# Characterization of Markov semigroups on $\mathbb{R}$ Associated to Some Families of Orthogonal Polynomials

Dominique Bakry[1] and Olivier Mazet[2]

[1] Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France, bakry@cict.fr
[2] Laboratoire de Mathématiques appliquées de Lyon, INSA Lyon, Centre de Mathématiques, 20 avenue A. Einstein, 69621 Villeurbanne Cedex, France, olivier.mazet@insa-lyon.fr

**Summary.** We give a characterization of the eigenvalues of Markov operators which admit an orthogonal polynomial basis as eigenfunctions, in the Hermite and the Laguerre cases, as well as for the sequences of orthogonal polynomials associated to some probability measures on $\mathbb{N}$. In the Hermite case, we also give a description of the path of the associated Markov processes, as well as a geometric interpretation.

## 1 Introduction

The aim of this work is to describe all reversible Markov operators on $\mathbb{R}$, which have a spectral decomposition along some families of orthogonal polynomials. For any exponentially integrable probability measure on $\mathbb{R}$ (or on an interval), there is a natural family of orthogonal polynomials which forms a $L^2$-basis. The question arises to classify all Markov processes associated to this family, and more precisely we shall require this family of polynomials to be the spectral decomposition of some Markov operator, or of some generator of a Markov semigroup. There are many examples of this situation, where one may describe all possible eigenvalues. Among them the Hermite polynomials, the Laguerre polynomials, the Jacobi ones, and many other examples also in the discrete case, that is, when the underlying measure is carried by a discrete set. The question of whether or not there exists a Markov generator associated to a given family of orthogonal polynomials seems out of reach in such generality. The only known result in this direction is when a diffusion generator is associated to a family of orthogonal polynomials ([9]); in this case, the classification is quite simple, and the only classes of polynomials are the classical ones: the Hermite, Laguerre and Jacobi polynomials. (But there are many more examples, like Mexnier, Charlier and Hahn polynomials, which are not associated with diffusion operators.)

J. Azéma et al. (Eds.): LNM 1832, pp. 60–80, 2003.

The problem then arises in those cases to describe all Markov operators associated to these different families. The case of Jacobi processes was completely resolved in [8], through the use of the underlying associated hypergroup structure. In the cases of Laguerre and Hermite polynomials, this underlying structure fails to be valid, because the support of the reference measure is not compact. But in some sense, this makes the classification simpler, since there is a kind of degenerate hypergroup structure which is always valid, with the Dirac mass at infinity playing the rôle of identity.

More precisely, we consider a probability measure $\mu$ on $\mathbb{R}$, such that $\int \exp(\alpha|x|)\,\mu(\mathrm{d}x) < \infty$ for some $\alpha > 0$, and we assume that the measure $\mu$ is not supported by a finite set. Then we know that the set of polynomials is dense in $L^2(\mu)$, and therefore there is an $L^2$ basis made of a sequence $(P_n)$ of orthogonal polynomials, $P_n$ being of degree $n$; and this sequence is unique provided we assume the leading coefficient of $P_n$ to be positive and the polynomials to have norm 1 in $L^2(\mu)$.

A Markov operator, defined by a kernel of probability measures $K(x, \mathrm{d}y)$, is defined on all positive or bounded functions on $\mathbb{R}$ by

$$K(f)(x) = \int f(y)\,K(x, \mathrm{d}y).$$

It maps positive functions into positive functions, and is such that $K(1) = 1$. We are interested here in such Markov operators which are bounded on $L^2(\mu)$ and have the property that $K(P_n) = c_n P_n$, for some sequence $(c_n)$ of real numbers. This just means that the Markov operator $K$ is symmetric in $L^2(\mu)$ and has the family $(P_n)$ as spectral decomposition.

For simplicity, we shall call such a sequence a Markov sequence associated to the family $(P_n)$.

In the same way, we shall say that a sequence $(\lambda_k)$ is a Markov generator sequence if, for every $t > 0$, the sequence $(\mathrm{e}^{-\lambda_k t})$ is Markov. In this case, the family of Markov operators with eigenvalues $(\mathrm{e}^{-\lambda_k t})$ is a Markov semigoup, and the family $(\lambda_k)$ is the eigenvalues of its generator.

When $(c_n)$ is a Markov sequence, the operator $K$ is a symmetric operator on $L^2(\mu)$, is a contraction in $L^2(\mu)$, and therefore the sequence $(c_n)$ lies in $[-1, 1]$. If the sequence $(c_n)$ is square summable, then the operator $K$ is Hilbert–Schmidt, and

$$K(x, \mathrm{d}y) = \left(\sum_n c_n P_n(x) P_n(y)\right) \mu(\mathrm{d}y),$$

where the kernel $k(x, y) = \sum_n c_n P_n(x) P_n(y)$ is in $L^2(\mu \otimes \mu)$ and positive.

The purpose of this work is to provide a description of all possible sequences $(c_n)$ associated to Markov operators, in many different situations. For pedagogical reasons, we first give the classification in the case of the Jacobi polynomials, and then show how the method carries over to the non compact case for a quite general family of orthogonal polynomials.

## 2 The Jacobi Polynomials Case

Gasper ([8]) gave the complete classification of Markov sequences for the Jacobi polynomials (which depend on two parameters $\alpha$ and $\beta$), through the use of an hypergroup structure, that is, a proper convolution of measures. These Jacobi polynomials in the symmetric case ($\alpha = \beta$) are the ultraspherical polynomials (obtained, when $n$ is an integer, from projections of spheres).

Let us consider the probabilized interval $([-1,1], \mu)$, with

$$\mu^{\alpha,\beta}(\mathrm{d}x) = \mu(\mathrm{d}x) = C(1-x)^\alpha(1+x)^\beta\,\mathrm{d}x,$$

$\alpha > -1$, $\beta > -1$, $C$ being a normalization constant. The corresponding family of orthogonal polynomials is the family $(J_k^{\alpha,\beta})_{k\in\mathbb{N}}$ of Jacobi polynomials. They also may be defined via their generating series:

$$2^{\alpha+\beta}A^{-1/2}\big(1-t+A^{1/2}\big)^{-\alpha}\big(1+t+A^{1/2}\big)^{-\beta} = \sum_{k=0}^{+\infty} t^k h_k^{\alpha,\beta} J_k^{\alpha,\beta}(x),$$

where $A = A(x,t) = 1 - 2xt + t^2$ and

$$\big(h_k^{\alpha,\beta}\big)^2 = k!\,\frac{2k+\alpha+\beta+1}{\alpha+\beta+1}\,\frac{\Gamma(\alpha+1)}{\Gamma(k+\alpha+1)}\,\frac{\Gamma(\beta+1)}{\Gamma(k+\beta+1)}\,\frac{\Gamma(k+\alpha+\beta+1)}{\Gamma(\alpha+\beta+1)}\,.$$

(See [12], p. 69).

They satisfy the differential equation

$$\big(1-x^2\big)P_k'' + \big(\beta-\alpha-x(\alpha+\beta+2)\big)P_k' = -k(k+\alpha+\beta+1)P_k.$$

Thus they are the eigenvectors for the operator

$$L^{\alpha,\beta}(f) = \big(1-x^2\big)f'' + \big(\beta-\alpha-x(\alpha+\beta+2)\big)f', \tag{1}$$

with eigenvalues $\lambda_k = -k(k+\alpha+\beta+1)$. Therefore, for every $t > 0$, the sequence $c_k = \exp(-t\lambda_k)$ is a Markov sequence for this family.

Since they are the orthogonal polynomials associated with a measure supported by $[-1,1]$, it is not hard to see, by the usual property of interlacing of zeros, that the sequence $(P_k(1))$ is always positive. In fact, the maximum of $P_k$ on $[-1,1]$ is always attained at the point $x = 1$.

In [7] and [8], Gasper gave a complete representation of the Markov sequences related to the family of Jacobi polynomials, extending a result of Bochner ([3]) in the particular case $\alpha = \beta$, which is related to the ultraspherical polynomials:

**Proposition 1.** *Assume that $\alpha \geqslant \beta > -1$, with either $\beta \geqslant -1/2$, or $\alpha \geqslant -\beta$. Then, the sequence $(c_n)$ is Markov with respect to the family $(P_n)$ if and only if there is a probability measure $\nu$ on $[-1,1]$ such that*

$$c_k = \frac{1}{P_k(1)}\int_{-1}^{1} P_k(x)\,\nu(\mathrm{d}x).$$

*Proof.* For $\alpha > -1/2$ and $\alpha \geqslant \beta \geqslant 1/2$, the proof relies on the following important property: the series

$$K(x,y,z) = \sum_k P_k(x)P_k(y)P_k(z)/P_k(1)$$

is convergent (in $(L^2(\mu^{\otimes 3}))$ and the sum is positive.

By construction, this kernel is symmetric and has integral 1 with respect of any of its variables.

Therefore, there is an explicit representation

$$P_n(x)P_n(y) = P_n(1)\int P_n(z)K(x,y,z)\,\mu(\mathrm{d}z),$$

and

$$\int P_n(x)P_k(y)K(x,y,z)\,\mu(\mathrm{d}x)\,\mu(\mathrm{d}y) = \delta_{n,k}P_n(z)/P_n(1).$$

We then may define a convolution of probability measures as

$$\nu_1 * \nu_2(\mathrm{d}z) = \Big(\int_{x,y} K(x,y,z)\,\nu_1(\mathrm{d}x)\,\nu_2(\mathrm{d}y)\Big)\mu(\mathrm{d}z),$$

which is commutative and satisfy

$$\nu * \delta_1 = \nu,$$

as may be seen directly when $\nu$ has an $L^2(\mu)$ density $f$ with respect to $\mu$. The fact that the result is a probability measure is the fact that the integral of the kernel $K$ with respect to $z$ is 1.

We may as well define the convolution between a measure and an integrable function, identifying a function $f$ with the (non-necessarily positive) measure $f(x)\,\mu(\mathrm{d}x)$. We then have $\delta_1 * f = f$.

If $\nu$ is a bounded measure, then $\nu * P_n = c_n P_n$, where

$$c_n = \frac{1}{P_n(1)}\int P_n(y)\,\nu(\mathrm{d}y).$$

Now, if an operator $K$ is Markov and symmetric with respect to $\mu$, we may define as well the action of $K$ on probability measures, and we see that

$$K(\nu_1 * \nu_2) = K(\nu_1) * \nu_2 = \nu_1 * K(\nu_2).$$

To see this, we may restrict ourselves to the case where $\nu_1$ and $\nu_2$ have $L^2$ densities with respect to $\mu$; in that case this is immediate in the $L^2$-basis $(P_n)$, starting from $KP_n = c_n P_n$.

Therefore, the Markov kernel $K$ has the following representation

$$K(f) = K(\delta_1 * f) = K(\delta_1) * f,$$

and the probability measure $\nu = K(\delta_1)$ gives the representation. □

Of course, the basic tool used in the preceding proof, namely the positivity of the kernel $K(x,y,z)$, is indeed a quite deep result. Observe that the existence of this kernel amounts to saying that, for almost every $x$, (with respect to the reference measure), the sequence $(P_n(x)/P_n(1))$ is Markov. Since Markov sequences are bounded by 1, this implies that the maxima of the polynomials $(P_n)$ are attained at 1. In any case, the possibility to find Markov sequences of the form $c_nP_n(x)$ implies that the polynomials $(P_n)$ are bounded, and therefore that the support of the reference measure is compact.

In order to understand where this convolution comes from, we shall concentrate on the case where $\alpha = \beta$ is a half-integer. We shall see that this property is in fact a by-product of the spectral decomposition of the spherical Laplace operator on spheres.

First, consider the unit sphere $S_{n-1}$ in $\mathbb{R}^n$, with $n \geqslant 2$, equipped with its uniform probability measure $\sigma$. We may project this measure on the interval $[-1,1]$, which is identified with the diameter of the sphere carried by the first unit vector $e_1$ of $\mathbb{R}^n$. Then, this measure is $\mu^{\alpha,\alpha}$, with $\alpha = (n-3)/2$.

Moreover, if we take a smooth function $f$ on $[-1,1]$, and if we lift it on the sphere through the map just described here, say $F(x) = f \circ \Phi(x)$, then $\Delta F = L^{\alpha,\alpha}(f) \circ \Phi$ where $\Delta$ is the Laplace operator on the sphere. Therefore, the Jacobi polynomials (in this case the ultraspherical polynomials) may be lifted to eigenvectors of the Laplace operator on the sphere.

Now, in the same way, we may lift a measure on $[-1,1]$ into a measure on the sphere. We shall say that such a lifted measure is radial around $e_1$. If we take any rotation which maps $e_1$ onto another point $x$ of the sphere, the image of a radial measure is radial around $x$, and this image does not depend on the choice of the rotation (since it is radial). Therefore, for any point $x$ of the sphere, and any radial measure $\nu$ around $e_1$, we may define in a unique way a radial measure $R_x\nu$ around $x$. Now, to define the convolution of two probability measures $\nu_1$ and $\nu_2$, we may first lift those two measures on the sphere into radial measures $m_1$ and $m_2$, consider a random variable $X_1$ with law $m_1$, and construct a new random variable $Y$ such that its conditional law given $X_1$ is $R_{X_1}m_2$. It turns out that this new random variable has a law which is radial around $e_1$. Its projection on $[-1,1]$ is $\nu_1 * \nu_2$.

Now, if we want to take the convolution of two functions, then we see by construction that

$$f * g(x) = \int_{S_{n-1}(1)} f(x \cdot y) g(y \cdot e_1)\, \sigma(\mathrm{d}y), \tag{2}$$

where $x \cdot e$ denotes the scalar product in $\mathbb{R}^n$. From this we see that convolution is symmetric.

Now, for any $x \in S_{n-1}$, the function $y \mapsto P_k(x \cdot y)$ is an eigenvector of the Laplace operator of the sphere, with eigenvalue $-k(k+n-2)$ (we know the result by projection when $x = e_1$, and everything is invariant by rotation).

Moreover, it is quite easy to observe that

$$\int P_k(x \cdot y) P_l(y \cdot e)\, \sigma(\mathrm{d}y) = \delta_{k,l} P_k(1) P_k(x \cdot e).$$

Therefore, we get

$$P_k * P_l = \delta_{k,l} P_k(1) P_k,$$

and the convolution is the one we were looking for.

Although it is not necessary, the previous construction gives us a way to compute the kernel in this case, which is

$$K(t, r, s)\, \mu(\mathrm{d}s) = \delta_s * \delta_t,$$

from which we get

$$K_n(t, r, s) = \kappa_n \frac{(1 - r^2 - s^2 - t^2 + 2rst)^{(n-4)/2}}{(1 - s^2)^{(n-3)/2}(1 - r^2)^{(n-3)/2}(1 - t^2)^{(n-3)/2}} \times \mathbf{1}_{\{1-r^2-s^2-t^2+2rst \geqslant 0\}}. \quad (3)$$

## 3 The Non Compact Setting

Following the same scheme, we shall now investigate a number of examples of orthogonal polynomials associated with non compactly supported measures. In this situation, there is no valid hypergroup structure, but in some sense it makes things simpler. The reason is, in this situation the point 1, which is the point at which every polynomial in the Jacobi family achieves its maximum, is then pushed at infinity.

We shall restrict ourselves to the following situation.

The reference measure $\mu$ is not supported by any interval $(-\infty, M]$, is exponentially integrable as described in the introduction (that is, there exists some constant $\varepsilon > 0$ such that $\int \exp(\varepsilon|x|)\, \mu(\mathrm{d}x) < \infty$). We call $(P_n)$ the sequence of orthogonal polynomials associated to it, with leading coefficient $d_n > 0$. As before, we shall call a bounded sequence $(c_n)$ Markov if $c_0 = 1$ and if the linear operator $K$ defined on $L^2(\mu)$ by $K(P_n) = c_n P_n$ preserves positivity. We define in the same way a Markov generator sequence $(\lambda_k)$ by the fact that, for every $t > 0$, $(\exp(-\lambda_k t))$ is a Markov sequence.

Our basic assumption is the following:

There exists a Markov generator sequence $(\lambda_k)$ such that for every $t > 0$, $\sum_k \mathrm{e}^{-\lambda_k t} < \infty$.

In this case, we shall call $P_t$ the Markov operator associated with the sequence $(\exp(-\lambda_k t))$. Because of the summation hypothesis on $(\exp(-\lambda_k t))$, we know that $P_t$ is a Hilbert–Schmidt operator, and therefore may be represented as

$$P_t(f)(x) = \int K_t(x, y) f(y)\, \mu(\mathrm{d}y),$$

with

$$K_t(x,y) = \sum_k \exp(-\lambda_k t) P_k(x) P_k(y).$$

This series converges in $L^2(\mu \otimes \mu)$ and the sum is almost everywhere positive. Moreover, almost everywhere in the product,

$$K_t(x,y)^2 \leqslant K_{2t}(x,x) K_{2t}(y,y),$$

and the integral

$$\int_x K_t(x,x)\, \mu(\mathrm{d}x) = \sum_k \mathrm{e}^{-\lambda_k t} < \infty.$$

Since the function $K_t(x,x)$ is defined almost everywhere, we have to be a bit careful about convergence. We shall say that a function $f$ converges almost everywhere (for $\mu$) to 0 at infinity if the sequence of functions $f_N = f\mathbf{1}_{[N,\infty)}$ converges almost surely to 0 when $N$ goes to infinity. In most cases, this will be irrelevant since we shall be able to find a "good" version of $K_t(x,x)$ (say continuous), and in the discrete case the problem simply does not arise.

Under an extra technical condition, we have a simple characterization of Markov sequences:

**Theorem 1.** *Suppose that* $\forall t > 0,\ \forall k \in \mathbb{N}$, *the function*

$$H_{t,k}(x) = \frac{\sqrt{K_t(x,x)}}{x^k} \int_{|y|>x} y^k \sqrt{K_t(y,y)}\, \mu(\mathrm{d}y) \tag{4}$$

*converges to* 0 *almost everywhere when* $x$ *goes to infinity. Then, for every Markov sequence* $(c_k)$ *associated to* $(P_k)$, *there exists a probability measure* $\nu$ *on* $[-1,1]$ *such that*

$$\forall k \in \mathbb{N}, \qquad c_k = \int x^k\, \nu(\mathrm{d}x).$$

*Moreover, if the measure is carried by* $\mathbb{R}_+$, *then the measure* $\nu$ *may be chosen with support* $[0,1]$.

*Proof.* For a Markov sequence $(c_k)$, we define for every $t > 0$ the Markov kernel

$$K_t^c(x,y) = \sum_k \exp(-\lambda_k t) c_k P_k(x) P_k(y). \tag{5}$$

This is again a positive kernel, which is square integrable with respect to the product measure. (Recall that the sequence $(c_k)$ lies in $[-1,1]$.) Since the kernel $K_t^c$ corresponds to a positive operator, for almost every $x$ the measure $K_t^c(x,y)\, \mu(\mathrm{d}y)$ is a positive measure.

On the other hand, we know that

$$K_t^c(x,y)^2 \leqslant K_{2t}^c(x,x) K_{2t}^c(y,y) \leqslant K_{2t}(x,x) K_{2t}(y,y),$$

almost everywhere in the product.

Therefore, there exists a sequence $(x_n)$ going to infinity such that the measures $\nu_{n,t}(\mathrm{d}y) = K_t^c(x_n, y)\,\mu(\mathrm{d}y)$ are probability measures, such that the previous inequality holds for any $(x_n)$ almost everywhere in $y$, and such that the sequence $(H_{t,k}(x_n))$ converges to 0, where $H_{t,k}$ is the function defined in 4. (This is the only place where we use the fact that the support of $\mu$ is not compact.)

We consider then the measures $\mu_{n,t}$ on $[-1,1]$ obtained in the following way: we restrict $\nu_{n,t}$ on $[-x_n, x_n]$, and take the image under the map $x \mapsto x/x_n$.

Then, we shall show that the sequence $(\nu_{n,t})$ converges to a measure $\nu$ whose moments are $(c_n)$. The limit is taken first in $n \to \infty$, then in $t \to 0$.

Since the measures $\nu_{n,t}$ are supported by the compact interval $[-1,1]$, to prove this convergence it suffices to show that the moments of the measures $\nu_{n,t}$ converge, and that the total mass converges to 1.

In fact, we shall prove that, for every $k \in \mathbb{N}$,

$$\lim_{n,t} \int x^k\, \nu_{n,t}(\mathrm{d}x) = c_k,$$

and this will complete the proof, since the set of probability measures on $[-1,1]$ is compact for the weak convergence.

First we observe that, $k$ being fixed, $P_k(xx_n)/(d_k x_n^k)$ converges uniformly on $[-1,1]$ to $x^k$ when $n$ goes to infinity, and therefore it suffices to check that

$$\lim_{n,t} \frac{1}{d_k x_n^k} \int P_k(xx_n)\, \nu_{n,t}(\mathrm{d}x) = c_k.$$

Then, by definition of $\nu_{n,t}$, we write the last integral as

$$\begin{aligned} &\frac{1}{d_k x_n^k} \int_{|y|\leqslant x_n} P_k(y) K_t^c(x_n, y)\,\mu(\mathrm{d}y) \\ &\qquad = \frac{1}{d_k x_n^k}\left(\int_{\mathbb{R}} P_k(y) K_t^c(x_n, y)\,\mu(\mathrm{d}y) - \int_{|y|\geqslant c_n} P_k(y) K_t^c(x_n, y)\,\mu(\mathrm{d}y)\right). \end{aligned}$$

The first integral is nothing else than $c_k \exp(-\lambda_k t) P_k(x_n)/(d_k x_n^k)$, whose limit in $n$ is $c_k \exp(-\lambda_k t)$, and we then take the limit in $t \to 0$.

It remains to show that the second integral goes to 0. But then we use $K_t^c(x,y)^2 \leqslant K_{2t}(x,x)K_{2t}(y,y)$ (remember that the sequence $(c_k^{2)}$ is bounded by 1), and $|P_k(x)| \leqslant C_k |x^k|$, on $|y| \geqslant x_n$, for $n$ large enough. The result then follows from the assumption. □

Applying Theorem 1 requires some knowledge about the functions $K_t(x,x)$, which is not always easy to obtain. We shall derive below another version, which avoids this difficulty.

**Theorem 2.** *Assume that the measure $\mu$ has a density $\rho$ with respect to the Lebesgue measure, and that, for any $k$, there exists a constant $C_k$ such that, for $x$ large enough,*

$$\int_{|y|\geqslant x} \frac{|y|^k}{x^k}\,\mu(\mathrm{d}x) \leqslant C_k\rho(x).$$

*Then, the same conclusion holds.*

*Proof.* In the previous theorem, we considered the measure $\mu_{n,t} = K_t^c(x_n, \mathrm{d}y)$, which we truncated on $[-x_n, x_n]$ and then concentrate on $[-1,1]$ by dilation. Here, we shall apply the same procedure to the measure

$$\mu_{n,t}(\mathrm{d}y) = \int_{x\in[n-1,n]} K_t^c(x,\mathrm{d}y)\,\mathrm{d}x,$$

which we truncate on $[-n,n]$ and carry onto $[-1,1]$.

The same proof works without any change, and we are led to prove that, for any $t$ and any $k$,

$$\int_{|y|\geqslant n} \frac{|y|^k}{n^k}\,\mu_{n,t}(\mathrm{d}y) \longrightarrow 0, \qquad (n\to\infty).$$

We majorize again

$$|K_t^c(x,y)| \leqslant \sqrt{K_{2t}(x,x)}\,\sqrt{K_{2t}(y,y)},$$

and the latter expression is bounded by

$$\int_{[n-1,n]} \sqrt{K_{2t}(x,x)}\left(\int_{|y|\geqslant|x|} \frac{|y|^k}{n^k}\sqrt{K_{2t}(y,y)}\,\mu(\mathrm{d}y)\right)\mathrm{d}x.$$

We know that

$$\int K_{2t}(y,y)\,\mu(\mathrm{d}y) < \infty.$$

We may use Schwarz' inequality and the hypothesis (with $k$ replaced by $2k$), and we are led to prove that

$$\int_{[n-1,n]} \sqrt{K_{2t}(x,x)\rho(x)}\,\mathrm{d}x \longrightarrow 0 \qquad (n\to\infty).$$

But we then use Schwarz's inequality again and we know that

$$\int K_{2t}(x,x)\rho(x)\,\mathrm{d}x < \infty,$$

and therefore that

$$\int_{[n-1,n]} K_{2t}(x,x)\rho(x)\,\mathrm{d}x \longrightarrow 0 \qquad (n\to\infty),$$

which shows that the previous sequence goes to 0. $\square$

From the representation theorem for Markov sequences, it is easy to deduce the representation theorem for Markov generators sequences.

**Proposition 2.** *Under the conditions of theorems 1 or 2, if* $(\lambda_k)$ *is a Markov generator sequence associated with* $(P_n)$*, there exist two non negative constants* $\theta$ *and* $c$ *and a probability measure* $\nu$ *on* $(-1,1)$ *such that*

$$\lambda_k = \theta k + c\int_{-1}^{1} \frac{1-s^k}{1-s}\,\nu(\mathrm{d}s).$$

*Proof.* This Lévy–Khintchine representation theorem is straightforward once we have the representation theorem for the Markov sequences. Assume that $(\lambda_k)$ is a Markov generator sequence. Then, for any $t > 0$, there exists a probability measure $\mu_t$ on $[-1,1]$ such that

$$\exp(-\lambda_k t) = \int x^k \mu_t(\mathrm{d}x).$$

If we define the convolution of two measures on $[-1,1]$ by

$$\mu * \nu(f) = \int f(xy)\,\mu(\mathrm{d}x)\,\nu(\mathrm{d}y),$$

we see that $(\mu_t)_{t\geqslant 0}$ is a convolution semigroup for this structure.

Therefore, the result comes from classical results of harmonic analysis on groups (see [2], for example).

Nevertheless, for the sake of completeness and since the arguments are really easy to obtain in this case, we cannot resist to briefly sketch the proof.

Let $(\mu_t^\lambda)_{t\geqslant 0}$ be the convolution semigroup associated to the Markov generator sequence $\lambda$. First, we remark that the set $\mathcal{L}$ of Markov generator sequences is a convex cone (with $\mu_{at}^\lambda * \mu_{bt}^{\lambda'}$ associated with the sequence $a\lambda + b\lambda'$). We endow it with the topology of pointwise convergence, which corresponds to the narrow convergence of the associated measures $\mu_t^\lambda$. Remark that each $\lambda$ in $\mathcal{L}$ satisfies $\lambda_0 = 0$. Then we observe that, by Jensen's inequality,

$$\lambda(2k) \leqslant 2k\lambda(1),$$

and also that

$$\int_{-1}^{1} (x^{2k} - x^{2k+1})\,\mu_t(\mathrm{d}x) \leqslant \int (1-x)\,\mu_t(\mathrm{d}x),$$

from which we get, at $t = 0$, that

$$\lambda(2k+1) \leqslant \lambda(2k) + \lambda(1) \leqslant (2k+1)\lambda(1).$$

Therefore, for any $k$, $\lambda(k) \leqslant k\lambda(1)$, and the cone $\mathcal{L}$ has compact basis.

Observe also that, if there is an even $h$ for which $\lambda(h) = 0$, then the measure $\mu_t^\lambda$ is supported by $\{-1,1\}$, and therefore all even $h$ satisfy $\lambda(h) = 0$

and $\lambda(2k+1) = \lambda(1)$. The representation is given with any probability measure $\nu = \alpha\delta_1 + (1-\alpha)\delta_{-1}$, $\lambda(1) = c(1-\alpha)$ and $\theta = c\alpha$.

So we may suppose that all even $\lambda(h)$ are non zero.

In this case, the trick is to show that, if $h$ is even, then the sequences (in $k$) $\lambda_1(h,k) = \lambda(k+h) - \lambda(h)$ and $\lambda_2(h,k) = \lambda(k) + \lambda(h) - \lambda(k+h)$ are again in $\mathcal{L}$.

The sequence $\lambda_1$ is associated to the semigroup $(\mathrm{e}^{t\lambda(h)} x^h \mu_t^\lambda(\mathrm{d}x))_{t\geqslant 0}$. For the sequence $\lambda_2$, we observe that the sequence

$$\nu_n(\mathrm{d}x) = \frac{1-x^h}{1-\exp\bigl(-\lambda(h)/n\bigr)}\,\mu^\lambda_{1/n}(\mathrm{d}x)$$

has a weak limit (considering its moments) which is associated to the Markov sequence $\lambda_2(h,k)/\lambda(h)$.

Also, if $\lambda$ is a Markov sequence, associated to some measure $\mu$, then $1-\lambda$ is in $\mathcal{L}$, associated to

$$\mu_t = \mathrm{e}^{-t} \sum_k \frac{t^k}{k!}\,\mu^{*k}.$$

If we apply this to the Markov sequence just obtained and multiply by $\lambda(h)$ we get $\lambda_2$.

Now, if $\lambda$ is an extremal element of the cone $\mathcal{L}$, then $\lambda_2$ and $\lambda_1$ are proportional to $\lambda$, we get

$$\lambda(k+h) = \lambda(h) + c(h)\lambda(k),$$

for all even $h$ and any $k$. Applying that with $h+h'$, and comparing, we get $c(h)c(h') = c(h+h')$, and therefore $c(h) = x^h$, for some $x \in (0,1]$.

The case $x = 1$ leads us to $\lambda(2p) = p\lambda(2)$ and $\lambda(2p+1) = p\lambda(2) + \lambda(1)$, and the case $x \neq 1$ gives

$$\lambda(k) = \frac{\lambda(2)}{1-x^2}(1-x^k), \qquad (k \text{ even}),$$

and

$$\lambda(k) = \frac{\lambda(2)}{1-x^2}(1-x^{k-1}) + \lambda(1)x^{k-1}, \qquad (k \text{ odd}).$$

Therefore, the extremal elements of $\mathcal{L}$ are included in the following function set:

$$\begin{cases} \lambda : \begin{array}{rcl} 2p & \longmapsto & 2p\theta \\ 2p+1 & \longmapsto & (2p+1)\theta + \gamma \end{array} \\ \\ \lambda : \begin{array}{rcl} 2p & \longmapsto & \alpha(1-x^{2p}) \\ 2p+1 & \longmapsto & \alpha(1-x^{2p}) + \beta x^{2p} \end{array} \end{cases}$$

with $\theta = \lambda(2)/2$, $\gamma = \lambda(1) - \lambda(2)/2$, $\alpha = \lambda(2)/(1-x^2)$ et $\beta = \lambda(1)$.

The first family corresponds to

$$\lambda(k) = \theta k + \int_{-1}^{1} \frac{1-x^k}{1-x} \, (\gamma\delta_{-1})(\mathrm{d}x).$$

For the second, it may be written as a combination of

$$\lambda(k) = \frac{1-x^k}{1-x} \qquad \text{and} \qquad \lambda'(k) = \frac{1-(-x)^k}{1+x},$$

which corresponds to the measures $\nu = \delta_x$ and $\nu = \delta_{-x}$ respectively.

Last, we use the Choquet representation theorem to get the general form of the representation. □

## 4 The Hermite Polynomials Case

We shall now illustrate our main theorems on some examples. Let us begin with the family of Hermite polynomials.

They are the orthonormal family associated with the Gaussian measure

$$\gamma(\mathrm{d}x) = \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-x^2/2} \, \mathrm{d}x.$$

The Hermite polynomials $(H_k(x))_{k\in\mathbb{N}}$ are defined by their generating series:

$$\forall t \in \mathbb{R}, \quad \forall x \in \mathbb{R}, \qquad \sum_{k\in\mathbb{N}} \frac{t^k}{\sqrt{k\,!}} \, H_k(x) = \mathrm{e}^{tx-t^2/2}.$$

From their generating series, it is not hard to deduce the following property: if $L(f)(x) = f''(x) - xf'(x)$, then

$$L(H_n) = -nH_n.$$

But the operator $L$ is the generator of a diffusion semigroup, namely the Ornstein–Uhlenbeck semigroup, which may be defined in the following way:

$$P_t(f)(x) = \int f\Big(\mathrm{e}^{-t}x + \sqrt{1-\mathrm{e}^{-2t}}y\Big) \, \gamma(\mathrm{d}y).$$

Once again, starting from the generating series, it is easy to check that

$$P_t(H_n) = \exp(-nt)H_n.$$

Therefore, the sequence $\lambda_n = -n$ is a Markov generator sequence for the family $(H_n)$, and this semigroup $(P_t)_{t\geqslant 0}$ will be used for the semigroup $(K_t)_{t\geqslant 0}$ described in theorem 1.

We get from the previous section the following result, due to [10].

**Proposition 3.** *The sequence of real numbers $(c_k)$ is Markov with respect to $(H_n)$ if and only if there exists a probability measure $\mu$ on $[-1,1]$ such that*

$$c_k = \int_{-1}^{1} x^k \, \mu(\mathrm{d}x).$$

*Proof.* We first check that the conditions of the theorem apply.

From the definition of $P_t$, we can see that its kernel

$$p_t(x,y) = \sum_k \exp(-kt) H_k(x) H_k(y)$$

may be written as

$$p_t(x,y) = \big(2\pi(1-\rho^2)\big)^{-1/2} \, \mathrm{e}^{-\frac{(y-x\rho)^2}{2(1-\rho^2)} + \frac{y^2}{2}},$$

where $\rho = \mathrm{e}^{-t}$.

From this, by a change of variables, we see that

$$\int_{|y| \geqslant x} \frac{y^k}{x^k} \sqrt{p_t(y,y)} \, \gamma(\mathrm{d}y) \leqslant c(k,t) \, \mathrm{e}^{-\frac{x^2}{2}\frac{1}{1+\rho}},$$

while

$$\sqrt{p_t(x,x)} = c'(t) \, \mathrm{e}^{\frac{x^2}{2}\frac{\rho}{1+\rho}}.$$

Therefore, the product of these two quantities converges to 0 when $x$ goes to $\infty$.

It remains to show that all the moments of measures are effectively Markov sequences, and, by convexity, it is enough to show that, for any $x \in [-1,1]$, the sequence $(x^k)$ is a Markov sequence.

If $0 < x < 1$, then the sequence $(\exp(-kt))$ answers the question (and corresponds to the Markov generator sequence described before).

The case $x = 0$ corresponds to the projection onto the constant functions (the integration with respect to $\gamma$, which is always a Markov sequence whatever the model).

The case $x = 1$ corresponds to the identity operator (same remark).

We have to show that the same remains true for $-x$, where $x \in [0,1]$. But the product of two Markov sequences is always a Markov sequence, and therefore we only have to show that the sequence $((-1)^k)$ is Markov.

This corresponds to the operator $K(f)(x) = f(-x)$. In fact, for the Hermite polynomials $(H_n)$, we have $H_n(-x) = (-1)^n H_n(x)$, a property which reflects the symmetry around 0 of the Gaussian measure. □

It is interesting to notice that the convolution structure associated to the Markov sequence for Hermite polynomials is inherited from the hypergroup structure for the ultraspherical polynomials.

Recall that the ultraspherical measure (which corresponds to the case $\alpha = \beta$ of Jacobi polynomials) is

$$\gamma_n(\mathrm{d}x) = c_n\big(1 - x^2\big)^{n/2-1}\,\mathbf{1}_{(-1,1)}\,\mathrm{d}x,$$

which for $n \in \mathbb{N}$ corresponds to the projection of the sphere onto a diameter. If we scale this measure by $\sqrt{n}$ and let $n \to \infty$, it is quite clear that this sequence of measures converges to the Gaussian measure (the celebrated Poincaré limit). Now, the whole structure of orthogonal polynomials converges in the same way.

If we look closely at the hypergroup convolution described in section 2, call it $*_n$, a simple exercise shows that

$$\nu *_n \mu(f) \longrightarrow \int f(xy)\,\mu(\mathrm{d}x)\,\nu(\mathrm{d}y) \qquad (n \to \infty).$$

As we just saw, there are in this case two extremal Markov generators. The Markov generator associated to the Ornstein–Uhlenbeck process, and the generator associated to the "sign" process, corresponding to a measure $\nu$ which is a Dirac mass at $-1$. To be more precise, we shall notice that we may always construct any Markov semigroup by subordination to the two corresponding semigroups.

To clarify the ideas, and to avoid complications, we just consider the case where in the Lévy–Khintchine representation 2 the measure $\mathrm{d}\hat{\nu} = c\,\mathrm{d}\nu/(1-x)$ on $[-1,1]$ is a probability measure, and where $\theta = 0$. Then, we can easily construct a Markov process with generator sequence given by the associated family of $\lambda_k$ with the help of a Ornstein–Uhlenbeck process and an independent Markov process on $[-1,1]$ with generator $\nu$. To do that, we just consider a sequence of independent random variables $(Y_n)$ on $[-1,1]$ with common law $\hat{\nu}$, and an independent Poisson process $(N_t)_{t\geqslant 0}$ with intensity 1 on the integers. Then, we set

$$M_t = \prod_0^{N_t} Y_i.$$

This defines a Markov process $(M_t)_{t\geqslant 0}$ on $[-1,1]$ with semigroup $(\mu_t)_{t\geqslant 0}$. Then, we set $T_t = -\log(|M_t|)$, and $\varepsilon_t = \operatorname{sign}(M_t)$, where $\operatorname{sign}(x) = \mathbf{1}_{x\geqslant 0} - \mathbf{1}_{x<0}$.

Let $(X_t)_{t\geqslant 0}$ be an independent Ornstein–Uhlenbeck process. Then the process $(Y_t)_{t\geqslant 0}$ defined by $Y_t = \varepsilon_t X_{T_t}$ is a Markov process, and its generator is given by the Markov generator sequence $(\lambda_k)$ associated to $\nu$. To see this, it is enough to show that, for any $k$,

$$E_x[H_k(\varepsilon_t X_{T_t})] = \mathrm{e}^{-\lambda_k t} H_k(x).$$

We distinguish two cases.

*The integer $k$ is even:* $H_k(x)$ is then an even polynomial, and $H_k(\varepsilon_t x) = H_k(x)$. Let $\rho_t$ be the law of $T_t$. Using independence, we have

$$\begin{aligned}
E_x[H_k(\varepsilon_t X_{T_t})] &= E_x[H_k(X_{T_t})] \\
&= \int_{\mathbb{R}} E_x[H_k(X_s)]\,\rho_t(\mathrm{d}s) \\
&= \int_{\mathbb{R}} \mathrm{e}^{-ks}\,\rho_t(\mathrm{d}s) H_k(x) \\
&= E\big[|M_t|^k\big] H_k(x) \\
&= E\big[M_t^k\big] H_k(x) \\
&= \mathrm{e}^{-\lambda_k(t)} H_k(x).
\end{aligned}$$

*The integer $k$ is odd:* $H_k(x)$ is then an odd polynomial, and $H_k(\varepsilon_t x) = \varepsilon_t H_k(x)$. Then, $\rho_t(\mathrm{d}\varepsilon, \mathrm{d}s)$ being the joint law of $(\varepsilon_t, T_t)$

$$\begin{aligned}
E_x[H_k(\varepsilon_t X_{T_t})] &= E_x[\varepsilon_t H_k(X_{T_t})] \\
&= \int_{\mathbb{R}\times\{-1,1\}} \varepsilon E_x[H_k(X_s)]\,\rho_t(\mathrm{d}\varepsilon, \mathrm{d}s), \\
&= E\big[\varepsilon_t |M_t|^k\big] H_k(x) \\
&= E\big[M_t^k\big] H_k(x) \\
&= \mathrm{e}^{-\lambda_k(t)} H_k(x).
\end{aligned}$$

Of course, when the total mass of the measure $\nu$ is finite and not 1, we may as well use a Poisson process with intensity $\lambda \neq 1$ to represent the semigroup.

When $\theta \neq 0$, we simply replace the process $(X_{T_t})_{t\geqslant 0}$ by $(X_{\theta t+T_t})_{t\geqslant 0}$ to get the same representation.

Thus, we always get a Bochner representation for the Markov semigroups associated to $(H_n)$ of the form

$$K_t(f) = \int_{\mathbb{R}_+\times\{-1,1\}} P_s\big(f(\varepsilon x)\big)\,\rho_t(\mathrm{d}\varepsilon, \mathrm{d}s),$$

with $(\rho_t)_{t\geqslant 0}$ a convolution semigroup on $\mathbb{R}_+ \times \{-1, 1\}$.

Back to the Poincaré limit, it would be nice if we could interpret those processes as limits of jump processes on spheres. To make things simpler, we only consider the case where $\theta = 0$, $c = 1$, $\nu(\mathrm{d}s) = \delta_l(\mathrm{d}s)$.

Let $(X_t)_{t\geqslant 0}$ be a jump process on $S_{n-1}(1)$, whose jump times have an exponential distribution law of parameter 1, and whose each jump amplitude has a distribution law, which is invariant under any rotation around the point from where the jump began, that is, we can write: $X_t = Y_{N_t}$, where $(N_t)_{t\geqslant 0}$ is a Poisson process of parameter 1, and $(Y_n)_{n\in\mathbb{N}}$ an homogeneous Markov chain, whose transition kernel is $P\{Y_n \in \mathrm{d}y \,|\, Y_{n-1} = x\} = f(x \cdot y)\,\sigma(\mathrm{d}y)$, where $x$ and $y$ are some points of $S_{n-1}(1)$, and where $f$ is a function on $S_{n-1}(1)$.

Assume that $f$ is a function of $L^2(S_{n-1}(1), m)$, it then satisfies the following expression, in the ultraspherical polynomial:

$$f(x \cdot y) = \sum_{k \in \mathbb{N}} f_k P_k^n(x \cdot y).$$

Let $(P_t)_{t \geqslant 0}$ be the semigroup associated to the process $(X_t)_{t \geqslant 0}$, and let $g$ be an element of $L^2(S_{n-1}(1), \sigma)$:

$$P_t g(x) = g(x)\,\mathrm{e}^{-t} + \left(1 - \mathrm{e}^{-t}\right) \int_{S_{n-1}(1)} f(x \cdot y) g(y)\, \sigma(\mathrm{d}y) + o(t^2),$$

hence if we denote by $L$ the associated infinitesimal generator, and if we assume that $g$ lies in the domain of $L$:

$$Lg(x) = -g(x) + \int_{S_{n-1}(1)} g(y) f(x \cdot y)\, \sigma(\mathrm{d}y).$$

Let us focus on the action of the operator $L$ on the radial functions. Let $g$ be a radial function of $L^2(S_{n-1}(1), m)$, it then satisfies the following expression:

$$g(x) = \sum_{k \in \mathbb{N}} g_k P_k^n(x \cdot e).$$

Then we have:

$$\begin{aligned} Lg(x) &= \int_{S_{n-1}(1)} \sum_{k \in \mathbb{N}} f_k g_k P_k^n(x \cdot y) P_k^n(y \cdot e)\, \sigma(\mathrm{d}y) - g(x) \\ &= \sum_{k \in \mathbb{N}} f_k g_k \int_{S_{n-1}(1)} P_k^n(x \cdot y) P_k^n(y \cdot e)\, \sigma(\mathrm{d}y) - g(x). \end{aligned}$$

Now, we already observed that

$$\int P_k^n(x.y) P_k^n(z \cdot y)\, \sigma(\mathrm{d}y) = \frac{P_k^n(x \cdot z)}{P_k^n(1)}.$$

Then, the Markov generator sequence associated to $L$ is $(f_k / P_k^n(1) - 1)$. When the measure $\nu$ is $\delta_l$, we simply get $(P_k^n(l)/P_k^n(1) - 1)$.

Now, some analysis on the polynomials $(P_k^n)$ shows that

$$\lim_{n \to +\infty} \frac{P_k^n(l)}{P_k^n(1)} = l^k.$$

It is easy to understand if we recall that

$$P_k^n(l) \sim H_k(l\sqrt{n}),$$

and on the other hand that

$$\lim_{n\to+\infty} \frac{H_k(l\sqrt{n})}{H_k(\sqrt{n})} = l^k.$$

So that we deduce the following

$$\lambda_k = \lim_{n\to+\infty} \lambda_k^n = 1 - l^k.$$

The Ornstein–Uhlenbeck process can be thought of as a limit, when $l \to 1$, of such limits of jump processes on the sphere.

## 5 The Laguerre Polynomials Case

The Laguerre polynomials are the family of orthogonal polynomials related to the measure on $\mathbb{R}_+$

$$\mu_\alpha(\mathrm{d}x) = K_d\,\mathrm{e}^{-x} x^\alpha\,\mathrm{d}x \qquad (\alpha > -1).$$

When $\alpha = d/2 - 1$, with $d \in \mathbb{N}$, this measure is obtained from the Gaussian measure in $\mathbb{R}^d$ by taking its image under $x \mapsto |x|^2/2$. In this case, if $n$ is even and if $H_n$ is the relevant Hermite polynomial, then $\sum_1^d H_n(x_i)$ is indeed a polynomial in $|x|^2/2$ of degree $n/2$, and this is exactly (up to a constant) the Laguerre polynomial of degree $n/2$. This explains the strong connections between Laguerre and Hermite polynomials.

We have the generating function

$$F(t,x) = (1-t)^{-a-1}\mathrm{e}^{-xt/(1-t)} = \sum_k t^k c_{k,\alpha} L_k(x),$$

for the sequence of orthonormal polynomials $(L_n)$ associated to this measure. (The values of $c(n,\alpha)$ may be computed by taking the integral of $F(t,x)F(s,x)$ and identifying the series.) From this generating series, we deduce that, if $L(f)(x) = xf''(x) + (\alpha - x)f'(x)$, then

$$LP_k = -kP_k,$$

and therefore, since $L$ is the generator of a diffusion semigroup on $\mathbb{R}_+$, the sequence $(\exp(-kt))$ is a Markov sequence.

In this situation, it is not completely straightforward to get an explicit upperbound on the kernel $K_t(x,x)$. But a simple computation gives

$$\int_{y\geqslant x} \frac{y^k}{x^k}\,\mu_\alpha(\mathrm{d}x) \leqslant C_k \mathrm{e}^{-x} x^\alpha.$$

Therefore, the second criterion 2 applies and we get

**Theorem 3.** *A sequence $(c_k)$ is Markov with respect to the family of Laguerre polynomials if and only if there exists a probability measure $\nu$ on $[0,1]$ such that*

$$c_k = \int x^k \, \nu(\mathrm{d}x).$$

*A sequence $(\lambda_k)$ is a Markov generator sequence with respect to the Laguerre polynomials if and only if there exists a probability measure $\nu$ on $[0,1)$, and two non negative reals $\theta$ and $c$ such that*

$$\lambda_k = \theta k + c \int_0^1 \frac{1-s^k}{1-s} \, \nu(\mathrm{d}s).$$

*Proof.* The second assertion is a direct consequence of the first one. For the first one, the only thing to prove is that any probability measure $\nu$ provides a Markov sequence, and by convexity, that for any $0 \in [0,1]$, $(x^k)$ is a Markov sequence. We already know the result for $x = 0$ and $x = 1$, and we also know that, for any $t > 0$, the sequence $(\mathrm{e}^{-\lambda_k t})$ is Markov. The proof is completed. □

Once again, in this case, the Lévy–Khintchine formula gives a Bochner representation of any semigroup with respect to the diffusion semigroup associated to the generator

$$L(f)(x) = xf''(x) + (\alpha - x)f'(x).$$

The situation is even simpler than the Hermite case, since we do not need to use the "sign" semigroup.

## 6 Discrete Measures

In this section we briefly investigate some examples of measures on the set $\mathbb{N}$ of integers.

The first remark is that, provided there exists a Markov generator sequence $(\lambda_k)$ such that $\sum \mathrm{e}^{-\lambda_k t} < \infty$ for any $t > 0$, the criterion of theorem 1 is always satisfied for measures with exponential decay. Indeed, we have

**Proposition 4.** *Assume that $(\lambda_k)$ is a Markov generator sequence such that*

$$t > 0 \implies \sum_k \mathrm{e}^{-\lambda_k t} < \infty,$$

*and let $K_t(x,y)$ be the associated kernel. If*

$$\limsup_{n\to\infty} \frac{\mu(n+1)}{\mu(n)} < 1,$$

*then, for any $k \in \mathbb{N}$, for any $t > 0$,*

$$\lim_{n\to\infty} \frac{\sqrt{K_t(n,n)}}{n^k} \sum_{p>n} p^k \sqrt{K_t(p,p)}\mu(p) = 0.$$

*Proof.* We just have to follow the lines of the proof of theorem 2.

We write

$$\sum_{p>n} p^k \sqrt{K_t(p,p)}\,\mu(p) \leqslant \left(\sum_{p\geqslant 0} K_t(p,p)\mu(p)\right)^{1/2} \left(\sum_{p>n} p^{2k}\mu(p)\right)^{1/2}.$$

For $n$ large enough, the hypothesis shows that,

$$p \geqslant n \Longrightarrow \mu(p) \leqslant c^{p-n}\mu(n),$$

for some $c < 1$.

Therefore,

$$\sum_{p>n} p^{2k}\mu(p) \leqslant C_k n^{2k}\mu(n),$$

and it remains to observe that $K_t(n,n)\mu(n)$ goes to 0 when $n$ goes to infinity, which comes from the summability of the series. □

We may then apply this result to different classical families.

### 6.1 The Charlier Polynomials

They are the polynomials associated to the Poisson measure

$$\mu_a(n) = \exp(-a)\,\frac{a^n}{n!} \qquad (a > 0).$$

The generating series is

$$\mathrm{e}^{-t}\left(1 + \frac{t}{a}\right)^x = \sum \left(\frac{t}{\sqrt{a}}\right)^n \frac{P_n(x)}{\sqrt{n!}},$$

(see [12]), from which it is not hard to deduce that, if

$$L_a(f)(k) = f(k+1) + \frac{k}{a} f(k-1) - \left(\frac{k}{a} + 1\right) f(k),$$

then

$$L_a P_n = -\frac{n}{a}\,P_n.$$

Therefore, the Charlier polynomials are the eigenvectors of the finite difference operator $L_a$, which is the generator of a Markov semigroup $K_t$. The sequence $(\lambda_n) = (n/a)$ is therefore a Markov generator sequence, and the result applies. We get

**Proposition 5.** *A sequence* $(c_k)$ *is Markov for the Charlier polynomials if and only if there exists a probability measure* $\mu$ *on* $[0,1]$ *such that*

$$c_k = \int x^k \, \mu(\mathrm{d}x).$$

*A sequence* $(\lambda_k)$ *is a Markov generator sequence for the Charlier polynomials if and only if there exists a probability measure* $\nu$ *on* $[0,1)$ *and two non negative real parameters* $\theta$ *and* $c$ *such that*

$$\lambda_k = \theta k + c \int \frac{1-s^k}{1-s} \, \nu(\mathrm{d}s).$$

*Every Markov semigroup is Bochner subordinated to the Markov semigroup with generator* $L_a$*.*

### 6.2 The Meixner Polynomials

They are the polynomials associated with the measure

$$\mu(n) = \frac{\Gamma(b+n)}{\Gamma(b)} \frac{c^n}{\Gamma(n+1)}, \qquad (b > 0, \; 0 < c < 1).$$

Notice that for $b = 1$ we get the geometric distribution with parameter $c$. They satisfy

$$LP_n = -n(1-c)P_n,$$

where the finite difference operator $L$ may be defined as

$$L(f)(n) = n\Delta(f)(n) + \big(bc - (1-c)n\big)D(f)(n),$$

with $D(f)(n) = f(n+1) - f(n)$ and $\Delta(f)(n) = f(n+1) + f(n-1) - 2f(n)$. (A nice account of the properties of classical orthogonal polynomials associated with discrete measures may be found in [1].)

Once again, the sequence $(\lambda_k = (1-c)k)$ is a Markov generator sequence and we get the same result as in proposition 5, replacing everywhere Charlier polynomials by Meixner polynomials.

## References

1. J. Arvesù, J. Coussement, and W. Van Assche. Some discrete multiple orthogonal polynomials. Preprint, www.wis.kuleuven.ac.be/applied/intas/RomeJon.pdf, 2001.
2. C. Berg, J.P.R. Christensen, and P. Ressel. Positive definite functions on abelian semigroups. *Math. Ann.*, vol. 223, p. 163–169, 1976.
3. S. Bochner. Sturm-Liouville and heat equations whose eigenfunctions are ultraspherical polynomials or associated Bessel functions. *Proc. Conf. Differential Equations*, pages 23–48, 1955.

4. P. Feinsilver and R. Schott. *Algebraic structures and operator calculus. Vol. II*, vol. 292 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1994. Special functions and computer science.
5. P. Feinsilver and R. Schott. *Algebraic structures and operator calculus. Vol. III*, vol. 347 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
6. P. Feinsilver. *Special functions, probability semigroups, and Hamiltonian flows*, vol. 696 of *Lecture Notes in Mathematics*. Springer, Berlin, 1978.
7. G. Gasper. Positivity and the convolution structure for Jacobi series. *Ann. of Math.*, vol. 2(93), p. 112–118, 1971.
8. G. Gasper. Banach algebras for Jacobi series and positivity of a kernel. *Ann. of Math.*, 2(95), p. 261–280, 1972.
9. O. Mazet. Classification des semi-groupes de diffusion sur $\mathbb{R}$ associés à une famille de polynômes orthogonaux. In *Séminaire de probabilités XXXI, Lectures notes in Mathematics*, vol. 1655, p. 40–54, Springer, 1997.
10. O.V. Sarmanov and Z.N. Bratoeva. Probabilistic properties of bilinear expansions of Hermite polynomials. *Teor. Verujatnost. i Primenen*, vol.12, 470–481, 1967.
11. W. Schoutens. *Stochastic processes and orthogonal polynomials*, vol. 146 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 2000.
12. G. Szegö. *Orthogonal Polynomials*. American Mathematical Society, 4th edition, 1975.