

# Random walks and orthogonal polynomials: some challenges

F. ALBERTO GRÜNBAUM

To Henry, teacher and friend, with gratitude and admiration.

**ABSTRACT.** The study of several naturally arising “nearest neighbour” random walks benefits from the study of the associated orthogonal polynomials and their orthogonality measure. I consider extensions of this approach to a larger class of random walks. This raises a number of open problems.

## 1. Introduction

Consider a birth and death process, i.e., a discrete time Markov chain on the nonnegative integers, with a one step transition probability matrix  $\mathbb{P}$ . There is then a time-honored way of writing down the  $n$ -step transition probability matrix  $\mathbb{P}^n$  in terms of the orthogonal polynomials associated to  $\mathbb{P}$  and the spectral measure. This goes back to Karlin and McGregor [1957] and, as they observe, it is nothing but an application of the spectral theorem. One can find some precursors of these powerful ideas, see for instance [Harris 1952; Ledermann and Reuter 1954]. Inasmuch as this is such a deep and general result, it holds in many setups, such as a nearest neighbours random walk on the  $N$ th-roots of unity. In general this representation of  $\mathbb{P}^n$  allows one to relate properties of the Markov chain, such as recurrence or other limiting behaviour, to properties of the orthogonality measure.

---

*Mathematics Subject Classification:* 33C45, 22E45.

*Keywords:* Matrix-valued orthogonal polynomials, Karlin–McGregor representation, Jacobi polynomials.

The author was supported in part by NSF Grant # DMS 0603901.

In the few cases when one can get one's hands on the orthogonality measure and the polynomials themselves this gives fairly explicit answers to various questions.

The two main drawbacks to the applicability of this representation (to be recalled below) are:

- a) typically one cannot get either the polynomials or the measure explicitly.
- b) the method is restricted to “nearest neighbour” transition probability chains that give rise to tridiagonal matrices and thus to orthogonal polynomials.

The challenge that we pose in this paper is very simple: to try to extend the class of Markov processes whose study can benefit from a similar association.

There is an important collection of papers that study in detail the cases where the entries in  $\mathbb{P}$  depends linearly, quadratically or even rationally on the index  $n$ . We make no attempt to review these results, but we just mention that the linear case involves (associated) Laguerre and Meixner polynomials, and the other cases involve associated dual Hahn polynomials. For a very small sample of important sources dealing with this connection see [Chihara 1978; van Doorn 2003; Ismail et al. 1990].

The plan for this paper is as follows. In Section 2, we review briefly the approach of S. Karlin and J. McGregor. In Sections 3, 4, and 5, we consider a few examples of physically important Markov chains that happen to feature rather well known families of orthogonal polynomials. In Sections 6–10 we propose a way of extending this representation to the case of certain Markov chains where the one-step transition probability matrix is not necessarily tridiagonal. For concreteness we restrict ourselves to the case of pentadiagonal matrices or more generally block tridiagonal matrices. This is illustrated with some examples. A number of open problems are mentioned along the way; a few more are listed in Section 11. The material in Sections 2–5 is well known, while the proposal developed in Sections 3–10 appears to be new.

After this paper was completed we noticed that [Karlin and McGregor 1959] contains an explicit expression for the spectral matrix corresponding to the example that we treat in Section 10. The same example, as well as the connection with matrix valued orthogonal polynomials is discussed in [Dette et al. 2006]. See also [Grünbaum and de la Iglesia 2007] for a fruitful interaction with group representation theory.

## 2. The Karlin–McGregor representation

If we have

$$\mathbb{P}_{i,j} = \Pr\{X(n+1) = j \mid X(n) = i\}$$

for the 1-step transition probability of our Markov chain, and we put  $p_i = \mathbb{P}_{i,i+1}$ ,  $q_{i+1} = \mathbb{P}_{i+1,i}$ , and  $r_i = \mathbb{P}_{i,i}$  we get for the matrix  $\mathbb{P}$ , in the case of a birth and death process, the expression

$$\mathbb{P} = \begin{pmatrix} r_0 & p_0 & 0 & 0 \\ q_1 & r_1 & p_1 & 0 \\ 0 & q_2 & r_2 & p_2 \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

We will assume that  $p_j > 0$ ,  $q_{j+1} > 0$  and  $r_j \geq 0$  for  $j \geq 0$ . We also assume  $p_j + r_j + q_j = 1$  for  $j \geq 1$  and by putting  $p_0 + r_0 \leq 1$  we allow for the state  $j = 0$  to be an absorbing state (with probability  $1 - p_0 - r_0$ ). Some of these conditions can be relaxed.

If we introduce the polynomials  $Q_j(x)$  through the conditions  $Q_{-1}(0) = 0$ ,  $Q_0(x) = 1$  and, using the notation

$$Q(x) = \begin{pmatrix} Q_0(x) \\ Q_1(x) \\ \vdots \end{pmatrix},$$

we insist on the recursion relation

$$\mathbb{P}Q(x) = xQ(x),$$

we can prove the existence of a unique measure  $\psi(dx)$  supported in  $[-1, 1]$  such that

$$\pi_j \int_{-1}^1 Q_i(x) Q_j(x) \psi(dx) = \delta_{ij},$$

and obtain the Karlin–McGregor representation formula

$$(\mathbb{P}^n)_{ij} = \pi_j \int_{-1}^1 x^n Q_i(x) Q_j(x) \psi(dx).$$

Many general results can be obtained from this representation formula, some of which will be given for certain examples in the next three sections.

Here we just remark that the existence of

$$\lim_{n \rightarrow \infty} (\mathbb{P}^n)_{ij}$$

is equivalent to  $\psi(dx)$  having no mass at  $x = -1$ . If this is the case this limit is positive exactly when  $\psi(dx)$  has some mass at  $x = 1$ .

If one notices that  $Q_n(x)$  is nothing but the determinant of the  $(n+1) \times (n+1)$  upper-left corner of the matrix  $xI - \mathbb{P}$ , divided by the factor

$$p_0 p_1 \cdots p_{n-1},$$





any time that  $x$  is one of  $0, 1, \dots, 2N$ . This means that the eigenvalues of the matrix  $\mathbb{P}$  above are given by the values of  $1 - (x/N)$  at these values of  $x$ , that is,

$$1, 1 - \frac{1}{N}, \dots, -1,$$

and that the corresponding eigenvectors are the values of

$$[K_0(x), K_1(x), \dots, K_{2N}(x)]^T$$

at these values of  $x$ .

Since the matrix  $\mathbb{P}$  above is the one step transition probability matrix for our urn model we conclude that

$$(\mathbb{P}^n)_{ij} = \pi_j \sum_{x=0}^{2N} \left(1 - \frac{x}{N}\right)^n K_i(x) K_j(x) \frac{\binom{2N}{x}}{2^{2N}}.$$

We can use these expressions to rederive some results given in [Kac 1947].

We have

$$(\mathbb{P}^n)_{00} = \sum_{x=0}^{2N} \left(1 - \frac{x}{N}\right)^n \frac{\binom{2N}{x}}{2^{2N}}$$

and the “generating function” for these probabilities, defined by

$$U(z) \equiv \sum_{n=0}^{\infty} z^n (\mathbb{P}^n)_{00}$$

becomes

$$U(z) = \sum_{x=0}^{2N} \frac{N}{N(1-z) + xz} \frac{\binom{2N}{x}}{2^{2N}}.$$

In particular  $U(1) = \infty$  and then the familiar “renewal equation” (see [Feller 1967]) given by

$$U(z) = F(z)U(z) + 1,$$

where  $F(z)$  is the generating function for the probabilities  $f_n$  of returning from state 0 to state 0 for the first time in  $n$  steps

$$F(z) = \sum_{n=0}^{\infty} z^n f_n$$

gives

$$F(z) = 1 - \frac{1}{U(z)}$$

Therefore we have  $F(1) = 1$ , indicating that one returns to state 0 with probability one in finite time.

These results allow us to compute the expected time to return to state 0. This expected value is given by  $F'(1)$ , and we have

$$F'(z) = \frac{U'(z)}{U^2(z)}.$$

Since

$$U'(z) = \sum_{x=0}^{2N} \frac{N(N-x)}{(N(1-z) + xz)^2} \frac{\binom{2N}{x}}{2^{2N}}$$

we get  $F'(1) = 2^{2N}$ . The same method shows that any state  $i = 0, \dots, 2N$  is also recurrent and that the expected time to return to it is given by

$$\frac{2^{2N}}{\binom{2N}{i}}.$$

The moral of this story is clear: if  $i = 0$  or  $2N$ , or if  $i$  is close to these values, meaning that we start from a state where most balls are in one urn, it will take on average a huge amount of time to get back to this state. On the other hand if  $i = N$ , that is, we are starting from a very balanced state, then we will (on average) return to this state fairly soon. Thus we see how the issues of irreversibility and recurrence are rather subtle.

In a very precise sense these polynomials are discrete analogs of those of Hermite in the case of the real line. For interesting material regarding this section the reader should consult [Askey 2005].

#### 4. A Chebyshev-type example

The example below illustrates nicely how certain recurrence properties of the process are related to the presence of point masses in the orthogonality measure. This is seen by comparing the two integrals at the end of the section.

Consider the matrix

$$\mathbb{P} = \begin{pmatrix} 0 & 1 & 0 & & & \\ q & 0 & p & & & \\ 0 & q & 0 & p & & \\ & & \ddots & \ddots & \ddots & \\ & & & & & \ddots \end{pmatrix}$$

with  $0 \leq p \leq 1$  and  $q = 1 - p$ . We look for polynomials  $Q_j(x)$  such that

$$Q_{-1}(x) = 0, \quad Q_0(x) = 1$$

and if  $Q(x)$  denotes the vector

$$Q(x) \equiv \begin{pmatrix} Q_0(x) \\ Q_1(x) \\ Q_2(x) \\ \vdots \end{pmatrix}$$

we ask that we should have

$$\mathbb{P}Q(x) = xQ(x).$$

The matrix  $\mathbb{P}$  can be conjugated into a symmetric one and in this fashion one can find the explicit expression for these polynomials.

We have

$$Q_j(x) = \left(\frac{q}{p}\right)^{j/2} \left( (2-2p)T_j\left(\frac{x}{2\sqrt{pq}}\right) + (2p-1)U_j\left(\frac{x}{2\sqrt{pq}}\right) \right)$$

where  $T_j$  and  $U_j$  are the Chebyshev polynomials of the first and second kind.

If  $p \geq 1/2$  we have

$$\left(\frac{p}{1-p}\right)^n \int_{-\sqrt{4pq}}^{\sqrt{4pq}} Q_n(x)Q_m(x) \frac{\sqrt{4pq-x^2}}{1-x^2} dx = \delta_{nm} \begin{cases} 2(1-p)\pi & \text{if } n=0, \\ 2p(1-p)\pi & \text{if } n \geq 1, \end{cases}$$

while if  $p \leq 1/2$  we get a new phenomenon, namely the presence of point masses in the spectral measure

$$\begin{aligned} \left(\frac{p}{1-p}\right)^n & \left( \int_{-\sqrt{4pq}}^{\sqrt{4pq}} Q_n(x)Q_m(x) \frac{\sqrt{4pq-x^2}}{1-x^2} dx \right. \\ & \left. + (2-4p)\pi [Q_n(1)Q_m(1) + Q_n(-1)Q_m(-1)] \right) \\ & = \delta_{nm} \begin{cases} 2(1-p)\pi & \text{if } n=0, \\ 2p(1-p)\pi & \text{if } n \geq 1. \end{cases} \end{aligned}$$

From a probabilistic point of view these results are very natural.

## 5. The Hahn polynomials, Laplace and Bernoulli

As has been pointed out before, a limitation of this method is given by the sad fact that given the matrix  $\mathbb{P}$  very seldom can one write down the corresponding polynomials and their orthogonality measure. In general there is no reason why physically interesting Markov chains will give rise to situations where these mathematical objects can be found explicitly.

The example below shows that one can get lucky: there is a very old model of the exchange of heat between two bodies going back to Laplace and Bernoulli,



see [Feller 1967, p. 378]. It turns out that in this case the corresponding orthogonal polynomials can be determined explicitly.

The Bernoulli–Laplace model for the exchange of heat between two bodies consists of two urns, labeled 1 and 2. Initially there are  $W$  white balls in urn 1 and  $B$  black balls in urn 2. The transition mechanism is as follows: a ball is picked from each urn and these two balls are switched. It is natural to expect that eventually both urns will have a nice mixture of white and black balls.

The state of the system at any time is described by  $w$ , defined to be the number of white balls in urn 1. It is clear that we have, for  $w = 0, 1, \dots, W$

$$\begin{aligned}\mathbb{P}_{w,w+1} &= \frac{W-w}{W} \frac{W-w}{B}, & \mathbb{P}_{w,w-1} &= \frac{w}{W} \frac{B-W+w}{B}, \\ \mathbb{P}_{w,w} &= \frac{w}{W} \frac{W-w}{B} + \frac{W-w}{W} \frac{B-W+w}{B}.\end{aligned}$$

Notice that

$$\mathbb{P}_{w,w-1} + \mathbb{P}_{w,w} + \mathbb{P}_{w,w+1} = 1.$$

Now introduce the dual Hahn polynomials by means of

$$\begin{aligned}R_n(\lambda(x)) &= {}_3\tilde{F}_2 \left( \begin{matrix} -n, -x, x-W-B-1 \\ -W, -W \end{matrix} \middle| 1 \right) \\ n &= 0, \dots, W; \quad x = 0, \dots, W.\end{aligned}$$

These polynomials depend in general on one more parameter.

Notice that these are polynomials of degree  $n$  in

$$\lambda(x) \equiv x(x-W-B-1).$$

One has

$$\mathbb{P}_{w,w-1} R_{w-1} + \mathbb{P}_{w,w} R_w + \mathbb{P}_{w,w+1} R_{w+1} = \left( 1 - \frac{x(B+W-x+1)}{BW} \right) R_w.$$

This means that for each value of  $x = 0, \dots, W$  the vector

$$\begin{pmatrix} R_0(\lambda(x)) \\ R_1(\lambda(x)) \\ \vdots \\ R_W(\lambda(x)) \end{pmatrix}$$

is an eigenvector of the matrix  $\mathbb{P}$  with eigenvalue  $1 - \frac{x(B+W-x+1)}{BW}$ . The relevant orthogonality relation is given by

$$\pi_j \sum_{x=0}^W R_i(\lambda(x)) R_j(\lambda(x)) \mu(x) = \delta_{ij}$$

with

$$\mu(x) = \frac{w!(-w)_x(-w)_x(2x - W - B - 1)}{(-1)^{x+1}x!(-B)_x(x - W - B - 1)_{w+1}}, \quad \pi_j = \frac{(-w)_j}{j!} \frac{(-B)_{w-j}}{(w-j)!}.$$

The Karlin–McGregor representation gives

$$(\mathbb{P}^n)_{ij} = \pi_j \sum_{x=0}^W R_i(\lambda(x)) R_j(\lambda(x)) e^n(x) \mu(x)$$

$$\text{with } e(x) = 1 - \frac{x(B+W-x+1)}{BW}.$$

These results can be used, once again, to get some quantitative results on this process.

Interestingly enough, these polynomials were considered in great detail by S. Karlin and J. McGregor [1961] and used by these authors in the context of a model in genetics describing fluctuations of gene frequency under the influence of mutation and selection. The reader will find useful remarks in [Diaconis and Shahshahani 1987].

## 6. The classical orthogonal polynomials and the bispectral problem

The examples discussed above illustrate the following point: quite often the orthogonal polynomials that are associated with important Markov chains belong to the small class of polynomials usually referred to as *classical*. By this one means that they satisfy not only three term recursion relations but that they are also the common eigenfunctions of some fixed (usually second order) differential operator. The search for polynomials of this kind goes back at least to [Bochner 1929]. In fact this issue is even older; see [Routh 1884] and also [Ismail 2005] for a more complete discussion.

In the context where both variables are continuous, this problem has been raised in [Duistermaat and Grünbaum 1986]. For a view of some related subjects see [Harnad and Kasman 1998]. The reader will find useful material in [Askey and Wilson 1985; Andrews et al. 1999; Ismail 2005].

## 7. Matrix-valued orthogonal polynomials

Here we recall a notion due to M. G. Krein [1949; 1971]. Given a self adjoint positive definite matrix-valued smooth weight function  $W(x)$  with finite moments, we can consider the skew symmetric bilinear form defined for any pair of matrix-valued polynomial functions  $P(x)$  and  $Q(x)$  by the numerical matrix

$$(P, Q) = (P, Q)_W = \int_{\mathbb{R}} P(x) W(x) Q^*(x) dx,$$

where  $Q^*(x)$  denotes the conjugate transpose of  $Q(x)$ . By the usual construction this leads to the existence of a sequence of matrix-valued orthogonal polynomials with nonsingular leading coefficient.

Given an orthogonal sequence  $\{P_n(x)\}_{n \geq 0}$  one gets by the usual argument a three term recursion relation

$$xP_n(x) = A_n P_{n-1}(x) + B_n P_n(x) + C_n P_{n+1}(x), \quad (7-1)$$

where  $A_n$ ,  $B_n$  and  $C_n$  are matrices and the last one is nonsingular.

We now turn our attention to an important class of orthogonal polynomials which we will call *classical matrix-valued orthogonal polynomials*. Very much as in [Duran 1997; Grünbaum et al. 2003; Grünbaum et al. 2005] we say that the weight function is *classical* if there exists a second order ordinary differential operator  $D$  with matrix-valued polynomial coefficients  $A_j(x)$  of degree less or equal to  $j$  of the form

$$D = A_2(x) \frac{d^2}{dx^2} + A_1(x) \frac{d}{dx} + A_0(x), \quad (7-2)$$

such that for an orthogonal sequence  $\{P_n\}$ , we have

$$DP_n^* = P_n^* \Lambda_n, \quad (7-3)$$

where  $\Lambda_n$  is a real-valued matrix. This form of the eigenvalue equation (7-3) appears naturally in [Grünbaum et al. 2002] and differs only superficially with the form used in [Duran 1997], where one uses right handed differential operators.

During the last few years much activity has centered around an effort to produce families of matrix-valued orthogonal polynomials that would satisfy differential equations as those above. One of the examples that resulted from this search, see [Grünbaum 2003], will be particularly useful later on.

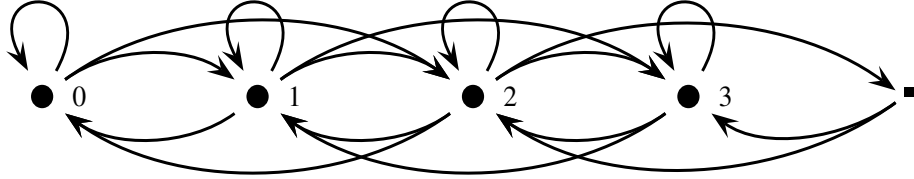
## 8. Pentadiagonal matrices and matrix-valued orthogonal polynomials

Given a pentadiagonal scalar matrix it is often useful to think of it either in its original unblocked form or as being made, let us say, of  $2 \times 2$  blocks. These two ways of seeing a matrix, and the fact that matrix operations like multiplication can be performed “by blocks”, has proved very important in the development of fast algorithms.

In the case of a birth and death process it is useful to think of a graph like

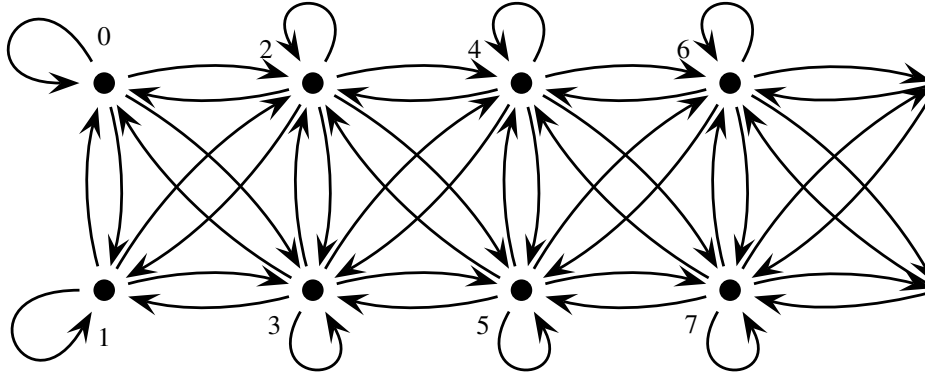


Suppose that we are dealing with a more complicated Markov chain in the same probability space, where the elementary transitions can go beyond “nearest neighbours”. In such a case the graph may look as follows:



The matrix  $\mathbb{P}$  going with the graph above is now pentadiagonal. By thinking of it in the manner mentioned above we get a block tridiagonal matrix. As an extra bonus, its off-diagonal blocks are triangular.

The graph



clearly corresponds to a general block tridiagonal matrix, with blocks of size  $2 \times 2$ .

If  $\mathbb{P}_{i,j}$  denotes the  $(i,j)$ -block of  $\mathbb{P}$  we can generate a sequence of  $2 \times 2$  matrix-valued polynomials  $Q_i(t)$  by imposing the three-term recursion of Section 8. Using the notation of Section 2, we would have

$$\mathbb{P}Q(x) = xQ(x),$$

where the entries of the column vector  $Q(x)$  are now  $2 \times 2$  matrices.

Proceeding as in the scalar case, this relation can be iterated to give

$$\mathbb{P}^n Q(x) = x^n Q(x),$$

and if we assume the existence of a weight matrix  $W(x)$  as in Section 7, with the property

$$(Q_j, Q_j)\delta_{i,j} = \int_{\mathbb{R}} Q_i(x)W(x)Q_j^*(x)dx,$$

it is then clear that one can get an expression for the  $(i, j)$  entry of the block matrix  $\mathbb{P}^n$  that would look exactly as in the scalar case, namely

$$(\mathbb{P}^n)_{ij}(Q_j, Q_j) = \int x^n Q_i(x) W(x) Q_j^*(x) dx.$$

Just as in the scalar case, this expression becomes useful when we can get our hands on the matrix-valued polynomials  $Q_i(x)$  and the orthogonality measure  $W(x)$ . Notice that we have not discussed conditions on the matrix  $\mathbb{P}$  to give rise to such a measure. For this issue the reader can consult [Durán and Polo 2002; Duran 1999] and the references in these papers.

The spectral theory of a scalar double-infinite tridiagonal matrix leads naturally to a  $2 \times 2$  semi-infinite matrix. This has been looked at in terms of random walks in [Pruitt 1963]. In [Ismail et al. 1990] this work is elaborated further to get a formula that could be massaged to look like the right-hand side of the one above. See also the last section in [Karlin and McGregor 1959].

## 9. An explicit example

Consider the matrix-valued polynomials given by the three-term recursion relation

$$A_n \Phi_{n-1}(x) + B_n \Phi_n(x) + C_n \Phi_{n+1}(x) = t \Phi_n(x), \quad n \geq 0,$$

with

$$\Phi_{-1}(x) = 0, \quad \Phi_0(t) = I,$$

and where the entries in  $A_n, B_n, C_n$  are given by

$$\begin{aligned} A_n^{11} &:= \frac{n(\alpha+n)(\beta+2\alpha+2n+3)}{(\beta+\alpha+2n+1)(\beta+\alpha+2n+2)(\beta+2\alpha+2n+1)}, \\ A_n^{12} &:= \frac{2n(\beta+1)}{(\beta+2n+1)(\beta+\alpha+2n+2)(\beta+2\alpha+2n+1)}, \quad A_n^{21} := 0, \\ A_n^{22} &:= \frac{n(\alpha+n+1)(\beta+2n+3)}{(\beta+2n+1)(\beta+\alpha+2n+2)(\beta+\alpha+2n+3)}, \\ C_n^{11} &:= \frac{(\beta+n+2)(\beta+2n+1)(\beta+\alpha+n+2)}{(\beta+2n+3)(\beta+\alpha+2n+2)(\beta+\alpha+2n+3)}, \\ C_n^{21} &:= \frac{2(\beta+1)(\beta+n+2)}{(\beta+2n+3)(\beta+\alpha+2n+3)(\beta+2\alpha+2n+5)}, \quad C_n^{12} := 0, \\ C_n^{22} &:= \frac{(\beta+n+2)(\beta+\alpha+n+3)(\beta+2\alpha+2n+3)}{(\beta+\alpha+2n+3)(\beta+\alpha+2n+4)(\beta+2\alpha+2n+5)}, \end{aligned}$$

$$\begin{aligned}
B_n^{11} &:= 1 + \frac{n(\beta+n+1)(\beta+2n-1)}{(\beta+2n+1)(\beta+\alpha+2n+1)} - \frac{(n+1)(\beta+n+2)(\beta+2n+1)}{(\beta+2n+3)(\beta+\alpha+2n+3)} \\
&\quad - \frac{2(\beta+1)^2}{(\beta+2n+1)(\beta+2n+3)(\beta+2\alpha+2n+3)}, \\
B_n^{12} &:= \frac{2(\beta+1)(\alpha+\beta+n+2)}{(\beta+2n+3)(\beta+\alpha+2n+2)(\beta+2\alpha+2n+3)}, \\
B_n^{21} &:= \frac{2(\alpha+n+1)(\beta+1)}{(\beta+2n+1)(\beta+\alpha+2n+3)(\beta+2\alpha+2n+3)}, \\
B_n^{22} &:= 1 + \frac{n(\beta+n+1)(\beta+2n+3)}{(\beta+2n+1)(\beta+\alpha+2n+2)} - \frac{(n+1)(\beta+n+2)(\beta+2n+5)}{(\beta+2n+3)(\beta+\alpha+2n+4)} \\
&\quad + \frac{2(\beta+1)^2}{(\beta+2n+1)(\beta+2n+3)(\beta+2\alpha+2n+3)}.
\end{aligned}$$

Notice that the matrices  $A_n$  and  $C_n$  are upper and lower triangular respectively.

If the matrix  $\Psi_0(x)$  is given by

$$\Psi_0(x) = \begin{pmatrix} 1 & 1 \\ 1 & \frac{(\beta+2\alpha+3)x}{\beta+1} - \frac{2(\alpha+1)}{\beta+1} \end{pmatrix}$$

one can see that the polynomials  $\Phi_n(x)$  satisfy the orthogonality relation

$$\int_0^1 \Phi_i(x)W(x)\Phi_j^*(x)dx = 0 \quad \text{if } i \neq j,$$

where

$$W(x) = \Psi_0(x) \begin{pmatrix} (1-x)\beta x^{\alpha+1} & 0 \\ 0 & (1-x)\beta x^\alpha \end{pmatrix} \Psi_0^*(x).$$

The polynomials  $\Phi_n(x)$  are classical in the sense that they are eigenfunctions of a fixed second order differential operator. More precisely, we have

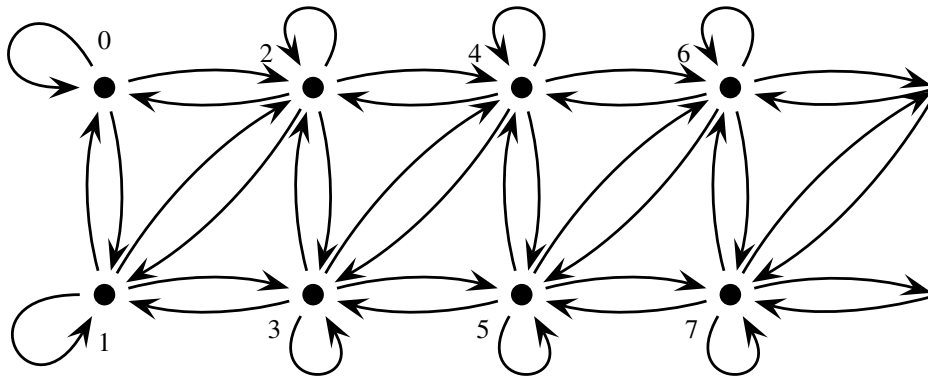
$$\mathcal{F}\Phi_n^* = \Phi_n^* \Lambda_n,$$

where  $\Lambda_n = \text{diag}(-n^2 - (\alpha + \beta + 2)n + \alpha + 1 + \frac{1}{2}(\beta + 1), -n^2 - (\alpha + \beta + 3)n)$  and

$$\begin{aligned}
\mathcal{F} &= x(1-x) \left( \frac{d}{dx} \right)^2 \\
&\quad + \begin{pmatrix} \frac{(\alpha+1)(\beta+2\alpha+5)}{\beta+2\alpha+3} - (\alpha+\beta+3)x & \frac{2\alpha+2}{2\alpha+\beta+3} + x \\ \frac{\beta+1}{\beta+2\alpha+3} & \frac{(\alpha+2)\beta+2\alpha^2+5\alpha+4}{\beta+2\alpha+3} - (\alpha+\beta+4)x \end{pmatrix} \frac{d}{dx} \\
&\quad + \begin{pmatrix} \alpha+1 + \frac{\beta+1}{2} & 0 \\ 0 & 0 \end{pmatrix} I.
\end{aligned}$$

As mentioned earlier this is the reason why this example has surfaced recently; see [Grünbaum 2003]. An explicit expression for the polynomials themselves is given in [Tirao 2003, Corollary 3].

Now we observe that the entries of the corresponding pentadiagonal matrix are all nonnegative, and that the sum of the entries on any given row are all equal to 1. This allows for an immediate probabilistic interpretation of the pentadiagonal matrix as the one step transition probability matrix for a Markov chain whose state space could be visualized in the graph



I find it remarkable that this example, which was produced for an entirely different purpose, should have this extra property. Finding an appropriate combinatorial mechanism, maybe in terms of urns, that goes along with this example remains an interesting challenge.

Two final observations dealing with these state spaces that can be analyzed using matrix-valued orthogonal polynomials. If we were using matrix-valued polynomials of size  $N$  we would have as state space a semiinfinite network consisting of  $N$  (instead of two) parallel collection of nonnegative integers with connections going from each of the  $N$  states on each vertical rung to every one in the same rung and the two neighbouring ones. The examples in [Grünbaum et al. 2002] give instances of this situation with a rather local connection pattern.

In the case of  $N = 2$  one could be tempted to paraphrase a well known paper and say that “it has not escaped our notice that” some of these models could be used to study transport phenomena along a DNA segment.

## 10. Another example

Here we consider a different example of matrix-valued orthogonal polynomials whose block tridiagonal matrix can be seen as a scalar pentadiagonal matrix with nonnegative elements. In this case the sum of the elements in the rows of this scalar matrix is not identically one, but this poses no problem in terms of a Karlin–McGregor-type representation formula for the entries of the powers  $\mathbb{P}^n$ .

This example has the important property that the orthogonality weight matrix  $W(x)$ , as well as the polynomials themselves are explicitly known. This is again a classical situation; see [Castro and Grünbaum 2006].

Consider the block tridiagonal matrix

$$\begin{pmatrix} B_0 & I & & & \\ A_1 & B_1 & I & & \\ 0 & A_2 & B_2 & I & \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

with  $2 \times 2$  blocks given by  $B_0 = \frac{1}{2}I$ ,  $B_n = 0$  if  $n \geq 1$ , and  $A_n = \frac{1}{4}I$  if  $n \geq 1$ . In this case one can compute explicitly the matrix-valued polynomials  $P_n$  given by

$$A_n P_{n-1}(x) + B_n P_n(x) + P_{n+1}(x) = x P_n(x), \quad P_{-1}(x) = 0, \quad P_0(x) = I.$$

One gets

$$P_n(x) = \frac{1}{2^n} \begin{pmatrix} U_n(x) & -U_{n-1}(x) \\ -U_{n-1}(x) & U_n(x) \end{pmatrix}$$

where  $U_n(x)$  is the  $n$ -th Chebyshev polynomial of the second kind.

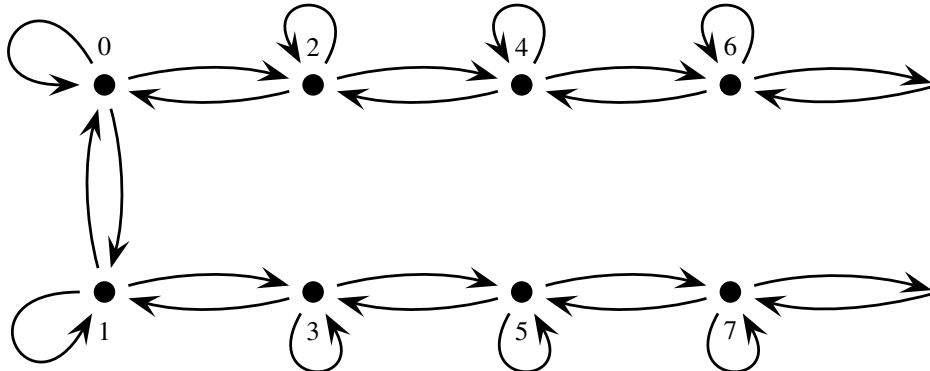
The orthogonality measure is read off from the identity

$$\frac{4^i}{\pi} \int_{-1}^1 P_i(x) \frac{1}{\sqrt{1-x^2}} \begin{pmatrix} 1 & x \\ x & 1 \end{pmatrix} P_j(x) dx = \delta_{ij} I.$$

We get, for  $n = 0, 1, 2, \dots$

$$\frac{4^i}{\pi} \int_{-1}^1 x^n P_i(x) \frac{1}{\sqrt{1-x^2}} \begin{pmatrix} 1 & x \\ x & 1 \end{pmatrix} P_j(x) dx = (\mathbb{P}^n)_{ij},$$

where, as above,  $(\mathbb{P}^n)_{ij}$  stands for the  $i, j$  block of the matrix  $\mathbb{P}^n$ .



In this way one can compute the entries of the powers  $\mathbb{P}^n$ , with  $\mathbb{P}^n$  thought of as a pentadiagonal matrix, namely



$$\mathbb{P} = \begin{pmatrix} 0 & \frac{1}{2} & 1 & 0 & 0 & 0 & \\ \frac{1}{2} & 0 & 0 & 1 & 0 & 0 & \ddots \\ \frac{1}{4} & 0 & 0 & 0 & 1 & 0 & \ddots \\ 0 & \frac{1}{4} & 0 & 0 & 0 & 1 & \ddots \\ & 0 & \frac{1}{4} & 0 & 0 & 0 & \ddots \\ & & 0 & \frac{1}{4} & 0 & 0 & \ddots \\ & & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

This example goes along with the graph on the previous page.

## 11. A few more challenges

We have already pointed out a few challenges raised by our attempt to extend the Karlin–McGregor representation beyond its original setup. Here we list a few more open problems. The reader will undoubtedly come up with more.

Is there a natural version of the models introduced by Bernoulli–Laplace and by P. and T. Ehrenfest whose solution features matrix-valued polynomials?

Is it possible to modify the simplest Chebyshev-type examples in [Duran 1999] to accommodate cases where some of the blocks in the tridiagonal matrix give either absorption or reflection boundary conditions?

One could consider the emerging class of polynomials of several variables and find here interesting instances where the state space is higher-dimensional. For a systematic study of polynomials in several variables one should consult [Dunkl and Xu 2001] as well as the monograph [Macdonald 2003], on Macdonald polynomials of various kinds. A look at the pioneering work of Tom Koornwinder (see [Koornwinder 1975], for instance) is always a very good idea.

\* \* \* \*

After this paper was finished I came up with two independent sources of multivariable polynomials of the type alluded to in the previous paragraph. One is the series of papers by Hoare and Rahman [1979; 1984; 1988;  $\geq 2007$ ]. The other deals with the papers [Milch 1968; Iliev and Xu 2007; Geronimo and Iliev 2006].

In queueing theory one finds the notion of quasi-birth-and-death processes; see [Latouche and Ramaswami 1999; Neuts 1989]. Within those that are non-homogeneous one could find examples where the general approach advocated here might be useful.

## References

- [Andrews et al. 1999] G. E. Andrews, R. Askey, and R. Roy, *Special functions*, Encyclopedia of Mathematics and its Applications **71**, Cambridge University Press, Cambridge, 1999.
- [Askey 2005] R. Askey, “Evaluation of Sylvester type determinants using orthogonal polynomials”, pp. 1–16 in *Advances in analysis*, World Scientific, Singapore, 2005.
- [Askey and Wilson 1985] R. Askey and J. Wilson, “Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials”, *Mem. Amer. Math. Soc.* **54**:319 (1985), iv+55.
- [Bochner 1929] S. Bochner, “Über Sturm-Liouvillesche Polynomsysteme”, *Math. Z.* **29**:1 (1929), 730–736.
- [Castro and Grünbaum 2006] M. Castro and F. A. Grünbaum, “The algebra of matrix valued differential operators associated to a given family of matrix valued orthogonal polynomials: five instructive examples”, *Internat. Math. Res. Notices* **2006**:7 (2006), 1–33.
- [Chihara 1978] T. S. Chihara, *An introduction to orthogonal polynomials*, Gordon and Breach, New York, 1978.
- [Dette et al. 2006] H. Dette, B. Reuther, W. J. Studden, and M. Zygmunt, “Matrix measures and random walks with a block tridiagonal transition matrix”, *SIAM J. Matrix Anal. Appl.* **29**:1 (2006), 117–142.
- [Diaconis and Shahshahani 1987] P. Diaconis and M. Shahshahani, “Time to reach stationarity in the Bernoulli-Laplace diffusion model”, *SIAM J. Math. Anal.* **18**:1 (1987), 208–218.
- [van Doorn 2003] E. A. van Doorn, “On associated polynomials and decay rates for birth-death processes”, *J. Math. Anal. Appl.* **278**:2 (2003), 500–511.
- [Duistermaat and Grünbaum 1986] J. J. Duistermaat and F. A. Grünbaum, “Differential equations in the spectral parameter”, *Comm. Math. Phys.* **103**:2 (1986), 177–240.
- [Dunkl and Xu 2001] C. F. Dunkl and Y. Xu, *Orthogonal polynomials of several variables*, Encyclopedia of Mathematics and its Applications **81**, Cambridge University Press, Cambridge, 2001.
- [Duran 1997] A. J. Duran, “Matrix inner product having a matrix symmetric second order differential operator”, *Rocky Mountain J. Math.* **27**:2 (1997), 585–600.
- [Duran 1999] A. J. Duran, “Ratio asymptotics for orthogonal matrix polynomials”, *J. Approx. Theory* **100**:2 (1999), 304–344.
- [Durán and Polo 2002] A. J. Durán and B. Polo, “Gaussian quadrature formulae for matrix weights”, *Linear Algebra Appl.* **355** (2002), 119–146.
- [Ehrenfest 1907] P. Ehrenfest and T. Eherenfest, “Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem”, *Phys. Z.* **8** (1907), 311–314.
- [Feller 1967] V. Feller, *An introduction to probability theory and its applications*, vol. 1, 3rd ed., Wiley, New York, 1967.
- [Geronimo and Iliev 2006] J. Geronimo and P. Iliev, “Two variable deformations of the Chebyshev measure”, preprint, 2006. Available at [arXiv.org/abs/math/0612489](http://arXiv.org/abs/math/0612489).

- [Grünbaum 2003] F. A. Grünbaum, “Matrix valued Jacobi polynomials”, *Bull. Sci. Math.* **127**:3 (2003), 207–214.
- [Grünbaum and de la Iglesia 2007] F. A. Grünbaum and M. de la Iglesia, “Matrix valued orthogonal polynomials arising from group representation theory and a family of quasi-birth-and-death processes”, preprint, 2007. Submitted.
- [Grünbaum et al. 2002] F. A. Grünbaum, I. Pacharoni, and J. Tirao, “Matrix valued spherical functions associated to the complex projective plane”, *J. Funct. Anal.* **188**:2 (2002), 350–441.
- [Grünbaum et al. 2003] F. A. Grünbaum, I. Pacharoni, and J. Tirao, “Matrix valued orthogonal polynomials of the Jacobi type”, *Indag. Math. (N.S.)* **14**:3–4 (2003), 353–366.
- [Grünbaum et al. 2005] F. A. Grünbaum, I. Pacharoni, and J. Tirao, “Matrix valued orthogonal polynomials of Jacobi type: the role of group representation theory”, *Ann. Inst. Fourier (Grenoble)* **55**:6 (2005), 2051–2068.
- [Harnad and Kasman 1998] J. Harnad and A. Kasman (editors), *The bispectral problem* (Montreal, 1997), CRM Proceedings & Lecture Notes **14**, American Mathematical Society, Providence, RI, 1998.
- [Harris 1952] T. E. Harris, “First passage and recurrence distributions”, *Trans. Amer. Math. Soc.* **73** (1952), 471–486.
- [Hoare and Rahman 1979] M. R. Hoare and M. Rahman, “Distributive processes in discrete systems”, *Phys. A* **97**:1 (1979), 1–41.
- [Hoare and Rahman 1984] M. R. Hoare and M. Rahman, “Cumulative Bernoulli trials and Krawtchouk processes”, *Stochastic Process. Appl.* **16**:2 (1984), 113–139.
- [Hoare and Rahman 1988] M. R. Hoare and M. Rahman, “Cumulative hypergeometric processes: a statistical role for the  ${}_nF_{n-1}$  functions”, *J. Math. Anal. Appl.* **135**:2 (1988), 615–626.
- [Hoare and Rahman  $\geq$  2007] M. Hoare and M. Rahman, “A probabilistic origin for a new class of bivariate polynomials”, Technical report. To appear.
- [Iliev and Xu 2007] P. Iliev and Y. Xu, “Discrete orthogonal polynomials and difference equations of several variables”, *Adv. Math.* **212**:1 (2007), 1–36.
- [Ismail 2005] M. E. H. Ismail, *Classical and quantum orthogonal polynomials in one variable*, Encyclopedia of Mathematics and its Applications **98**, Cambridge University Press, Cambridge, 2005.
- [Ismail et al. 1990] M. E. H. Ismail, D. R. Masson, J. Letessier, and G. Valent, “Birth and death processes and orthogonal polynomials”, pp. 229–255 in *Orthogonal polynomials* (Columbus, OH, 1989), edited by P. Nevai, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **294**, Kluwer Acad. Publ., Dordrecht, 1990.
- [Kac 1947] M. Kac, “Random walk and the theory of Brownian motion”, *Amer. Math. Monthly* **54** (1947), 369–391.
- [Karlin and McGregor 1957] S. Karlin and J. McGregor, “The classification of birth and death processes”, *Trans. Amer. Math. Soc.* **86** (1957), 366–400.

- [Karlin and McGregor 1959] S. Karlin and J. McGregor, “Random walks”, *Illinois J. Math.* **3** (1959), 66–81.
- [Karlin and McGregor 1961] S. Karlin and J. L. McGregor, “The Hahn polynomials, formulas and an application”, *Scripta Math.* **26** (1961), 33–46.
- [Koorwinder 1975] T. Koorwinder, “Two-variable analogues of the classical orthogonal polynomials”, pp. 435–495 in *Theory and application of special functions*, edited by R. Askey, Math. Res. Center, Univ. Wisconsin, Publ. **35**, Academic Press, New York, 1975.
- [Krein 1949] M. Krein, “Infinite  $J$ -matrices and a matrix-moment problem”, *Doklady Akad. Nauk SSSR (N.S.)* **69** (1949), 125–128.
- [Krein 1971] M. G. Krein, “Fundamental aspects of the representation theory of hermitian operators with deficiency index  $(m, m)$ ”, *Amer. Math. Soc. Transl. (2)* **97** (1971).
- [Latouche and Ramaswami 1999] G. Latouche and V. Ramaswami, *Introduction to matrix analytic methods in stochastic modeling*, ASA-SIAM Series on Statistics and Applied Probability, Soc. Ind. Appl. Math., Philadelphia, 1999.
- [Ledermann and Reuter 1954] W. Ledermann and G. E. H. Reuter, “Spectral theory for the differential equations of simple birth and death processes”, *Philos. Trans. Roy. Soc. London. Ser. A.* **246** (1954), 321–369.
- [Macdonald 2003] I. G. Macdonald, *Affine Hecke algebras and orthogonal polynomials*, Cambridge Tracts in Mathematics **157**, Cambridge University Press, Cambridge, 2003.
- [Milch 1968] P. R. Milch, “A multi-dimensional linear growth birth and death process”, *Ann. Math. Statist.* **39** (1968), 727–754.
- [Neuts 1989] M. F. Neuts, *Structured stochastic matrices of  $M/G/1$  type and their applications*, vol. 5, Dekker, New York, 1989.
- [Pruitt 1963] W. E. Pruitt, “Bilateral birth and death processes”, *Trans. Amer. Math. Soc.* **107** (1963), 508–525.
- [Routh 1884] E. Routh, “On some properties of certain solutions of a differential equation of the second order”, *Proc. London Math. Soc.* **16** (1884), 245–261.
- [Schrödinger and Kohlrausch 1926] E. Schrödinger and F. Kohlrausch, “Das Ehrenfestche Model der H-Kurve”, *Phys. Z.* **27** (1926), 306–313.
- [Tirao 2003] J. A. Tirao, “The matrix-valued hypergeometric equation”, *Proc. Natl. Acad. Sci. USA* **100**:14 (2003), 8138–8141.

F. ALBERTO GRÜNBAUM  
DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF CALIFORNIA  
BERKELEY, CA 94720  
UNITED STATES  
grunbaum@math.berkeley.edu