthogonal polynomials. Moreover, we prove the bi-orthogonal property between the Jacobi and the Hahn polynomials, which unveils Lancaster sequences associated to beta and beta-binomial marginals, and thus lead to the Wright-Fisher diffusion.

The rest of the paper is organized as follows: Section 2 summarizes the construction of continuous-time reversible Markov processes introduced in Mena and Walker [22]. In Section 3, we derive a methodology to build time-dependent Lancaster probabilities, and its relation with the spectral expansion of some transition probabilities. Finally, to illustrate our results, the last sections present some widely known models, for which we construct time-dependent Lancaster probabilities.

2. On a construction of continuous-time reversible Markov processes

Consider the measurable spaces $(\mathbb{X}, \mathscr{X})$ and $(\mathbb{Y}, \mathscr{Y})$, where $\mathbb{X}, \mathbb{Y} \subseteq \mathbb{R}$. Let π be a probability distribution over $(\mathbb{X}, \mathscr{X})$; $\Theta : \mathbb{R}^+ \to \mathbb{R}^+$ be a real-valued function; and for every $x \in \mathbb{X}$, $\mathtt{f}\,(\cdot; \Theta_t, x)$ be a probability distribution over $(\mathbb{Y}, \mathscr{Y})$. Here, we are assuming that the support of π coincides with the set $\{x \in \mathbb{X} : \mathtt{f}\,(\,\cdot\,; \Theta_t, x) > 0\}$. Also, for the sake of notation, we will assume that the density functions are continuous with respect to the Lebesgue measure. Given this, one can define a joint distribution over the product space $(\mathbb{X} \times \mathbb{Y}, \mathscr{X} \times \mathscr{Y})$ as follows

$$\int_B \int_A \mathtt{f}\,(u; \Theta_t, v)\pi(v)\mathrm{d}v\mathrm{d}u, \tag{2.1}$$

for $A \in \mathscr{X}$, $B \in \mathscr{Y}$ and $t > 0$. Thus, the marginal distributions of (2.1) over $(\mathbb{X}, \mathscr{X})$ and $(\mathbb{Y}, \mathscr{Y})$ are π and

$$\mathsf{m}(\,\cdot\,; \Theta_t) = \int_{\mathbb{X}} \mathtt{f}\,(\,\cdot\,; \Theta_t, v)\pi(v)\mathrm{d}v,$$

respectively. Thus, the posterior distribution of π, with respect to $\mathtt{f}$, is

$$\nu_0(x; y, \Theta_t) = \frac{\mathtt{f}\,(y; \Theta_t, x)\pi(x)}{\mathsf{m}(y; \Theta_t)}, \tag{2.2}$$

for $x \in \mathbb{X}$, $y \in \mathbb{Y}$. Then, one can build a time-homogeneous process continuous in time, $(\mathbb{X}, \mathscr{X})$-valued, here denoted by $X = (X_t)_{t \geq 0}$, driven by transition probabilities

$$\mathtt{k}_t(x_0, x) = \int_{\mathbb{Y}} \nu_0(x; u, \Theta_t)\mathtt{f}\,(u; \Theta_t, x_0)\mathrm{d}u, \tag{2.3}$$

for $x_0, x \in \mathbb{X}$. Hence, if $\mathtt{k}_t$ satisfies Chapman-Kolmogorov property, then X turns out to be a standard reversible Markov process whose law is completely characterized by $\mathtt{k}_t$ and its invariant measure π.

Clearly, the main problem related to the above construction is finding the appropriate choice of the probability distributions π and $\mathtt{f}(\cdot; \Theta_t, x)$. That is to say, those distributions that make the transition probabilities $\mathtt{k}_t$ satisfy Chapman-Kolmogorov's equations. This, which is not a straightforward task, simplifies with use of Lancaster probabilities.

It is worth emphasizing that, for the discrete-time case, there have been several efforts in similar directions. For instance, Diaconis *et al.* [4] studied the sharp rate of convergence to stationary of Gibbs sampler, whereas

Letac [18] characterizes Lancaster families with given margins. Furthermore, Griffiths [8] characterized and studied the behavior of reversible stochastic processes with given stationary distributions.

3. Time-dependent Lancaster probabilities

Consider bivariate distributions $\sigma(x,y)$ with margins π and $\mathtt{m}$ defined over the measurable spaces $(\mathbb{X},\mathscr{X})$ and $(\mathbb{Y},\mathscr{Y})$, respectively. Also, we denote by (X,Y) a random vector with distribution σ. If there exists bi-orthonormal polynomials, $\{\mathtt{P}_n\}_{n\geq 0}$ and $\{\mathtt{Q}_n\}_{n\geq 0}$, associated to π and $\mathtt{m}$, respectively, *i.e.* when

$$\mathbb{E}(\mathtt{P}_n(X)\mathtt{P}_k(X)) = \delta_{n,k},\ \mathbb{E}(\mathtt{Q}_n(Y)\mathtt{Q}_k(Y)) = \delta_{n,k} \quad \text{and} \quad \mathbb{E}(\mathtt{P}_n(X)\mathtt{Q}_k(Y)) = \rho_n\delta_{n,k},$$

for $n,k\geq 0$, then the class of bivariate distributions σ are known as Lancaster probabilities. Here, the sequence $\rho_n := \mathbb{E}[\mathtt{P}_n(X)\mathtt{Q}_n(Y)]$ is known as the Lancaster sequence of σ. Recall that, bivariate distributions with marginals within the Meixner class (*cf.* [21]) possess the polynomial bi-orthogonal property (*cf.* [5]). Furthermore, it is known that for marginal distributions in the Meixner class we have

$$\sigma(x,y) = \sum_{n\geq 0}\rho_n\mathtt{P}_n(x)\mathtt{Q}_n(y)\pi(x)\mathtt{m}(y). \tag{3.1}$$

where $\mathtt{P}_0 = \mathtt{Q}_0 = 1$ (*cf.* [8]). In such a case, one can decompose Lancaster probabilities as $\sigma(x,y) = \pi(x)\mathtt{f}(y;x) = \mathtt{m}(y)\nu_0(x;y)$. Hence, given the marginal π, the conditional distribution $\mathtt{f}$ is given by

$$\mathtt{f}(y;x) = \sum_n \rho_n\mathtt{P}_n(x)\mathtt{Q}_n(y)\mathtt{m}(y).$$

Thus, the orthogonal property of $\mathtt{Q}$ leads to one-step transition probabilities given by

$$\begin{aligned}
\mathtt{k}(x_n,x_{n+1}) &= \int_{\mathbb{Y}} \nu_0(x_{n+1};y)\mathtt{f}(y;x_n)\mathtt{d}y \\
&= \int_{\mathbb{Y}} \sum_{j\geq 0}\rho_j\mathtt{P}_j(x_{n+1})\mathtt{Q}_j(y)\pi(x_{n+1}) \sum_{k\geq 0}\rho_k\mathtt{P}_k(x_n)\mathtt{Q}_k(y)\mathtt{m}(y)\mathtt{d}y \\
&= \sum_{k\geq 0}\rho_k^2\mathtt{P}_k(x_n)\mathtt{P}_k(x_{n+1})\pi(x_{n+1}),
\end{aligned} \tag{3.2}$$

for $x_n, x_{n+1}\in\mathbb{X}$, which can be associated to a discrete-time Markov process. To generalize the above methodology to the continuous-time case, our purpose here reduces to make ρ_n continuous-time dependent.

On the other hand, the spectral expansion of the transition probabilities associated to some stationary diffusion processes has the form (3.2) with $\rho_n(t) = e^{-\gamma_n t}$, for $n\geq 0$, *i.e.*

$$\mathtt{k}_t(x_0,x) = \sum_{j\geq 0}\rho_j^2(t)\mathtt{P}_j(x_0)\mathtt{P}_j(x)\pi(x),$$

where π is the invariant measure of a continuous-time homogeneous reversible Markov process X. In the case of birth and death processes, the so-called birth and death polynomials are determined uniquely by the recurrence relation

$$-x\mathtt{P}_n(x) = \mu_n\mathtt{P}_{n-1}(x) - (\lambda_n+\mu_n)\mathtt{P}_n(x) + \lambda_n\mathtt{P}_{n+1}(x),$$

for $n\geq 0$, with $\mathtt{P}_{-1} = 0$ and $\mathtt{P}_0(x) = 1$. Here, λ_i and μ_i, for $i\in S = \{-1,0,1,2,\ldots\}$, denote the birth and death rates, respectively. The state -1 is absorbing and, ignoring this state, if $\mu_0 = 0$, it becomes a reflecting state. Additionally, denoting by π the stationary distribution, when it exists, Karlin and McGregor [12] proved the

the transition function can be represented as

$$\mathbb{P}[X_t = j | X_0 = i] = \pi_j \int_0^\infty e^{-xt} \mathtt{P}_i(x) \mathtt{P}_j(x) \mathrm{d}\phi(x), \tag{3.3}$$

for $i, j = 0, 1, 2, \ldots$, and $t > 0$, where ϕ is a positive Borel measure with total mass 1 and with support on the nonnegative real axis. ϕ is known as the spectral measure of X. Taking $t = 0$ in (3.3) one easily sees that the polynomials $\{\mathtt{P}_n\}_{n\geq 0}$ are orthogonal with respect to ϕ. Note that, the dependence on t in the right-hand side of (3.3) only enters in the exponential term e^{-tx}. Furthermore, the integral representation (3.3) remains valid if one allows the possibility of absorption into the bottom state from any other state (*cf.* [25]). It is also worth emphasizing that, the transition function (3.3) is a different type of spectral expansion to the Lancaster expansion (3.2), nonetheless for self-dual polynomials, *i.e.* those polynomials satisfying $\mathtt{P}_n(x) = \mathtt{P}_x(n)$ (*cf.* [24]), we obtain the desired expression. Indeed, the orthogonal polynomials associated to the Markov chain models presented below turn out to be self-dual.

Therefore, time-dependent Lancaster probabilities can be used to build reversible Markov processes *via* their corresponding spectral expansion. On the other hand, given the spectral expansion of some continuous-time stationary Markov process, we can build time-dependent Lancaster probabilities. Either way, in the subsequent sections we explore the construction presented in Section 2 to derive new identities associated to known orthogonal polynomials.

4. Some known Lancaster probabilities associated to diffusion processes

4.1. Gamma and negative-binomial marginals

Let π be the gamma$(a, 1)$ distribution, with $a > 0$. The orthonormal polynomials associated to π, denoted by $\{\mathtt{P}_n\}_{n\geq 0}$, are given by $\mathtt{P}_n = \sqrt{n!/(a)_n} L_n^a$, where $\{L_n^a\}_{n\geq 0}$ are known as the Laguerre polynomials, and are given by

$$L_n^a(x) = \sum_{k=0}^{n} \frac{(-1)^k (1+a)_n x^k}{k!(n-k)!(1+a)_k}$$

for $x \in (0, \infty)$ (*cf.* [23]). Also, let $\mathtt{m}$ be the negative-binomial$(r, \mathtt{p})$ distribution, with $r > 0$ and $\mathtt{p} \in (0, 1)$. The orthonormal polynomials associated to $\mathtt{m}$ are the normalized Meixner polynomials, which are given by

$$\mathtt{Q}_n(y) = \sqrt{\frac{\mathtt{p}^n (r)_n}{n!}} \sum_{j=0}^{n} \frac{(-n)_j (-y)_j}{(r)_j j!} \left(1 - \frac{1}{\mathtt{p}}\right)^j,$$

(*cf.* [1]). Furthermore, Koudou [14] proved that the Lancaster probabilities with marginals π and $\mathtt{m}$ has extreme Lancaster sequences of the form $(\rho^n)_{n\geq 0}$ for $0 \leq \rho \leq \sqrt{\mathtt{p}}$. Moreover, they obtained that the following equality

$$\sum_{n\geq 0} (\sqrt{\mathtt{p}})^n \mathtt{P}_n(x) \mathtt{Q}_n(y) = \frac{x^y}{(a)_y (1-\mathtt{p})^{y+a}} \exp\left\{ - \frac{x\mathtt{p}}{1-\mathtt{p}} \right\}$$

for $(x, y) \in (0, \infty) \times \mathbb{N}$.

Then, in order to use the above Lancaster probability to build a transition probability function that leads a continuous-time Markov process, it is necessary to make the Lancaster sequence time-dependent. If we consider the function $\mathtt{p} : \mathbb{R}^+ \to (0, 1)$, defined by $\mathtt{p}(t) = e^{-\mathtt{c}t}$ for $\mathtt{c} > 0$, we obtain that

$$\rho_n(t) = e^{-\frac{\mathtt{c}n}{2} t} = \left[\sqrt{\mathtt{p}(t)}\right]^n.$$

Moreover, letting $\Theta_t = \mathtt{p}(t)/(1-\mathtt{p}(t)) = (e^{\mathtt{c}t}-1)^{-1}$, and $r = a$, the Lancaster probability takes the form,

$$\sigma(x,y) = e^{-x\Theta_t}\frac{(x\Theta_t)^y}{y!}\pi(x),$$

for $(x,y) \in (0,\infty)\times\mathbb{N}$. Hence, $\mathtt{f}$ and ν_0 have Poisson(λ) and gamma($a+y, 1+\Theta_t$) distribution, respectively. Thus, the transition probabilities $\mathtt{k}_t$ are given by

$$\mathtt{k}_t(x_0,x) = \frac{\exp\{-[\,\Theta_t(x_0+x)+x\,]\}}{(\Theta_t+1)^{-\frac{a+1}{2}}\Theta_t^{\frac{a-1}{2}}}\times\left(\frac{x}{x_0}\right)^{\frac{a-1}{2}}\times \mathtt{I}_{a-1}\left(2\sqrt{x_0 x\Theta_t(1+\Theta_t)}\right),$$

for x_0, $x \in \mathbb{R}^+$ and $\mathtt{I}_\nu(\cdot)$ denote the modified Bessel function of the first kind with argument ν. In this case, if $\Theta_t = 1/(e^{\mathtt{c}t}-1)$, with $\mathtt{c} > 0$, then $\mathtt{k}_t$ satisfies Chapman-Kolmogorov's equations (*cf.* [22]). Moreover, the process X is the only solution to the following stochastic differential equation

$$\mathrm{d}X_t = \mathtt{c}\,(a - X_t)\,\mathrm{d}t + \sqrt{2\mathtt{c}X_t}\mathrm{d}W_t,$$

where $W = (W_t)_{t\geq 0}$ is a standard Brownian motion. This process constitutes a re-parameterization of the Cox-Ingersoll-Ross (*cf.* [3]).

For this example we deduce the value of the eigenvalues are $\gamma_n = \frac{\mathtt{c}n}{2}$. Hence, this model allows us to see how we can use Lancaster probabilities to build continuous-time reversible Markov processes.

4.2. Normal marginals

The mean reverting Ornstein-Uhlenbeck diffusion (*cf.* [2]) is a real-valued process, here denoted by $X = \{X_t\}_{t\geq 0}$, driven by the following SDE

$$\mathrm{d}X_t = -\mathtt{c}(X_t - \gamma)\mathrm{d}t + \sigma\mathrm{d}W_t \tag{4.1}$$

for $\mathtt{c} > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$ and, where $W = (W_t)_{t\geq 0}$ is a standard Brownian motion. Such a diffusion turns out to be reversible with invariant $\mathrm{N}(\gamma, \frac{\sigma^2}{2\mathtt{c}})$ distribution, where N stands for the normal distribution. Moreover, the spectral expansion for the transition probabilities associated to X, are given by

$$\mathtt{k}_t(x_0,x) = \sum_{n=0}^{\infty} e^{-\mathtt{c}nt}\mathtt{P}_n(x_0)\mathtt{P}_n(x)\pi(x),$$

for x_0, $x \in \mathbb{R}^+$, where π denotes the density function associated to the invariant measure of X and $\{\mathtt{P}_n\}_{n\geq 0}$ are the Hermite polynomials standardized to a $\mathrm{N}(\gamma, \frac{\sigma^2}{2\mathtt{c}})$ distribution.

On the other hand, letting π be a $\mathrm{N}(\gamma,\alpha)$ distribution and $\mathtt{f}(\cdot;\Theta_t,x)$ be a $\mathrm{N}(x\Theta_t, \alpha(1-\Theta_t)\Theta_t)$ distribution, where $\Theta : \mathbb{R}^+ \to (0,1)$, the posterior distribution of ν_0 has $\mathrm{N}(y+\gamma(1-\Theta_t), (1-\Theta_t)\alpha)$ distribution, *i.e.* π is conjugate with respect to $\mathtt{f}$. Thus, the transition probabilities of X are

$$\mathtt{k}_t(x_0,x) = \frac{1}{\sqrt{2\pi\alpha(1-\Theta_t^2)}}\exp\Big\{-\frac{\big[x - x_0\Theta_t - \gamma(1-\Theta_t)\big]^2}{2\alpha(1-\Theta_t^2)}\Big\},$$

for $x_0, x \in \mathbb{R}$. Moreover, $\mathtt{k}_t$ satisfies Chapman-Kolmogorov's equations if and only if $\Theta_t = e^{-\mathtt{c}t}$ for $\mathtt{c} > 0$ (*cf.* [22]). In particular, letting $\alpha = \sigma^2/(2\mathtt{c})$ we have that $\mathtt{k}_t$ drives a diffusion whose SDE coincides with the equation (4.1).

Furthermore, since $\mathtt{Q}_n$ are the Hermite polynomials standardized to a $\mathrm{N}(x\Theta_t, \alpha\Theta_t)$ distribution, whose density function is denoted by $\mathtt{m}(\cdot;\Theta_t)$, we notice that $\mathtt{P}_n$ and $\mathtt{Q}_n$ are bi-orthonormal polynomials with respect to π and $\mathtt{m}$, respectively. As a consequence, we are able to define a Lancaster probability *via* the product

$\sigma_t(x,y) = \pi(x)\mathtt{f}(y;x,\Theta_t)$, whose Lancaster sequence is given by $\rho_n(t) = \exp\{-\mathtt{c}nt\}$, with $\mathtt{c} > 0$. Thus, we obtain the following equalities

$$\mathrm{N}(y; x\Theta_t, \alpha\Theta_t(1-\Theta_t)) = \sum_{n\geq 0} \rho_n(t)\mathtt{P}_n(x)\mathtt{Q}_n(y)\mathtt{m}(y),$$
$$\mathrm{N}(x; y+\gamma(1-\Theta_t), \alpha(1-\Theta_t)) = \sum_{n\geq 0} \rho_n(t)\mathtt{P}_n(x)\mathtt{Q}_n(y)\pi(x).$$

where $\mathrm{N}(x;\gamma,\alpha)$ denotes the density function of the $\mathrm{N}(\gamma,\alpha)$ distribution. Moreover, the Lancaster sequence satisfies the equality $\rho_n(t) = \mathbb{E}[\mathtt{P}_n(X)\mathtt{Q}_n(Y)]$, here (X,Y) denotes the random vector with distribution σ.

5. Time-dependent Lancaster probabilities derived from some birth and death processes

5.1. The $M/M/\infty$ queue model

The telephone trunking problem is defined as a reversible birth and death process with birth and death rates $\lambda_n = \lambda$ and $\mu_n = \mu n$, respectively, with invariant Poisson(λ/μ) distribution (*cf.* [11]). Such a process is also known as the $M/M/\infty$ queue model (*cf.* [13]). Further, letting $\mu = \mathtt{c}$ and $\lambda = \beta\mathtt{c}$, or equivalently, $\beta = \lambda/\mu$ and $\mathtt{c} = \mu$, the infinitesimal generator $\{q_{i,j}\}$ of the $M/M/\infty$ model takes the form

$$q_{ij} = \begin{cases} -\mathtt{c}(i+\beta), & i = j, \\ \mathtt{c}\beta, & i = j+1, \\ \mathtt{c}i, & i = j-1. \end{cases}$$

for $i,j = 0,1,2,\ldots$ The above definition has the inconvenient that the transition probabilities associated to the model are unknown. Hence, an expression for such a transition can be obtained from (3.3) where the polynomials $\{\mathtt{P}_n\}_{n\geq 0}$ are given by the Charlier polynomials. Let us recall that, the Charlier polynomials, denoted by $\mathtt{C} = \{\mathtt{C}_n(\cdot;\beta)\}_{n\geq 0}$, are given by

$$\mathtt{C}_n(x;\beta) = \sum_{k=0}^{n} (-1)^k \binom{n}{k}\binom{x}{k}\frac{k!}{\beta^x},$$

for $\beta \geq 0$. Moreover, $\mathtt{C}$ satisfy an orthogonal relation with respect to the Poisson distribution, *i.e.*

$$\sum_{x=0}^{\infty} \mathtt{C}_n(x;\beta)\mathtt{C}_m(x;\beta)\mathrm{Po}(x;\beta) = \frac{n!}{\beta^n}\delta_{n,m}.$$

where $\mathrm{Po}(x;\beta)$ denotes the probability function of the Poisson distribution with mean β, evaluated in x. Also, it is known that, the Charlier polynomials are self-dual, *i.e.* $\mathtt{C}_n(x;\beta) = \mathtt{C}_x(n;\beta)$ for $n, x \geq 0$ (*cf.* [24]). Thus, the spectral expansion of the transition probability function of the $M/M/\infty$ model takes the form

$$\mathtt{p}_{i,j}(t) = \frac{\beta^j}{j!}\sum_{n=0}^{\infty} e^{-n\mu t}\mathtt{C}_n(i;\beta)\mathtt{C}_n(j;\beta)\pi(n)$$
$$= \sum_{n=0}^{\infty} e^{-n\mu t}\mathtt{P}_n(i;\beta)\mathtt{P}_n(j;\beta)\pi(j),$$

for $i,j = 0,1,2,\ldots$, where $\mathtt{P}_n(\cdot;\beta) = \frac{\beta^{n/2}}{\sqrt{n!}}\mathtt{C}_n(\cdot;\beta)$ for $n \geq 0$.

On the other hand, Mena and Walker [22] derived a tractable expression for the transition probabilities of the $M/M/\infty$ queue model. Indeed, let π be the Poisson(β) distribution, with $\beta > 0$; for each $x \in \mathbb{Z}^+ \cup \{0\}$, let

$\mathtt{f}\,(\cdot;\Theta_t,x)$ be the binomial(x,Θ_t) distribution, whose probability function is denoted by $\mathrm{Bin}(\mathtt{n},\mathtt{p})$, and also we consider a function $\Theta:\mathbb{R}^+\to[0,1]$. In this case, the marginal distribution $\mathtt{m}$ has Poisson$(\beta\Theta_t)$ distribution and the posterior distribution of π is given by

$$\nu_0(x;y,\Theta_t)=e^{-\beta(1-\Theta_t)}\frac{\left[\beta(1-\Theta_t)\right]^{x-y}}{(x-y)!},$$

for $x\in\{y,\ldots,\infty\}$. Thus, the transition probabilities of the model are

$$\mathtt{k}_t(x_0,x)=e^{-\beta(1-\Theta_t)}\sum_{u=0}^{x_0\wedge x}\frac{\left[\beta(1-\Theta_t)\right]^{x-u}}{(x-u)!}\binom{x_0}{u}\Theta_t^u(1-\Theta_t)^{x_0-u},$$

for $x_0,\,x\in\mathbb{Z}^+\cup\{0\}$. For this case, if $\Theta_t=e^{-\mathtt{c}t}$, with $\mathtt{c}>0$, then the transition $\mathtt{k}_t$ satisfies the Chapman-Kolmogorov's equations.

Now, we define the polynomials $\mathtt{P}=\{\mathtt{P}_n(\cdot;\beta)\}_{n\geq 0}$ and $\mathtt{Q}=\{\mathtt{Q}_n(\cdot;\beta\Theta_t)\}_{n\geq 0}$ by $\mathtt{P}_n(\cdot;\beta)=(\beta^{n/2}/\sqrt{n!})\mathtt{C}_n(\cdot;\beta)$ and $\mathtt{Q}_n(\cdot;\beta\Theta_t)=([\beta\Theta_t]^{n/2}/\sqrt{n!})\mathtt{C}_n(\cdot;\beta\Theta_t)$. Then, $\mathtt{P}$ and $\mathtt{Q}$ are orthonormal polynomials with respect to the probability measures π and $\mathtt{m}$, respectively. Let us notice that $\mathtt{P}_n(\cdot;\beta)=\mathtt{Q}_n(\cdot;\beta)$ for all $n\geq 0$ and that the self-duality of the Charlier polynomials implies that $\mathtt{P}$ and $\mathtt{Q}$ are also self-dual. Furthermore, our proposal allows to define a random vector (X,Y) with bivariate distribution

$$\sigma(x,y):=\mathrm{Po}(x;\beta)\mathrm{Bin}(y;x,\Theta_t)=\mathrm{Po}(x-y;\beta(1-\Theta_t))\mathrm{Po}(y;\beta\Theta_t),$$

for $t\geq 0$, $y\in\{0,1,2,\ldots,x\}$ and $x\geq 0$. Hence, self-duality of $\mathtt{P}$ and $\mathtt{Q}$ implies that

$$\begin{aligned}
\mathbb{E}[\mathtt{P}_i(X;\beta)\mathtt{Q}_j(Y;\beta\Theta_t)] &= \sum_{y=0}^{\infty}\sum_{x=y}^{\infty}\mathtt{P}_i(x;\beta)\mathtt{Q}_j(y;\beta\Theta_t)\mathrm{Po}(x-y;\beta(1-\Theta_t))\mathrm{Po}(y;\beta\Theta_t)\\
&= \sum_{y=0}^{\infty}\mathtt{Q}_j(y;\beta\Theta_t)\mathrm{Po}(y;\beta\Theta_t)\left[\sum_{x=0}^{\infty}\mathtt{C}_{x+y}(i;\beta)\mathrm{Po}(x;\beta(1-\Theta_t))\right]\frac{\beta^{i/2}}{\sqrt{i!}}\\
&= \sum_{y=0}^{\infty}\mathtt{Q}_j(y;\beta\Theta_t)\mathrm{Po}(y;\beta\Theta_t)\left[\mathtt{C}_y(i;\beta\Theta_t)\Theta_t^i\right]\frac{\beta^{i/2}}{\sqrt{i!}}\\
&= \Theta_t^{i/2}\sum_{y=0}^{\infty}\mathtt{Q}_j(y;\beta\Theta_t)\mathtt{Q}_i(y;\beta\Theta_t)\mathrm{Po}(y;\beta\Theta_t)\\
&= e^{-\frac{\mathtt{c}ti}{2}}\delta_{i,j}
\end{aligned}$$

for $i,j=0,1,2,\ldots$, where the third equality holds since

$$\sum_{n=0}^{\infty}\mathtt{C}_{n+y}(x;\beta)\mathrm{Po}(n;y)=\mathtt{C}_y(x;\beta-y)\left(1-\frac{y}{\beta}\right)^x. \tag{5.1}$$

(*cf.* [20]). Therefore, the set of polynomials $\{\mathtt{P}_n\}_{n\geq 0}$ and $\{\mathtt{Q}_n(\cdot;\Theta_t)\}_{n\geq 0}$ are bi-orthogonal. Thus, the bivariate distribution (5.1) is a Lancaster probability, whose Lancaster sequences are given by $\rho_n(t)=\mathbb{E}[\mathtt{P}_n(X)\mathtt{Q}_n(Y;\Theta_t)]=e^{-n\mathtt{c}t/2}$, for $\mathtt{c}>0$.

Furthermore, since the bivariate distribution σ takes the form (3.1), we have

$$\mathrm{Bin}(y;x,\Theta_t)=\sum_{n\geq 0}\rho_n(t)\mathtt{P}_n(x;\beta)\mathtt{Q}_n(y;\beta\Theta_t)\mathtt{m}(y;\Theta_t), \tag{5.2}$$

$$\mathrm{Po}(x-y;\beta(1-\Theta_t)) = \sum_{n\geq 0} \rho_n(t)\mathtt{P}_n(x;\beta)\mathtt{Q}_n(y;\beta\Theta_t)\pi(x). \tag{5.3}$$

Let us notice that, to the best of our knowledge, the equalities (5.2) and (5.3) are new in literature. Clearly, this opens a mechanism to derive expressions for orthogonal polynomials.

5.2. The simple birth, death and immigration process

A stationary birth and death process with immigration is a process where individuals immigrate at rate ν, give birth to additional individuals at rate λ and die at rate μ, *i.e.* its infinitesimal rates are $\lambda_n = \lambda n + \nu$ and $\mu_n = \mu n$, for $n \geq 0$. In particular, the simple birth, death and immigration process has invariant negative-binomial distribution. This process is also known as linear growth process characterized as the process with birth and death rates $\lambda_n = \lambda n + \nu$ and $\mu_n = \mu n$, respectively (*cf.* [11]). The spectral expansion of the transition probabilities for this model is given with respect to the Meixner polynomials $\mathtt{M} = \{\mathtt{M}_n(\cdot; r, \mathtt{p})\}_{n\geq 0}$, where

$$\mathtt{M}_n(x;r,\mathtt{p}) = \sum_{k=0}^{n} \frac{(-n)_k(-x)_k}{(r)_k k!}\Big(1-\frac{1}{p}\Big)^k.$$

The Meixner polynomials are self-dual, *i.e.* $\mathtt{M}_n(x;r,\mathtt{p}) = \mathtt{M}_x(n;r,\mathtt{p})$ for $n, \geq 0$ (*cf.* [24]). Moreover, denoting the probability function of the negative-binomial(r,p) distribution by $\mathrm{NB}(\cdot;r,p)$, we have

$$\sum_{x=0}^{\infty} \mathtt{M}_n(x;r,\mathtt{p})\mathtt{M}_m(x;r,\mathtt{p})\mathrm{NB}(x;r,\mathtt{p}) = \frac{n!}{\mathtt{p}^n(r)_n}\delta_{n,m}.$$

This implies that, the spectral expansion of the simple birth, death and immigration process is

$$\begin{aligned}\mathtt{p}_{i,j}(t) &= \frac{(r)_j\mathtt{p}^j}{j!}\sum_{n=0}^{\infty} e^{-n(\mu-\lambda)t}\mathtt{M}_n(i;r,\mathtt{p})\mathtt{M}_n(j;r,\mathtt{p})\pi(n)\\ &= \sum_{n=0}^{\infty} e^{-n(\mu-\lambda)t}\mathtt{P}_n(i;r,\mathtt{p})\mathtt{P}_n(j;r,\mathtt{p})\pi(j)\end{aligned}$$

for $i, j = 0, 1, 2, \ldots$, where $\mathtt{P}_n(\cdot;r,\mathtt{p}) = \sqrt{\frac{(r)_n\mathtt{p}^n}{n!}}\mathtt{M}_n(\cdot;r,\mathtt{p})$ for $n \geq 0$ and, π has $\mathrm{NB}(r,\mathtt{p})$ distribution.

On the other hand, letting π be a negative-binomial distribution and; for $x \in \{0,1,2,\ldots\}$ and $\Theta : \mathbb{R}^+ \to (0,1)$, let $\mathtt{f}(\cdot;x,\Theta_t)$ be a $\mathrm{bin}(\cdot;x,\Theta_t)$ distribution. Then, $\mathtt{m}$ has $\mathrm{NB}(\mathtt{r},\mathtt{p}\Theta_t/[1-\mathtt{p}(1-\Theta_t)])$ distribution and, $\nu_0(x;y,\Theta_t)$, as a function of $x-y$, has $\mathrm{NB}(\mathtt{r}+y,\mathtt{p}(1-\Theta_t))$ distribution. Thus, we denote by $(X_t)_{t\geq 0}$, a process with transition probabilities

$$\mathtt{k}_t(x_0,x) = \sum_{y=0}^{x_0\wedge x}\binom{x+r-1}{x-y}\big[\mathtt{p}(1-\Theta_t)\big]^{x-y}\big[1-\mathtt{p}(1-\Theta_t)\big]^{y+r}\binom{x_0}{y}\Theta_t^y(1-\Theta_t)^{x_0-y},$$

for $x_0, x \in \{0,1,2,\ldots\}$. In fact, if $\Theta_t = (1-\mathtt{p})/(e^{\mathtt{c}t}-\mathtt{p})$ for $\mathtt{c} > 0$, then the transition $\mathtt{k}_t$ satisfies Chapman-Kolmogorov's equations. Indeed, $(X_t)_{t\geq 0}$ is the birth, death and immigration process described in this section with birth rate $\mu = \mathtt{c}/(1-\mathtt{p})$, death rate $\lambda = \mathtt{c}\mathtt{p}/(1-\mathtt{p})$ and immigration rate $\nu = \mathtt{c}\mathtt{p}r/(1-\mathtt{p})$, or equivalently, $\mathtt{p} = \lambda/\mu$, $r = \nu/\lambda$ and $\mathtt{c} = \mu - \lambda$.

Now, consider the polynomials $\mathtt{P} = \{\mathtt{P}_n(\cdot;r,\mathtt{p})\}_{n\geq 0}$ and $\mathtt{Q} = \{\mathtt{Q}_n(\cdot;r,\tilde{\mathtt{p}}_t)\}_{n\geq 0}$, defined by $\mathtt{P}_n(\cdot;r,\mathtt{p}) = \sqrt{\mathtt{p}^n(r)_n/n!}\mathtt{M}_n(\cdot;r,\mathtt{p})$ and $\mathtt{Q}_n(\cdot;r,\tilde{\mathtt{p}}_t) = \sqrt{\tilde{\mathtt{p}}_t^n(r)_n/n!}\mathtt{M}_n(\cdot;r,\tilde{\mathtt{p}}_t)$, where $\tilde{\mathtt{p}}_t = \mathtt{p}\Theta_t/[1-\mathtt{p}(1-\Theta_t)]$. Hence, $\mathtt{P}$ and $\mathtt{Q}$ are orthonormal polynomials with respect to π and $\mathtt{m}$, respectively. Thus, as in the $M/M/\infty$ queue case, we define the random vector (X, Y) with bivariate distribution

$$\sigma(x,y) := \text{NB}(x; r, \mathtt{p})\text{Bin}(y; x, \Theta_t) = \text{NB}(x-y; r+y, \mathtt{p}(1-\Theta_t))\text{NB}(y; r, \tilde{\mathtt{p}}_t), \tag{5.4}$$

for $t \geq 0$ and $y \in \{0, 1, 2, \ldots, x\}$ where $x \geq 0$. Then, the self-duality of the Meixner polynomials implies that $\mathtt{P}$ and $\mathtt{Q}$ are self-dual, so that

$$\begin{aligned}
&\mathbb{E}[\mathtt{P}_i(X; r, \mathtt{p})\mathtt{Q}_j(Y; r, \tilde{\mathtt{p}})] \\
&= \sum_{y=0}^{\infty}\sum_{x=y}^{\infty} \mathtt{P}_i(x; r, \mathtt{p})\mathtt{Q}_j(y; r, \tilde{\mathtt{p}})\text{NB}(x; r, \mathtt{p})\text{Bin}(y; x, \Theta_t) \\
&= \sum_{y=0}^{\infty} \mathtt{Q}_j(y; r, \tilde{\mathtt{p}}) \frac{(1-\mathtt{p})^r \Theta_t^y}{y!} \sum_{x=0}^{\infty} \mathtt{P}_i(x+y; r, \mathtt{p}) \frac{(r)_{x+y}}{x!} p^{x+y}(1-\Theta_t)^x \\
&= \sum_{y=0}^{\infty} \mathtt{Q}_j(y; r, \tilde{\mathtt{p}}) \frac{(1-\mathtt{p})^r \Theta_t^y \mathtt{p}^y}{y!} \sqrt{\frac{\mathtt{p}^i (r)_i}{i!}} \sum_{x=0}^{\infty} \mathtt{M}_{x+y}(i; r, \mathtt{p}) \frac{(r)_{x+y}}{x!} [p(1-\Theta_t)]^x \\
&= \left[\frac{\Theta_t}{1-\mathtt{p}(1-\Theta_t)}\right]^{i/2} \sum_{y=0}^{\infty} \mathtt{Q}_j(y; r, \tilde{\mathtt{p}})\mathtt{Q}_i(y; r, \tilde{\mathtt{p}})\pi(y) \\
&= e^{-\frac{cti}{2}} \delta_{i,j}.
\end{aligned}$$

for $i, j = 0, 1, 2, \ldots$, where the third equality holds since

$$\sum_{n=0}^{\infty} \mathtt{M}_{n+y}(i; r, \mathtt{p}) \frac{(r)_{n+y}}{n!} z^n = \left(1 - \frac{z}{\mathtt{p}}\right)^i (1-z)^{-i-r-y} \left[(r)_y \mathtt{M}_y\left(i; r, \frac{\mathtt{p}-z}{1-z}\right)\right]. \tag{5.5}$$

This statement holds since bivariate probabilities with marginals, in the class of Meixner distributions, are Lancaster probabilities of the form (3.1), then (5.5) follows and we also obtain

$$\text{Bin}(y; x, \Theta_t) = \sum_{n\geq 0} \rho_n(t)\mathtt{P}_n(x)\mathtt{Q}_n(y; \Theta_t)\mathsf{m}(y; \Theta_t), \tag{5.6}$$

$$\text{NB}(x-y; \mathtt{r}, \mathtt{p}(1-\Theta_t)) = \sum_{n\geq 0} \rho_n(t)\mathtt{P}_n(x)\mathtt{Q}_n(y; \Theta_t)\pi(x), \tag{5.7}$$

for $\rho_n(t) = \mathbb{E}[\mathtt{P}_n(X; r, \mathtt{p})\mathtt{Q}_n(Y; r, \tilde{\mathtt{p}})] = \exp\{-ctn/2\}$. To the best of our knowledge equalities (5.5), (5.6) and (5.7) are new in literature.

6. Wright-Fisher diffusion

In this section we use the fact that the Jacobi and the Hahn polynomials are orthogonal with respect to the beta and the beta-binomial distributions (*cf.* [24]), respectively, to build Lancaster probabilities. Let us recall that, the Hahn polynomials, here denoted by $\{\mathtt{Q}_n(y; \alpha, \beta, N)\}_{n\geq 0}$, can be expressed in terms of the generalized hypergeometric function as follows

$$\mathtt{Q}_n(y; \alpha, \beta, N) = {}_3F_2(-n, n+\alpha+\beta+1, -y; \alpha+1, -N; 1),$$

for $N \geq 1$ and $\alpha, \beta > -1$. Moreover, it is known that $\{\mathtt{Q}_n(y; \alpha, \beta, N)\}_{n\geq 0}$ satisfies

$$\sum_{y=0}^{N} \mathtt{Q}_n(y; \alpha, \beta, N)\mathtt{Q}_m(y; \alpha, \beta, N)\mathsf{m}(y) = \mathtt{d}_n^2 \delta_{n,m},$$

for $n, m = 0, 1, 2, \ldots, N$ (*cf.* [24], A. B9); where $\mathtt{m}(y)$ denotes the beta-binomial$(y; N, \alpha+1, \beta+1)$ distribution, and

$$\pi_n = \frac{1}{\mathtt{d}_n^2} = \binom{N}{n} \frac{2n+\alpha+\beta+1}{\alpha+\beta+1} \frac{\mathtt{B}(\beta+1,n)}{\mathtt{B}(\alpha+1,n)} \frac{\mathtt{B}(\alpha+\beta+2+N,n)}{\mathtt{B}(\alpha+\beta+1,n)}.$$

where $\mathtt{B}$ denotes the beta function. Thus, $\tilde{\mathtt{Q}}_n(y;\alpha,\beta,N) = \sqrt{1/\mathtt{d}_n^2}\mathtt{Q}_n(y;\alpha,\beta,N)$ form a set of orthonormal polynomials with respect to $\mathtt{m}$.

On the other hand, the Jacobi polynomials, here denoted by $\{\mathtt{P}_n^{(\alpha,\beta)}(x)\}_{n\geq 0}$, can be expressed in terms of the hypergeometric function as follows

$$\mathtt{P}_n^{(\alpha,\beta)}(x) = \frac{(\alpha+1)_n}{n!}\, {}_2\mathtt{F}_1\Big(-n, n+\alpha+\beta+1; \alpha+1; \frac{1-x}{2}\Big),$$

for $\alpha, \beta > -1$. Also from ([24], A. B8), $\{\mathtt{Q}_n(y;\alpha,\beta,N)\}_{n\geq 0}$ satisfies the equality

$$\int_{\mathbb{R}} \mathtt{P}_n^{(\alpha,\beta)}(x)\mathtt{P}_m^{(\alpha,\beta)}(x)\pi(x)\mathrm{d}x = \mathtt{g}_n^2\delta_{n,m},$$

where

$$\pi(x) = \frac{(1-x)^\alpha(1+x)^\beta}{2^{\alpha+\beta+1}B(\alpha+1,\beta+1)} = \frac{1}{2}\mathrm{Be}\Big(\frac{1-x}{2}, \alpha+1, \beta+1\Big),$$

for $x \in (-1,1)$, here $\mathrm{Be}(x;\alpha,\beta)$ denotes the density function of the beta(α,β) distribution, and

$$\mathtt{g}_n^2 = \frac{\Gamma(n+\beta+1)\Gamma(n+\alpha+1)}{\mathtt{B}(\alpha+1,\beta+1)(2n+\alpha+\beta+1)n!\Gamma(n+\alpha+\beta+1)}.$$

Thus, $\tilde{\mathtt{P}}_n^{(\alpha,\beta)}(x) = \sqrt{1/\mathtt{g}_n^2}\mathtt{P}_n^{(\alpha,\beta)}(x)$ form a set of orthonormal polynomials with respect to the probability measure π.

Thus, in order to build a Lancaster probability, we define the random vector (X,Y) with bivariate distribution

$$\sigma(x,y) = \pi(x)\mathrm{Bin}\Big(y; N, \frac{1-x}{2}\Big) = \mathtt{m}(y)\frac{1}{2}\mathrm{Be}\Big(\frac{1-x}{2}; \alpha+1+y, \beta+1+N-y\Big),$$

which has π and $\mathtt{m}$ as marginal distributions. Hence, if the set of polynomials $\tilde{Q}_n(y;\alpha,\beta,N)$ and $\tilde{P}_n^{(\alpha,\beta)}(x)$ are bi-orthonormal then σ turns out to be a Lancaster probability that takes the form (3.1). Indeed, since we have that

$$\begin{aligned}
&\mathbb{E}[\tilde{\mathtt{Q}}_m(Y;\alpha,\beta,N)\tilde{\mathtt{P}}_n^{(\alpha,\beta)}(X)] \\
&= \sum_{y=0}^{N} \tilde{\mathtt{Q}}_m(y;\alpha,\beta,N)\nu(y) \int_{\mathbb{R}} \tilde{\mathtt{P}}_n^{(\alpha,\beta)}(x)\frac{1}{2}\mathrm{Be}\Big(\frac{1-x}{2}; \alpha+1+y, \beta+1+N-y\Big)\mathrm{d}x \\
&= \sum_{y=0}^{N} \tilde{\mathtt{Q}}_m(y;\alpha,\beta,N)\nu(y)\sqrt{\frac{1}{\mathtt{g}_n^2}}\frac{(\alpha+1)_n}{n!}\, {}_3\mathtt{F}_2(-n, n+\alpha+\beta+1, \alpha+1+y; \alpha+1, \alpha+\beta+2+N; 1) \\
&= \rho_n \sum_{y=0}^{N} \tilde{\mathtt{Q}}_m(y;\alpha,\beta,N)\nu(y)\sqrt{\frac{1}{\mathtt{d}_n^2}}\, {}_3F_2(-n, n+\alpha+\beta+1, -y; \alpha+1, -N; 1) \\
&= \rho_n \sum_{y=0}^{N} \tilde{\mathtt{Q}}_m(y;\alpha,\beta,N)\tilde{\mathtt{Q}}_n(y;\alpha,\beta,N)\nu(y) \\
&= \rho_n\delta_{n,m},
\end{aligned}$$

where the second equality holds since

$$
\begin{aligned}
&{}_3\mathtt{F}_2(-n, n+\alpha+\beta+1, \alpha+1+y; \alpha+1, \alpha+\beta+2+N; 1)\\
&\quad = \frac{N!\Gamma(N+\alpha+\beta+2)}{(N-n)!\Gamma(N+\alpha+\beta+2+n)}\, {}_3F_2(-n, n+\alpha+\beta+1, -y; \alpha+1, -N; 1),
\end{aligned}
$$

and

$$
\rho_n = \sqrt{\frac{N!\Gamma(N+\alpha+\beta+2)}{(N-n)!\Gamma(N+\alpha+\beta+2+n)}}.
$$

Thus, $(\tilde{\mathtt{P}}_n^{(\alpha,\beta)})_{n\geq 0}$ and $(\tilde{\mathtt{Q}}_n)_{n\geq 0}$ are bi-orthogonal polynomials. Therefore, the bivariate distribution σ is a Lancaster probability with margins π and ν, and Lancaster sequences ρ_n. In addition, this result implies that

$$
\frac{1}{2}\mathrm{Be}\Big(\frac{1-x}{2}; \alpha+1+y, \beta+1+N-y\Big) = \sum_{n\geq 1} \rho_n \tilde{\mathtt{P}}_n^{(\alpha,\beta)}(x)\tilde{\mathtt{Q}}_n(y;\alpha,\beta,N)\pi(x), \tag{6.1}
$$

and

$$
\mathrm{Bin}\Big(y; N, \frac{1-x}{2}\Big) = \sum_{n\geq 1} \rho_n \tilde{\mathtt{P}}_n^{(\alpha,\beta)}(x)\tilde{\mathtt{Q}}_n(y;\alpha,\beta,N)\nu(y). \tag{6.2}
$$

Furthermore, one can define an homogeneous discrete-time Markov process $(X_n)_{n\geq 0}$ driven by the transition probabilities

$$
\begin{aligned}
\mathtt{k}(x_0, x) &= \sum_{y=0}^{N} \frac{1}{2}\mathrm{Be}\Big(\frac{1-x}{2}; \alpha+1+y, \beta+1+N-y\Big)\mathrm{Bin}\Big(y; N, \frac{1-x_0}{2}\Big)\\
&= \sum_{n=0}^{N} \rho_n^2 \tilde{\mathtt{P}}_n^{(\alpha,\beta)}(x)\tilde{\mathtt{P}}_n^{(\alpha,\beta)}(x_0)\pi(x),
\end{aligned} \tag{6.3}
$$

for $x_0, x \in (-1, 1)$. The second equality holds due to the construction presented in Section 3. Let us notice that, the Markov process $(X_n)_{n\geq 0}$ is stationary and reversible with invariant measure π.

Now, in order to generalize the above model to the continuous-time framework, it is necessary to aggregate a time-dependence, which in this case has to be done through the parameter N. In this case, since N takes values in $\{0, 1, 2, \ldots\}$, such dependence is not given through a deterministic function, instead we consider a Markov process $(N_t)_{t\geq 0}$. In fact, denoting by $q_N^{\alpha+\beta+2}(t)$, the transition probabilities of a death process with an entrance boundary of infinity, with death rates $j(j+\alpha+\beta+2-1)$, $k \geq 1$. The transition probabilities

$$
k_t(x_0, x) = \sum_{N=0}^{\infty} q_N^{\alpha+\beta+2}(t) \sum_{y=0}^{N} \frac{1}{2}\mathrm{Be}\Big(\frac{1-x}{2}; \alpha+1+y, \beta+1+N-y\Big)\mathrm{Bin}\Big(y; N, \frac{1-x_0}{2}\Big),
$$

and the invariant beta distribution π characterize the Wright-Fisher diffusion, with parameters $(\alpha+1, \beta+1)$ (*cf.* [6]). An explicit expression for such a transition function is

$$
q_N^{\alpha+\beta+2}(t) = \sum_{j=N}^{\infty} e^{-\frac{1}{2}j(j+\alpha+\beta+2-1)t}(-1)^{j-N}\frac{(2j+\alpha+\beta+2-1)(N+\alpha+\beta+2)_{j-1}}{N!(j-N)!}.
$$

See also Walker *et al.* [26] and Griffiths and Spanó [9]. Then, equality (6.3) leads us to the following equalities

$$\begin{aligned} k_t(x_0,x) &= \sum_{N=0}^{\infty} q_N^{\alpha+\beta+2}(t) \sum_{y=0}^{N} \frac{1}{2}\mathrm{Be}\Big(\frac{1-x}{2};\alpha+1+y,\beta+1+N-y\Big)\mathrm{Bin}\Big(y;N,\frac{1-x_0}{2}\Big) \\ &= \sum_{N=0}^{\infty} q_N^{\alpha+\beta+2}(t) \sum_{n=0}^{N} \rho_n^2 \tilde{\mathrm{P}}_n^{(\alpha,\beta)}(x)\tilde{\mathrm{P}}_n^{(\alpha,\beta)}(x_0)\pi(x) \\ &= \sum_{n=0}^{\infty} \tilde{\mathrm{P}}_n^{(\alpha,\beta)}(x)\tilde{\mathrm{P}}_n^{(\alpha,\beta)}(x_0)\pi(x) \sum_{N=n}^{\infty} q_N^{\alpha+\beta+2}(t) \frac{N!\Gamma(N+\alpha+\beta+2)}{(N-n)!\Gamma(N+\alpha+\beta+2+n)} \\ &= \sum_{n\geq 0} e^{-\frac{1}{2}n(n+\alpha+\beta+2-1)t}\tilde{\mathrm{P}}_n^{(\alpha,\beta)}(x)\tilde{\mathrm{P}}_n^{(\alpha,\beta)}(x_0)\pi(x). \end{aligned}$$

where the last equality holds since the later expression is the spectral expansion of the Wright-Fisher diffusion. This implies that, the transitions function $q_N^{\alpha+\beta+2}(t)$ satisfies the equality

$$e^{-\frac{1}{2}n(n+\alpha+\beta+2-1)t} = \sum_{N=n}^{\infty} q_N^{\alpha+\beta+2}(t)\frac{N!\Gamma(N+\alpha+\beta+2)}{(N-n)!\Gamma(N+\alpha+\beta+2+n)}. \tag{6.4}$$

Therefore, we proved that the Hahn and the Jacobi polynomials are bi-orthogonal, which led us to build a Lancaster probability with beta and beta-binomial margins. We also derive its corresponding Lancaster sequence. This allows us to find the equalities (6.1), (6.2), (6.3) and (6.4), which again, to our knowledge, are new in literature.

Acknowledgements. The first author acknowledges the support of the national scholarship program and project 241195, both by CONACyT. The second author thanks the support provided by CONTEX project 2018-9B. This work was concluded while the second author was on Global Chair Visiting position at the University of Bath. Finally, we thank to the referees for their detailed and helpful comments.

References

[1] H. Bateman and A. Erdélyi, Vol. 1 of *Higher transcendental functions.* McGraw-Hill, New York (1953).
[2] T. Björk, Arbitrage theory in continuous time. Oxford University Press (2009).
[3] J. Cox, J. Ingersoll and S. Ross, A theory of the term structure of interest rates. *Theory Valuat.* **53** (2005) 126–164.
[4] P. Diaconis, K. Khare and L. Saloff-Coste, Gibbs sampling, exponential families and orthogonal polynomials. *Stat. Sci.* **23** (2008) 151–178.
[5] G.K. Eagleson, Polynomial expansions of bivariate distributions. *Ann. Math. Stat.* **35** (1964) 1208–1215.
[6] S.N. Ethier and R.C. Griffiths, The transition function of a Fleming Viot process. *Ann. Probab.* **21** (1993) 1571–1590.
[7] S.N. Ethier and T.G. Kurtz, *Markov processes: characterization and convergence.* John Wiley & Sons (2009).
[8] B. Griffiths, Stochastic processes with orthogonal polynomial eigenfunctions. *J. Comp. Appl. Math.* **233** (2008) 739–744.
[9] R.C. Griffiths and D. Spanó, Diffusion processes and coalescent trees. *London Mathematical Society Lecture Notes Series* (2010).
[10] P.A. Jenkins and D. Spanó, Exact simulation of the Wright-Fisher diffusion. *Ann. Appl. Probab.* **27** (2017) 1478–1509.
[11] S. Karlin, A first course in stochastic processes. Academic Press (2014).
[12] S. Karlin and J.L. McGregor, The differential equations of birth and death processes, and the Stieltjes moment problem. *Trans. Am. Math. Soc.* **85** (1957) 489–546.
[13] F.P. Kelly, Reversibility and stochastic networks. Cambridge University Press (2011).
[14] A.E. Koudou, Lancaster bivariate probability distributions with Poisson, negative binomial and gamma margins. *Test* **7** (1998) 95–110.
[15] H.O. Lancaster, The structure of bivariate distributions. *Ann. Math. Statist.* **29** (1958) 719–736.
[16] H.O. Lancaster, Correlations and canonical forms of bivariate distributions. *Ann. Math. Stati.* **34** (1963) 532–538.
[17] H.O. Lancaster, Joint probability distributions in the Meixner class. *J. Roy. Statist. Soc.* **37** (1975) 434–443.
[18] G. Letac, Lancaster probabilities and Gibbs sampling. *Stat. Sci.* **23** (2008) 187–191.
[19] T.M. Liggett, vol. 113 of *Continuous time Markov processes: an introduction.* American Mathematical Soc. (2010).

[20] E.B. McBride, *Obtaining generating functions*. Springer-Verlag (1971).
[21] J. Meixner, Orthogonale polynom system mit einer besonderen gestalt der erzeugenden funktion. *J. London Math. Soc.* **9** (1934) 6–13.
[22] R. Mena and S. Walker, On a construction of Markov models in continuous time. *METRON* **67** (2009) 303–323.
[23] E.D. Rainville, *Special functions*. Chelsea (1971).
[24] W. Schoutens, Vol. 146 of *Stochastic processes and orthogonal polynomials*. Springer Science & Business Media (2012).
[25] E.A. vann Doorn and A.I. Zeifman, Birth-death processes with killing. *Stat. Prob. Lett.* **72** (2005) 33–42.
[26] S. Walker, S. Hatjispyros and T. Nicoleris, A Fleming-Viot process and bayesian nonparametrics. *Ann. Appl. Prob.* **17** (2007) 67–80.