On intertwining relations between Ehrenfest, Yule and Ornstein-Uhlenbeck processes

Laurent Miclo[†] and Pierre Patie[‡]

[†] Toulouse School of Economics, UMR 5314
Institut de Mathématiques de Toulouse, UMR 5219,
CNRS and University of Toulouse

[‡] School of Operations Research and Information Engineering,
Cornell University

Abstract

Markovian intertwining relations between two Markov semigroups are related to the partial inclusion of the spectra of their generators, at least for finite ergodic processes. We check the limitations of this observation by investigating the Markov intertwining relations between the Ehrenfest, Yule and Ornstein-Uhlenbeck processes, whose spectra are all included into $-\mathbb{Z}_+$.

Keywords: Markov intertwining, Ehrenfest process, Yule process, Ornstein-Uhlenbeck process.

MSC2010: primary: 60J35, secondary: 60J27, 60J60, 05A15, 15A18, 42C05.

1 Introduction

The state space reduction is an important question in Markov process theory and its applications. Given a Markov process $X \coloneqq (X_t)_{t\geqslant 0}$ on a large state space V, one is looking for another Markov process $Y \coloneqq (Y_t)_{t\geqslant 0}$ on a much smaller state space W and serving as a "relatively good image" of the evolution of X. The process Y corresponds to a limited quantity of information that one would like to extract from X while still providing for a sufficient knowledge about certain characteristics of the corresponding conditional distributions of the positions of X. Some qualitative features are desirable in such an approximation/prediction procedure:

(i) The "indicative process" Y takes into account the limited observation chosen to be made on X in an non-anticipative way: for any $t \geqslant 0$, to construct the piece of trajectory $Y_{[0,t]}$, we should only use what is extracted from X up to time t and maybe some additional independent randomness (which may require an enlargement of the underlying probability space, from a mathematical point of view). Furthermore, $Y_{[0,t]}$ is the only information we keep from our partial observations from $X_{[0,t]}$.

(ii) The process $(X, Y) \coloneqq (X_t, Y_t)_{t\geqslant 0}$ is Markovian, to enable for "online" constructions. It is time-homogeneous, as all the processes considered here.

(iii) For any $t \geqslant 0$, knowing the trajectory $Y_{[0,t]}$, the conditional law of X_t should depend only on Y_t, to avoid the storage of too much information, since this is the objective of state space reduction. To be quite restrictive, we do not allow either for an explicit dependence on time.

Namely, we want to use some partial observations of X to construct in an adapted way a Markov process Y whose current value Y_t enables to make an "as good as possible" prediction on some aspects of the position X_t, given that we only observed X through Y. It may looks like filtering theory but it is different: there, the observation process Y is given and we have to evaluate where is the signal process X. Here we choose what to observe from X, encapsulated in Y, and it is limited because we want its state space to be small.

Markov intertwinings meet the above requirements. Initially they were developed by Diaconis and Fill [3] in a discrete time and finite state space framework. Let us recall the underlying principle in continuous time, as subsequently extended by Fill [4]. The state spaces V and W are still assumed to finite and we are given L^X the generator of X on V. In the first step, we look for a Markov generator L^Y on W and a Markov kernel Λ from W to V such that the following **intertwining relation** (said to go from L^Y to L^X) holds

$$L^Y\Lambda = \Lambda L^X \tag{1}$$

Ideally, the Markov kernel Λ should be the most "informative" possible, in particular its rank as a matrix should be $\min(\mathrm{card}(V), \mathrm{card}(W))$, which we expect to be $\mathrm{card}(W)$ in the setting of state space reduction. In the second step, when Y is a Markov process generated by L^Y and when its initial law $\mathcal{L}(Y_0)$ satisfies $\mathcal{L}(Y_0)\Lambda = \mathcal{L}(X_0)$, we construct a coupling of X and Y such that (i), (ii) and (iii) are satisfied:

$$\forall\, t \geqslant 0, \qquad \begin{cases} \mathcal{L}(Y_{[0,t]}|X) = \mathcal{L}(Y_{[0,t]}|X_{[0,t]}) \\ \mathcal{L}((X,Y)_{[t,+\infty)}|(X,Y)_{[0,t]}) = \mathcal{L}((X,Y)_{[t,+\infty)}|(X_t,Y_t)) \\ \mathcal{L}(X_t|Y_{[0,t]}) = \Lambda(Y_t,\cdot) \end{cases} \tag{2}$$

where the notation $\mathcal{L}(\cdot|\cdot)$ stands for conditional laws.

To illustrate this procedure, let us come back to the historical example of the top-to-random shuffle due to Aldous and Diaconis [1], in discrete time. The state space is $V \coloneqq \mathcal{S}_N$, the symmetric group on N cards, and the transition of the Markov chain $X \coloneqq (X_n)_{n\in\mathbb{Z}_+}$ corresponds to taking the top card and replacing it at a uniformly chosen position in the deck of cards. Here we adopted the notation

$\mathbb{Z}_+ \coloneqq \{0, 1, 2, 3, ...\}$, while $\mathbb{N}$ will stand for $\{1, 2, 3, ...\}$. The Markov chain $Y \coloneqq (Y_n)_{n\in\mathbb{Z}_+}$ records the position of the card C which initially was at the bottom of the deck, up to the time when it reaches the top of the deck (when C is replaced at random in the deck, by convention the current position of Y is set at 0 and it stays there forever). Thus $W = [\![0, N]\!] \coloneqq \{0, 1, 2, ..., N\}$ and for large $N \in \mathbb{N}$, $\text{card}(W) = N + 1 \ll N! = \text{card}(V)$. In this example, we are interested in the distance in separation of the distribution of X_n at time $n \in \mathbb{Z}_+$ to the invariant measure, which is the uniform distribution υ on $\mathcal{S}_N$. Knowing the "indicative process" $(Y_m)_{m\in[\![0,n]\!]}$, we have a good idea of the distance in separation of the conditional distribution of X_n with υ, in particular when $Y_n = 0$, $\mathcal{L}(X_n|(Y_m)_{m\in[\![0,n]\!]}, Y_n = 0) = \upsilon$. Nevertheless, this example does not convey very well the idea that for the purpose of state space reduction, we are rather looking for Markov kernels Λ whose probability distributions $\Lambda(x, \cdot)$ do not spread much.

Let us come back to the general situation. Given L^X and W, there is usually a lot of Markov generators L^Y and Markov kernels Λ such that (1) is satisfied. So we must be a little more quantitative and wonder about what is a "good intertwining relation". Note that if φ is an eigenfunction associated to an eigenvalue $-\lambda \in \mathbb{C}$ of L^X, then we get $L^Y[\psi] = -\lambda\psi$ with $\psi \coloneqq \Lambda[\varphi]$. Namely, either $\psi = 0$ or ψ is an eigenfunction of L^Y for the eigenvalue $-\lambda$.

Conversely, assume that V and W are finite and that both the Markov generators L^X and L^Y are irreducible. Suppose that some part S of the spectrum of L^Y is included into the spectrum of L^X. Here spectrum has to be understood in an extended sense: it concerns the size of the Jordan block as well as the value of the eigenvalue, and multiplicity is taken into account. Then the computations of [11] enable to find a Markov kernel Λ satisfying (1) and such that the image of Λ contains the eigenspace for L^Y associated to S. There is a trivial instance of this principle: consider the case $S = \{0\}$, which is necessarily included into the spectra of L^X and L^Y. Then we can take for Markov kernel Λ the invariant measure π^X associated to X, namely we consider

$$\forall\ y \in W, \qquad \Lambda(y, \cdot) \quad = \quad \pi^X$$

The general case is obtained by perturbation of this trivial situation. In particular, Λ may be quite small (measured for instance with respect to the image by Λ of the unitary ball of $\mathbb{L}^2(\pi^X)$) and a problem remains to find the largest possible one.

In the folklore, when λ is an eigenvalue of L^X, the smaller (respectively the larger) is $|\lambda|$, the more λ corresponds to global (resp. local) features of the dynamics generated by L^X. For instance in the context of simulated annealing at small temperature, the smallest (non-zero) eigenvalue is directly related to the largest height of a well not containing a fixed global minima of the underlying potential. If one wants to summarize such a process with a two-points dynamics, in some sense, one has to cluster the well with the largest height into a unique point and its complementary set into the other point. This is the most global aspect of the full dynamics (after the fact that the process does not lose mass, which corresponds to the eigenvalue zero). The following eigenvalues correspond to secondary features, see for instance [7] for their geometric description. Another instance of this heuristic is Weyl's law on a compact Riemannian manifold whose total volume is one (see e.g. the book of Taylor [14]): the behavior of the large eigenvalues of the Laplacian mainly depends on dimension of the manifold, since locally, manifolds of the same dimension all look identical.

The two above examples are somewhat asymptotical (one at small temperature and the other at large eigenvalues), nevertheless they suggest that if we are interested in the global behavior of the evolution of X, we should rather look for intertwining relations (1) such that Λ preserves the low lying part of the spectrum of $-L^X$ (while crushing the eigenspaces corresponding to the remaining high lying part). The fact that the above examples are reversible (i.e. self-adjoint) is not relevant, it just insures that the eigenvalues of the corresponding generators are real and non-positive. In general one has to consider the modules of the eigenvalues.

These motivating observations lead us to the following problem. Given L^X and a finite state space W, find a Markov generator L^Y and a Markov kernel Λ from W to V such that (1) holds and the spectrum of L^Y is the low lying spectrum of L^X. In fact this is only the first part of the program described above, since furthermore we would like Λ to be the largest possible and also to couple the processes X and Y. The latter question is very important for applications, since given X, it amounts to knowing how to extract the important information Y from X. In [10] we proposed a way to do it via the introduction of some random mappings in some particular situations where Y is subset-valued (but then the state space of Y can end up being much larger than the state space of X).

Here we will only be concerned with a very special instance of this kind of issue, namely we will consider some famous processes with the same low lying spectrum and we will try to find "nice" intertwining relations between them. This is quite an academic point of view, but it will provide some preliminary insights on what it is possible to do and what is not, especially when the state space V is infinite.

The first example we consider is the **Ehrenfest** family. For $N \in \mathbb{Z}_+$, define on $[\![0,N]\!]$ the Markov generator L_N via

$$\forall\ x \neq x' \in [\![0,N]\!], \qquad L_N(x,x') \ := \ \frac{1}{2}\begin{cases} N-x & \text{, if } x'=x+1 \\ x & \text{, if } x'=x-1 \\ 0 & \text{, otherwise} \end{cases} \tag{3}$$

(the values on the diagonal are such that the row sums all vanish). It is well-known that the spectrum of $-L_N$ is $[\![0,N]\!]$.

The second example is the **Yule** family. For $N \in \mathbb{Z}_+$, consider on $[\![0,N]\!]$ the pure-death generator D_N defined by

$$\forall\ x \neq x' \in [\![0,N]\!], \qquad D_N(x,x') \ := \ \begin{cases} x & \text{, if } x'=x-1 \\ 0 & \text{, otherwise} \end{cases} \tag{4}$$

Since D_N is a lower triangular matrix, its eigenvalues are given by the diagonal, namely $[\![0,N]\!]$ is the spectrum of $-D_N$. Note that the generators of this family are not irreducible, as the associated processes are non-increasing.

There is a reverse family of Yule generators $(\widetilde{D}_N)_{N\in\mathbb{Z}_+}$ given by

$$\forall\ N \in \mathbb{Z}_+,\ \forall\ x \neq x' \in [\![-N,0]\!], \qquad \widetilde{D}_N(x,x') \ := \ \begin{cases} -x+1 & \text{, if } x'=x-1 \\ 0 & \text{, otherwise} \end{cases} \tag{5}$$

The spectrum of $-\widetilde{D}_N$ is also $[\![0,N]\!]$. We have already encountered these generators: up to a shift of the state space and a rescaling of time (after changing from discrete to continuous time), they correspond to the evolution of the last card in the top-to-random shuffle.

The family $(D_N)_{N\in\mathbb{Z}_+}$ admits an infinite version D_∞: it is the pure-death generator on $\mathbb{Z}_+$ whose infinite matrix $(D_\infty(y,y'))_{y,y'\in\mathbb{Z}_+}$ is imposed by its off-diagonal entries via:

$$\forall\ y \neq y' \in \mathbb{Z}_+, \qquad D_\infty(y,y') \ := \ \begin{cases} y & \text{, if } y'=y-1 \\ 0 & \text{, otherwise} \end{cases}$$

From a functional point of view, we see D_∞ as an operator on $\mathbb{C}^{\mathbb{Z}_+}$, via

$$\forall\ f \in \mathbb{C}^{\mathbb{Z}_+},\ \forall\ y \in \mathbb{Z}_+, \qquad D_\infty[f](y) \ = \ y(f(y-1)-f(y))$$

As it can be expected, the spectrum of $-D_\infty$ turns out to be $\mathbb{Z}_+$.

The reverse family $(\widetilde{D}_N)_{N\in\mathbb{Z}_+}$ equally admits an infinite version $\widetilde{D}_\infty$, on the state space $\mathbb{Z}_- := -\mathbb{Z}_+$:

$$\forall\ y \neq y' \in \mathbb{Z}_-, \qquad \widetilde{D}_\infty(y,y') \ := \ \begin{cases} 1-y & \text{, if } y'=y-1 \\ 0 & \text{, otherwise} \end{cases}$$

We will mainly work with the Yule family $(D_N)_{N\in\mathbb{Z}_+\sqcup\{\infty\}}$, since Assertion (f) below reduces the interest of the reverse Yule family (nevertheless, see the considerations at the beginning of Subsection 3.3 and Conjecture 2 at the end of this introduction).

Our last Markov operator is the **Ornstein-Uhlenbeck** operator L acting on $\mathcal{P}$, the space of polynomial functions on $\mathbb{R}$, via

$$\forall\, f \in \mathcal{P},\, \forall\, x \in \mathbb{R}, \qquad L[f](x) \;\coloneqq\; f''(x) - xf'(x) \tag{6}$$

This operator is non-positive and symmetric in $\mathbb{L}^2(\gamma)$, where γ is the standard normal distribution. Thus its Freidrichs extension provides a self-adjoint operator on $\mathbb{L}^2(\gamma)$, that is still denoted by L. The spectrum of $-L$ consists of the eigenvalues $n \in \mathbb{Z}_+$, all of them of multiplicity 1.

The goal of this paper is to show the following assertions (under appropriate integrability assumptions for (b) and (d)):

(a) There are surjective intertwinings from L_N to L_M, for all $M \geqslant N \in \mathbb{Z}_+$.

(b) The only intertwining from L_N to L is trivial, for all $N \in \mathbb{Z}_+$.

(c) There are surjective intertwinings from D_N to L_M, for all $M \geqslant N \in \mathbb{Z}_+$.

(d) The only intertwining from L_N to D_M is trivial, for any $N, M \in \mathbb{Z}_+$.

(e) There are surjective intertwinings from D_N to L for all $N \geqslant 2$, but not for $N = 1$.

(f) The only intertwining from $\widetilde{D}_N$ to L is trivial for all $N \geqslant 1$.

In these statements, an intertwining relation is said to be trivial (respectively surjective) if the corresponding Markov kernel Λ coincides with a probability distribution, i.e. if all its rows are the same (resp. if Λ is surjective). In (a) and (c), we will provide some explicit and quite natural intertwinings. Concerning (e), we will describe all the possible intertwinings for $N = 2$ and $N = 3$, but for $N \geqslant 4$ the argument will only be perturbative, so some room is left for improvements that could lead to a proof of Conjecture 1 below. The non-existence of (b) does not come from the fact we are trying to intertwine jump processes with diffusions, since in [12] we intertwined infinite birth and death processes with Laguerre diffusions in non-trivial ways.

Theorem 3.4 in Biane [2], provides an intertwining relation from the Yule process to the Ornstein-Uhlenbeck process with generator $\widetilde{L} \coloneqq (1/8)\partial^2 - (1/2)\partial \neq L$ (see also Remark 14 in Subsection 3.2 below). We are left wondering whether his group theoretical approach could be used to check the following:

Conjecture 1 There exists a non-trivial intertwining from D_∞ to L.

□

Despite (f), we equally believe in:

Conjecture 2 There exists a non-trivial intertwining from $\widetilde{D}_\infty$ to L.

□

In the next section we study the intertwining relations starting from an ergodic generator, namely (a), (b) and (d). Section 3 deals with the remaining cases, where the intertwining relations starts from an absorbed generator.

Acknowledgments:

The authors are grateful to the Toulouse School of Economics for the invitation of the second author. We are thankful to Philippe Biane for explaining us why there is no contradiction between some of his results [2] and our computations (see Remark 14 below). We are equally grateful to the referee who simplified our first proof of Lemma 17.

2 Intertwinings from an ergodic generator

Here we deal with the points (a), (b) and (d) of the introduction, respectively in the following subsections.

2.1 From Ehrenfest to Ehrenfest

Fix some $N \in \mathbb{Z}_+$. The invariant probability measure π_N associated to the generator L_N defined in (3) is the binomial distribution given by

$$\forall\, x \in [\![0, N]\!], \qquad \pi_N(x) \;\coloneqq\; 2^{-N}\binom{N}{x} \tag{7}$$

This measure is furthermore reversible for L_N, i.e. the operator L_N is self-adjoint in $\mathbb{L}^2(\pi_N)$. It follows that L_N is diagonalizable. It is well-known that the set of eigenvalues of $-L_N$ is $[\![0, N]\!]$, all with multiplicity 1. The eigenspace associated to the eigenvalue $n \in [\![0, N]\!]$ is generated by the Krawtchouk polynomial $K_{N,n}$. These polynomials can be defined via their generating function:

$$\begin{aligned} \forall\, x \in [\![0, N]\!],\ \forall\, z \in \mathbb{C}, \qquad G_N(z, x) \;&\coloneqq\; \sum_{n\in[\![0,N]\!]} K_{N,n}(x)\frac{z^n}{n!} \\ &=\; \left(1 + \frac{z}{2}\right)^x \left(1 - \frac{z}{2}\right)^{N-x} \end{aligned} \tag{8}$$

see for instance Griffiths [6].

Since the set of eigenvalues of L_N is included into the set of eigenvalues of L_{N+1} and the probability measure π_{N+1} charges all the points of $[\![0, N + 1]\!]$, from [11], we get that there exists a Markov kernel Λ_N from $[\![0, N]\!]$ to $[\![0, N + 1]\!]$ such that

$$L_N\Lambda_N \;=\; \Lambda_N L_{N+1} \tag{9}$$

and such that the rank of Λ_N, seen as a $[\![0, N]\!] \times [\![0, N + 1]\!]$-matrix, is $N + 1$.

Let us exhibit a very simple one:

Lemma 3 *Consider the Markov kernel Λ_N from $[\![0, N]\!]$ to $[\![0, N + 1]\!]$ given by*

$$\forall\, x \in [\![0, N]\!],\ \forall\, y \in [\![0, N + 1]\!], \qquad \Lambda_N(x, y) \;\coloneqq\; \begin{cases} 1/2 & \text{, if } y = x + 1 \text{ or } y = x \\ 0 & \text{, otherwise} \end{cases}$$

We have

$$\forall\, n \in [\![0, N + 1]\!], \qquad \Lambda_N[K_{N+1,n}] \;=\; K_{N,n} \tag{10}$$

with the convention that $K_{N,N+1} = 0$ on $[\![0, N]\!]$. In particular the intertwining relation (9) is satisfied.

Proof

We compute the generating function of the family $(\Lambda_N[K_{N+1,n}])_{n\in[\![0,N+1]\!]}$: for any $z \in \mathbb{C}$ and $x \in$

$[\![0,N]\!]$,

$$\begin{aligned}\sum_{n\in[\![0,N+1]\!]}\Lambda[K_{N+1,n}](x)\frac{z^n}{n!} &= \frac{1}{2}\sum_{n\in[\![0,N+1]\!]}(K_{N+1,n}(x)+K_{N+1,n}(x+1))\frac{z^n}{n!}\\ &= \frac{1}{2}\left(\sum_{n\in[\![0,N+1]\!]}K_{N+1,n}(x)\frac{z^n}{n!}+\sum_{n\in[\![0,N+1]\!]}K_{N+1,n}(x+1)\frac{z^n}{n!}\right)\\ &= \frac{1}{2}\left(\left(1+\frac{z}{2}\right)^x\left(1-\frac{z}{2}\right)^{N+1-x}+\left(1+\frac{z}{2}\right)^{x+1}\left(1-\frac{z}{2}\right)^{N-x}\right)\\ &= \frac{1}{2}\left(1+\frac{z}{2}\right)^x\left(1-\frac{z}{2}\right)^{N-x}\left(1-\frac{z}{2}+1+\frac{z}{2}\right)\\ &= \left(1+\frac{z}{2}\right)^x\left(1-\frac{z}{2}\right)^{N-x}\end{aligned}$$

where (8) was taken into account in the second equality. Using again (8), we deduce (10). To get the sought intertwining relation, it is sufficient to check it on the $(K_{N+1,n})_{n\in[\![0,N+1]\!]}$, which is a base of $\mathbb{L}^2(\pi_{N+1})$. We have for any $n\in[\![0,N+1]\!]$,

$$\begin{aligned}L_N\Lambda_N[K_{N+1,n}] &= L_N[K_{N,n}]\\ &= -nK_{N,n}\\ &= -n\Lambda_N[K_{N+1,n}]\\ &= \Lambda_N L_{N+1}[K_{N+1,n}]\end{aligned}$$

■

The relation (9) can be extended to other pairs of Ehrenfest generators. Indeed, writing

$$\forall\ M\leqslant N\in\mathbb{Z}_+,\qquad \Lambda_{M,N} \coloneqq \Lambda_M\Lambda_{M+1}\cdots\Lambda_{N-1}$$

we get that for any $M\leqslant N\in\mathbb{Z}_+$,

$$\forall\ x\in[\![0,M]\!],\forall\ y\in[\![0,N]\!],\qquad \Lambda_{M,N}(x,y) = \begin{cases}2^{M-N}\binom{N-M}{y-x} & \text{, if } x\leqslant y\leqslant x+N-M\\ 0 & \text{, otherwise}\end{cases} \tag{11}$$

By an immediate iteration of (10), we obtain the intertwining relation

$$\forall\ M\leqslant N\in\mathbb{Z}_+,\qquad L_M\Lambda_{M,N} = \Lambda_{M,N}L_N \tag{12}$$

2.2 From Ehrenfest to Ornstein-Uhlenbeck

Consider the family of Hermite polynomials $(h_n)_{n\in\mathbb{Z}_+}$, defined, similarly to (8), via their generating function:

$$\forall\ x\in\mathbb{R},\ \forall\ z\in\mathbb{C},\qquad \sum_{n\in\mathbb{Z}_+}h_n(x)\frac{z^n}{n!} = \exp(zx-z^2/2) \tag{13}$$

For any $n\in\mathbb{Z}_+$, the eigenspace associated to the eigenvalue $-n$ of the Ornstein-Uhlenbeck generator L defined in (6) is generated by h_n.

Fix some $N\in\mathbb{Z}_+$. We are wondering whether there exists a Markov kernel Λ from $[\![0,N]\!]$ to $\mathbb{R}$ such that

$$L_N\Lambda = \Lambda L \tag{14}$$

This equality is understood as holding on $\mathcal{P}$, so it is implicitly assumed that for any $n \in \mathbb{Z}_+$, the probability measure $\Lambda(n, \cdot)$ has moments of all orders.

Since the Gaussian measure γ is invariant for L, (14) is satisfied with $\Lambda = \gamma$, namely with the trivial Markov kernel given by

$$\forall\, n \in \mathbb{Z}_+, \qquad \Lambda(n, \cdot) \;=\; \gamma \tag{15}$$

Indeed, with this Markov kernel, both sides of (14) vanish.

Our goal here is to show that under a strengthened integrability assumption, γ is the unique Markov kernel such that (14) is satisfied.

Let $\mathcal{M}$ be the set of probability measures on $\mathbb{R}$ integrating the mapping $\mathbb{R} \ni x \mapsto \exp(x^2/4)$ and denote by $\mathcal{K}_N$ the set of Markov kernels Λ from $[\![0, N]\!]$ to $\mathbb{R}$ such that $\Lambda(n, \cdot) \in \mathcal{M}$ for all $n \in [\![0, N]\!]$. We have:

Proposition 4 *The only Markov kernel $\Lambda \in \mathcal{K}_N$ such that* (14) *is satisfied is the trivial kernel defined in* (15).

Proof

For any $n \in \mathbb{Z}_+$, consider $\psi_n := \Lambda[h_n]$. Due to (14), we have

$$\begin{aligned} L_N[\psi_n] &= L_N\Lambda[h_n] \\ &= \Lambda L[h_n] \\ &= -n\Lambda[h_n] \\ &= -n\psi_n \end{aligned}$$

It follows that ψ_n is an eigenvector associated to the eigenvalue $-n$ of L_N. According to Subsection 2.1, ψ_n is proportional to the Krawtchouk polynomial $K_{N,n}$, with the convention that $K_{N,n} = 0$ on $[\![0, N]\!]$ for $n > N$. Let $a_n \in \mathbb{R}$ be such that $\psi_n = a_n K_{N,n}$, in particular $a_n = 0$ for $n > N$.

With the help of (13) and Lemma 5 below, which justifies the application of Fubini's lemma, we compute that for any $z \in \mathbb{C}$ and $y \in [\![0, N]\!]$,

$$\int_{\mathbb{R}} \Lambda(y, dx) \exp(zx - z^2/2) \;=\; \sum_{n\in\mathbb{Z}_+} \Lambda[h_n](y)\frac{z^n}{n!} \tag{16}$$

$$\;=\; \sum_{n\in\mathbb{Z}_+} a_n \frac{z^n}{n!} K_{N,n}(y)$$

$$\;=\; \sum_{n\in[\![0,N]\!]} a_n \frac{z^n}{n!} K_{N,n}(y) \tag{17}$$

Conversely, observe that if Λ is a Markov kernel from $\mathbb{Z}_+$ to $\mathbb{R}$ satisfying (17) for all $z \in \mathbb{C}$ and $y \in [\![0, N]\!]$, then (14) is true. Indeed, from (17) we deduce that $\psi_n = a_n K_{N,n}$ for all $n \in \mathbb{Z}_+$, namely Λ maps each eigenspace of L into the corresponding eigenspace of L_N (with the convention that the eigenspaces associated to the eigenvalues $-n$ with $n > N$ are reduced to $\{0\}$) and this is sufficient to insure (14).

Taking into account (13) and the fact that $(h_n)_{n\in\mathbb{Z}_+}$ is an orthogonal family in $\mathbb{L}^2(\gamma)$, we get that for any $n \in \mathbb{Z}_+$,

$$\int h_n(x) \exp(zx - z^2/2)\, \gamma(dx) \;=\; z^n \tag{18}$$

due to the fact that

$$\forall\, n \in \mathbb{Z}_+, \qquad \int h_n^2(x)\, \gamma(dx) \;=\; n!$$

It follows that for any $y \in [\![0, N]\!]$,

$$\int_{\mathbb{R}} \left(\sum_{n \in [\![0,N]\!]} K_{N,n}(y) \frac{a_n}{n!} h_n(x) \right) \exp(zx)\, \gamma(dx) \;=\; \left(\sum_{n \in [\![0,N]\!]} K_{N,n}(y) \frac{a_n}{n!} z^n \right) \exp(z^2/2)$$
$$= \int_{\mathbb{R}} \Lambda(y, dx) \exp(zx)$$

We deduce that for all $y \in [\![0, N]\!]$, we have

$$\Lambda(y, dx) \;=\; \sum_{n \in [\![0,N]\!]} K_{N,n}(y) \frac{a_n}{n!} h_n(x) \gamma(dx) \tag{19}$$

In particular, we must have for all $y \in [\![0, N]\!]$ and a.e. $x \in \mathbb{R}$,

$$\sum_{n \in [\![0,N]\!]} K_{N,n}(y) \frac{a_n}{n!} h_n(x) \;\geqslant\; 0$$

By continuity in x of the left-hand side, this should be true for all $y \in [\![0, N]\!]$ and $x \in \mathbb{R}$.

Let $n_0 \in [\![0, N]\!]$ be the largest integer such that $a_{n_0} \neq 0$. Assume that $n_0 \geqslant 1$. Since for any $n \in \mathbb{Z}_+$, the polynomial h_n has degree n and its highest coefficient is 1, we have that as x goes to $+\infty$,

$$\sum_{n \in [\![0,N]\!]} K_{N,n}(y) \frac{a_n}{n!} h_n(x) \;\sim\; K_{N,n_0}(y) \frac{a_{n_0}}{n_0!} x^{n_0}$$

Since the integral of K_{N,n_0} with respect to π_N is zero (the scalar product in $\mathbb{L}^2(\pi_N)$ of K_{N,n_0} and $K_{N,0} = \mathbb{1}$, the function always taking the value 1, vanishes), we can find $y_-, y_+ \in [\![0, N]\!]$ such that $K_{N,n_0}(y_-) < 0$ and $K_{N,n_0}(y_+) > 0$. Thus we can find $y_0 \in \{y_-, y_+\}$ such that $K_{N,n_0}(y_0) \frac{a_{n_0}}{n_0!} < 0$, namely

$$\lim_{x \to +\infty} \sum_{n \in [\![0,N]\!]} K_{N,n}(y_0) \frac{a_n}{n!} h_n(x) \;=\; -\infty$$

This is in contradiction with the non-negativity of the left-hand-side, so necessarily $n_0 = 0$. We deduce that for all $y \in [\![0, N]\!]$, we have

$$\Lambda(y, dx) \;=\; K_{N,0}(y) \frac{a_0}{0!} h_0(x) \gamma(dx)$$
$$= a_0 \gamma(dx)$$

We get that $a_0 = 1$ and finally $\Lambda = \gamma$.

■

The above proof is not yet complete, since we did not validate the use of Fubini's lemma in (16). This is done in the next result, which will also justify the integrability assumption of Proposition 4.

Lemma 5 *For any given $y \in \mathbb{Z}_+$, the identity* (16) *is justified as soon as $\Lambda(y, \cdot) \in \mathcal{M}$.*

Proof

It is a consequence of Cramer's inequality, see for instance the book of Szegö [13], claiming that there exists a constant $c > 0$ such that

$$\forall\, n \in \mathbb{Z}_+, \forall\, x \in \mathbb{R}, \qquad |h_n(x)| \;\leqslant\; c\sqrt{n!} \exp(x^2/2)$$

Indeed, this bound yields that for any given $z \in \mathbb{C}$,

$$\begin{aligned}
\int_{\mathbb{R}} \Lambda(y,dx) \sum_{n\in\mathbb{Z}_+} \left| h_n(x)\frac{z^n}{n!}\right| &\leqslant c\int_{\mathbb{R}} \Lambda(y,dx) \sum_{n\in\mathbb{Z}_+} \frac{\exp(x^2/4)|z|^n}{\sqrt{n!}} \\
&= c\int_{\mathbb{R}} \Lambda(y,dx)\exp(x^2/4) \sum_{n\in\mathbb{Z}_+} \frac{|z|^n}{\sqrt{n!}} \\
&< +\infty
\end{aligned}$$

■

Remark 6 If one is not interested in Markov kernels, note that the signed kernels defined in (19) always provide intertwining links between Ehrenfest and Ornstein-Uhlenbeck generators.

□

Proposition 4 is an unpleasant fact for the program described in the introduction: consider any irreducible Markov generator G_N on $[\![0,N]\!]$ whose eigenvalues are $-[\![0,N]\!]$, where $N \in \mathbb{Z}_+$ is fixed. Then there is no non-trivial intertwining from G_N to L, namely it is not possible to have a good image (in the sense of the introduction) on $N+1$ points of the Ornstein-Uhlenbeck process. Indeed, assume on the contrary that we have such an non-trivial intertwining $G_N\Lambda = \Lambda L$. Since L_N and G_N have the same spectrum and are irreducible, we deduce from [11] there exists a Markov kernel $\widetilde{\Lambda}$ from $[\![0,N]\!]$ into itself such that $L_N\widetilde{\Lambda} = \widetilde{\Lambda}G_N$ and such that $\widetilde{\Lambda}$ is an invertible matrix. We would then obtain the non-trivial intertwining relation $L_N\widetilde{\Lambda}\Lambda = \widetilde{\Lambda}\Lambda L$, a contradiction.

2.3 From Ehrenfest to Yule

In general the Markov kernel entering into an intertwining relation transports the invariant measure of the first generator into the invariant measure of the second generator. So when the second generator is absorbing with a unique absorbing point, the Markov kernel must be trivial and equal to the Dirac mass on the absorbing point.

Let us illustrate this principle on intertwining relations between Ehrenfest and Yule processes. For any given $N, M \in \mathbb{Z}_+$, assume that Λ is a Markov kernel from $[\![0,N]\!]$ to $[\![0,M]\!]$ such that

$$L_N\Lambda = \Lambda D_M \tag{20}$$

Let π_N, described in (7), be the invariant probability of L_N, for instance seen as a row vector. Multiplying (20) on the left by π_N, we get

$$\pi_N\Lambda D_M = \pi_N L_N\Lambda = 0$$

Thus the probability $\pi_N\Lambda$ on $[\![0,M]\!]$ is an invariant probability for D_M. Since the Markov processes associated to D_M all end up being absorbed at 0, necessarily $\pi_N\Lambda = \delta_0$. Taking into account that π_N charges all the points of $[\![0,N]\!]$, we get that

$$\forall\, x \in [\![0,N]\!], \qquad \Lambda(x,\cdot) = \delta_0$$

namely Λ is trivial.

The above arguments are immediately extended to the case where $M = \infty$ and also to intertwining relations of the form $L\Lambda = \Lambda D_M$, with $M \in \mathbb{Z}_+ \sqcup \{\infty\}$. We conclude that a.e. in $x \in \mathbb{R}$, $\Lambda(x,\cdot) = \delta_0$. However for $L\Lambda = \Lambda D_M$ to make sense a priori, we must assume that Λ transforms $\mathbb{R}^{[\![0,M]\!]}$ into functions that are at least continuous (or seen as elements of $\mathbb{L}^2(\gamma)$, if we consider the Friedrich extension), so that $\Lambda(x,\cdot) = \delta_0$ holds for all $x \in \mathbb{R}$ (or $\Lambda = \delta_0$ in the $\mathbb{L}^2(\gamma)$ context).

3 Intertwinings from an absorbed generator

In the last subsection we have seen there is usually no non-trivial intertwining relation from an ergodic generator to an absorbed generator. It is not true in the reverse direction, as shown by the top-to-random card shuffle example of Aldous and Diaconis [1], and more generally such intertwining relations were exploited by Diaconis and Fill [3] to construct strong stationary times. Here we deal with the points (c), (e) and (f) of the introduction, respectively in the following subsections.

3.1 From Yule to Ehrenfest

The intertwining relation from Yule to Ehrenfest is well-known, as well as the relations with the discrete hypercube random walks, see Example 4.38 of Diaconis and Fill [3]. We give here a direct proof in the spirit of this paper, via generating functions.

Fix some $N \in \mathbb{Z}_+$. We have seen that the eigenvalues of the Yule generator D_N defined in (4) are the elements of $-[\![0, N]\!]$. Let us compute the corresponding eigenvectors:

Lemma 7 *For all $n \in [\![0, N]\!]$, the eigenspace associated to the eigenvalue n of $-D_N$ is generated by the function φ_n given by*

$$\forall\ x \in [\![0, N]\!], \qquad \varphi_n(x) \ := \ \binom{x}{n} \tag{21}$$

with the usual convention that $\binom{m}{n} = 0$ for any $m < n$.

Proof

Fix $n \in [\![0, N]\!]$ and let us check that for any $x \in [\![0, N]\!]$,

$$D_N[\varphi_n](x) \ = \ -n\varphi_n(x) \tag{22}$$

For $x \in [\![0, n-1]\!]$, both sides are zero, so the equality holds.

For $x = n$, the left-hand side is equal to $n(\varphi_n(n-1) - \varphi_n(n)) = -n\varphi_n(n)$, so the equality holds.

Assume that (22) is true for some $x \geqslant n$ and let us show it is also true with x replaced by $x + 1$, as long as $x + 1 \leqslant N$. We have

$$\begin{aligned}
D_N[\varphi_n](x+1) &= (x+1)(\varphi_n(x) - \varphi_n(x+1)) \\
&= (x+1)\left(\binom{x}{n} - \binom{x+1}{n}\right) \\
&= \frac{x+1}{n!}\big(x(x-1)\cdots(x-n+1) - (x+1)x\cdots(x-n+2)\big) \\
&= \frac{x+1}{n!}x(x-1)\cdots(x-n+2)\big(x-n+1-(x+1)\big) \\
&= -n\frac{x+1}{n!}x(x-1)\cdots(x-n+2) \\
&= -n\varphi_n(x+1)
\end{aligned}$$

which completes the proof.

■

We proceed by computing the generating function of the family $(2^{-n}n!\varphi_n)_{n\in[\![0,N]\!]}$:

Lemma 8 *We have*

$$\forall\ z \in \mathbb{C},\ \forall\ x \in [\![0, N]\!], \qquad \sum_{n\in[\![0,N]\!]} 2^{-n}n!\varphi_n(x)\frac{z^n}{n!} \ = \ \left(1 + \frac{z}{2}\right)^x$$

Proof

Indeed, we compute that

$$\begin{aligned}\sum_{n\in[\![0,N]\!]} 2^{-n}n!\varphi_n(x)\frac{z^n}{n!} &= \sum_{n\in[\![0,N]\!]} \varphi_n(x)(2^{-1}z)^n \\ &= \sum_{n\in[\![0,N]\!]} \binom{x}{n}(2^{-1}z)^n \\ &= \left(1+\frac{z}{2}\right)^x\end{aligned}$$

■

Consider the Markov kernel $\widehat{\Lambda}_N$ from $[\![0,N]\!]$ to $[\![0,N]\!]$ defined by

$$\forall\ x,y\in[\![0,N]\!],\qquad \widehat{\Lambda}_N(x,y) \ :=\ 2^{x-N}\binom{N-x}{y-x} \ =\ \Lambda_{x,N}(x,y)$$

with the notation introduced in (11). In particular the support of $\Lambda(x,\cdot)$ is $[\![x,N]\!]$.

Its introduction is motivated by:

Proposition 9 *We have*

$$\forall\ n\in[\![0,N]\!],\qquad \widehat{\Lambda}_N[K_{N,n}] \ =\ 2^{-n}n!\varphi_n$$

and in particular,

$$D_N\widehat{\Lambda}_N \ =\ \widehat{\Lambda}_N L_N \tag{23}$$

Proof

For the first equality, it is sufficient to check the equality of the generating functions associated to the families $(\widehat{\Lambda}_N[K_{N,n}])_{n\in[\![0,N]\!]}$ and $(2^{-n}n!\varphi_n)_{n\in[\![0,N]\!]}$, namely

$$\forall\ z\in\mathbb{C},\ \forall\ x\in[\![0,N]\!],\qquad \sum_{n\in[\![0,N]\!]}\widehat{\Lambda}_N[K_{N,n}](x)\frac{z^n}{n!} \ =\ \left(1+\frac{z}{2}\right)^x$$

With the notation of (8), this equality is equivalent to

$$\forall\ z\in\mathbb{C},\ \forall\ x\in[\![0,N]\!],\qquad \widehat{\Lambda}_N[G_N(z,\cdot)](x) \ =\ \left(1+\frac{z}{2}\right)^x \tag{24}$$

so let us compute the l.h.s.: for any $z\in\mathbb{C}$ and $x\in[\![0,N]\!]$,

$$\begin{aligned}\widehat{\Lambda}_N[G_N(z,\cdot)](x) &= 2^{x-N}\sum_{y\in[\![x,N]\!]}\binom{N-x}{y-x}G_N(z,y) \\ &= 2^{x-N}\sum_{y\in[\![x,N]\!]}\binom{N-x}{y-x}\left(1+\frac{z}{2}\right)^y\left(1-\frac{z}{2}\right)^{N-y} \\ &= 2^{x-N}\left(1+\frac{z}{2}\right)^x\sum_{y\in[\![x,N]\!]}\binom{N-x}{y-x}\left(1+\frac{z}{2}\right)^{y-x}\left(1-\frac{z}{2}\right)^{N-y} \\ &= 2^{x-N}\left(1+\frac{z}{2}\right)^x\left(1+\frac{z}{2}+1-\frac{z}{2}\right)^{N-x} \\ &= \left(1+\frac{z}{2}\right)^x\end{aligned}$$

This enables us to conclude the validity of (24).

Concerning the intertwining equality announced in the lemma, it is sufficient to check it on the basis $(K_{N,n})_{n\in[\![0,N]\!]}$ of $\mathbb{R}^{[\![0,N]\!]}$. Indeed, for any $n\in[\![0,N]\!]$, we have

$$\begin{aligned}\widehat{\Lambda}_N L_N[K_{N,n}] &= \Lambda[-nK_{N,n}]\\ &= -n\Lambda[K_{N,n}]\\ &= -n2^{-n}n!\varphi_n\\ &= D_N[2^{-n}n!\varphi_n]\\ &= D_N\widehat{\Lambda}_N[K_{N,n}]\end{aligned}$$

■

Note that the kernel $\widehat{\Lambda}_N$ is surjective. For integers $M\leqslant N$, one deduces a surjective intertwining relation

$$D_M\widehat{\Lambda}_{M,N} = \widehat{\Lambda}_{M,N}L_N \tag{25}$$

from (23) by restriction from $[\![0,N]\!]$ to $[\![0,M]\!]$. Formally, consider $I_{M,N}$ is the natural imbedding from $[\![0,M]\!]$ into $[\![0,N]\!]$, seen as a Markov kernel. Since D_N is a lower triangular matrix, we have the simple intertwining relation $I_{M,N}D_N = D_M I_{M,N}$. Thus multiplying (23) on the left by $I_{M,N}$, we obtain (25) with the surjective kernel $\widehat{\Lambda}_{M,N} \coloneqq I_{M,N}\widehat{\Lambda}_N$.

Alternatively, we can start from $D_M\widehat{\Lambda}_M = \widehat{\Lambda}_M L_M$ that we multiply on the right by the Markov kernel $\Lambda_{M,N}$ defined in (11). Taking into account (12), we end up with (25) with $\widehat{\Lambda}_{M,N} \coloneqq \widehat{\Lambda}_N\Lambda_{M,N}$. It is not difficult to see that the latter Markov kernel coincides with the former one, i.e. $\widehat{\Lambda}_M\Lambda_{M,N} = I_{M,N}\widehat{\Lambda}_N$. Indeed, it is sufficient to check this identity when it is applied to the family of Krawtchouk polynomials $(K_{N,n})_{n\in[\![0,N]\!]}$.

3.2 From Yule to Ornstein-Uhlenbeck

We start by considering the finite Yule generators and thus fix some $N\in\mathbb{Z}_+$. Recall that the set $\mathcal{K}_N$ of Markov kernels was defined before Proposition 4. We are interested in

$$\mathcal{L}_N \coloneqq \{\Lambda\in\mathcal{K}_N : D_N\Lambda = \Lambda L\}$$

To any $a\coloneqq(a_n)_{n\in[\![0,N]\!]}$, we associate the mapping $\lambda_a : [\![0,N]\!]\times\mathbb{R}\to\mathbb{R}$ and the signed kernel Λ_a from $[\![0,N]\!]$ to $\mathbb{R}$ via

$$\forall\, y\in[\![0,N]\!],\ \forall\, x\in\mathbb{R},\qquad \lambda_a(y,x) \coloneqq \sum_{n\in[\![0,y]\!]}\frac{a_n}{n!}\binom{y}{n}h_n(x) \tag{26}$$

$$\forall\, y\in[\![0,N]\!],\qquad \Lambda_a(y,dx) \coloneqq \lambda_a(y,x)\,\gamma(dx) \tag{27}$$

Finally, consider $\mathcal{A}_N$ the set of $a\coloneqq(a_n)_{n\in[\![0,N]\!]}\in\mathbb{R}^{N+1}$ with $a_0=1$ such that

$$\forall\, y\in[\![0,N]\!],\ \forall\, x\in\mathbb{R},\qquad \lambda_a(y,x) \geqslant 0 \tag{28}$$

The interest of these definitions is:

Proposition 10 *We have*

$$\mathcal{L}_N = \{\Lambda_a : a\in\mathcal{A}_N\}$$

As in Remark 6, the following proof shows that the signed kernels Λ_a always provide intertwining links between the finite Yule and Ornstein-Uhlenbeck generators.

Proof

The arguments are similar to those of the proof of Proposition 4. Consider $\Lambda \in \mathcal{L}_N$ and for $n \in [\![0, N]\!]$ define $\psi_n := \Lambda[h_n]$. The intertwining relation $D_N\Lambda = \Lambda L$ implies that $D_N[\psi_n] = -n\psi_n$ and according to Lemma 7, ψ_n is proportional to φ_n. Denote $a_n \in \mathbb{R}$ such that $\psi_n = a_n\varphi_n$. Note that when $n = 0$, we have $\psi_0 = \mathbb{1} = \varphi_0$, so that $a_0 = 1$.

With the help of (13) and Lemma 5, we get that for any $z \in \mathbb{C}$ and $y \in \mathbb{Z}_+$,

$$\begin{aligned}\int_{\mathbb{R}} \Lambda(y, dx) \exp(zx - z^2/2) &= \sum_{n\in\mathbb{Z}_+} \Lambda[h_n](y)\frac{z^n}{n!} \\ &= \sum_{n\in\mathbb{Z}_+} a_n\frac{z^n}{n!}\varphi_n(y) \\ &= \sum_{n\in\mathbb{Z}_+} a_n\frac{z^n}{n!}\binom{y}{n} \\ &= \sum_{n\in[\![0,y]\!]} a_n\frac{z^n}{n!}\binom{y}{n} \qquad (29)\end{aligned}$$

Conversely, observe that if Λ is a Markov kernel from $\mathbb{Z}_+$ to $\mathbb{R}$ satisfying (29) for all $z \in \mathbb{C}$ and $y \in \mathbb{Z}_+$, then the intertwining relation $D_N\Lambda = \Lambda L$ holds. Indeed, from (29) we deduce that $\psi_n = a_n\varphi_n$ for all $n \in [\![0, N]\!]$, namely Λ maps each eigenspace of L into the corresponding eigenspace of D (again with the convention that the eigenspaces associated to the eigenvalues $-n$ with $n > N$ are reduced to $\{0\}$) and this is sufficient to insure that $D_N\Lambda = \Lambda L$.

Taking into account (18), it appears that for any $z \in \mathbb{C}$ and $y \in \mathbb{Z}_+$,

$$\int_{\mathbb{R}} \Lambda(y, dx) \exp(zx) = \int_{\mathbb{R}} \Lambda_a(y, dx) \exp(zx)$$

with $a := (a_n)_{n\in[\![0,N]\!]}$, namely $\Lambda = \Lambda_a$. Since Λ is a non-negative kernel, we get that $a \in \mathcal{A}_N$. Conversely, when $a \in \mathcal{A}_N$, it remains to check that the total mass of $\Lambda_a(y, \cdot)$ is 1, for any $y \in [\![0, N]\!]$. It comes from the fact that $a_0 = 1$ and that

$$\forall\, n \in [\![1, N]\!], \qquad \int h_n\, d\gamma = 0$$

■

Here are two observations about $\mathcal{A}_N$:

Lemma 11 *Consider $a := (a_n)_{n\in[\![0,N]\!]} \in \mathcal{A}_N$. Then for any $n \in [\![0, N]\!]$, $a_n = 0$ if n is odd and $a_n \geqslant 0$ if n is even. Furthermore, for any $M \in [\![0, N]\!]$, we have $(a_n)_{n\in[\![0,M]\!]} \in \mathcal{A}_M$.*

Proof

For $y \in [\![0, N]\!]$ odd, $\lambda_a(y, x)$ cannot stay non-negative as x goes to $\pm\infty$ if the coefficient of higher order a_y is non-zero, so we must have $a_y = 0$. For $y \in [\![0, N]\!]$ even, $\lambda_a(y, x)$ cannot stay non-negative if the coefficient of higher order a_y is negative, so we must have $a_y \geqslant 0$.

The second assertion is an immediate consequence of the fact that for $a \in \mathbb{R}^{N+1}$, $y \in [\![0, N]\!]$ and $x \in \mathbb{R}$, the definition of the quantity $\lambda_a(y, x)$ does not depend on N.

■

It follows from the second part of Lemma 11 that for any $M \leqslant N \in \mathbb{Z}_+$, an element of $\mathcal{L}_N$ can be seen as a element of $\mathcal{L}_M$ by restriction, namely we have

$$\Lambda \in \mathcal{L}_N \quad \Longrightarrow \quad I_{M,N}\Lambda \in \mathcal{L}_M \tag{30}$$

where we recall that $I_{M,N}$ is the natural imbedding from $[\![0, M]\!]$ into $[\![0, N]\!]$. To see that $\mathcal{L}_M$ can be strictly larger than $I_{M,N}\mathcal{L}_N$, let us compute the first sets $\mathcal{L}_N$. It will appear that $\mathcal{L}_2$ is strictly included into $I_{2,3}\mathcal{L}_3$.

From Proposition 10 and Lemma 11, we get that $\mathcal{L}_0 = \{\gamma\}$ (the only possible motion on a singleton is to stay still, so an intertwining relation from a singleton state space is equivalent to the existence of an invariant measure for the second Markov generator) and $\mathcal{L}_1 = \{\gamma\}$. The next two cases $\mathcal{L}_2$ and $\mathcal{L}_3$ are no longer singletons:

Lemma 12 *We have*

$$\mathcal{L}_2 \quad = \quad \Big\{\Lambda \in \mathcal{K}_2 \,:\, \Lambda(0,\cdot) = \Lambda(1,\cdot) = \gamma,\, \Lambda(2,dx) = (1 + a_2h_2(x)/2)\gamma(dx),\ \textit{with } a_2 \in [0,2]\Big\}$$

and

$$\begin{aligned}\mathcal{L}_3 \quad = \quad &\Big\{\Lambda \in \mathcal{K}_3 \,:\, \Lambda(0,\cdot) = \Lambda(1,\cdot) = \gamma,\, \Lambda(2,dx) = (1 + a_2h_2(x)/2)\gamma(dx),\\ &\Lambda(3,dx) = (1 + 3a_2h_2(x)/2)\gamma(dx),\ \textit{with } a_2 \in [0,2/3]\Big\}\end{aligned}$$

Proof

From Proposition 10 and Lemma 11, $\mathcal{L}_2$ and $\mathcal{L}_3$ must have the above form, except that for $\mathcal{L}_2$ (respectively $\mathcal{L}_3$) the condition on $a_2 \geqslant 0$ is that $\mathbb{R} \ni x \mapsto 1 + a_2h_2(x)/2$ (resp. $\mathbb{R} \ni x \mapsto 1 + a_2h_2(x)/2$ and $\mathbb{R} \ni x \mapsto 1 + 3a_2h_2(x)/2$) must stay non-negative.

Since $h_2(x) = x^2 - 1$ for all $x \in \mathbb{R}$, these conditions end up being equivalent to $a_2 \in [0,2]$ for $\mathcal{L}_2$ and $a_2 \in [0,2/3]$ for $\mathcal{L}_3$.

■

More generally, consider $N \in \mathbb{Z}_+$ with $N \geqslant 2$. Since each of the mappings h_{2n}, for $n \in \mathbb{Z}_+$, is bounded below, we see that we can find $(a_{2n})_{n\in[\![1,\lfloor N/2\rfloor]\!]}$ small enough so that the corresponding Markov kernel belongs to $\mathcal{L}_N$. In particular, $\mathcal{L}_N$ is not reduced to a singleton, as announced in (e) of the introduction.

There is no difficulty in replacing the finite sets $[\![0, N]\!]$, for $N \in \mathbb{Z}_+$, by $\mathbb{Z}_+$. The only supplementary information we need is the following result, extending Lemma 7:

Lemma 13 *There exist $l \in \mathbb{C}$ and $f \in \mathbb{C}^{\mathbb{Z}_+}\setminus\{0\}$ such that $D_\infty[f] = -lf$ if and only if $l \in \mathbb{Z}_+$ and f is proportional to $\varphi_l \in \mathbb{C}^{\mathbb{Z}_+}$ that was defined in* (21) *as*

$$\forall\ y \in \mathbb{Z}_+, \qquad \varphi_l(y) \ \coloneqq\ \binom{y}{l}$$

Proof

Since f is not vanishing everywhere, there exists $N \in \mathbb{Z}_+$ such that $f(N) \neq 0$. It follows that $I_N[f] \in \mathbb{C}^{[\![0,N]\!]}\setminus\{0\}$, where I_N is the natural embedding of $[\![0, N]\!]$ into $\mathbb{Z}_+$. Since $D_NI_N = I_ND$, we deduce that

$$D_N[I_N[f]] \quad = \quad -lI_N[f]$$

namely $-l$ is an eigenvalue of D_N, thus $l \in [\![0, N]\!]$. Furthermore, according to Lemma 7, f is proportional to φ_l on $[\![0, N]\!]$, say $f = z\varphi_l$ on $[\![0, N]\!]$ for some $z \in \mathbb{C}$. Using the relation $D_\infty[f] = -lf$, we show by an iteration on $n \in \mathbb{Z}_+$ that $f = z\varphi_l$ on $[\![0, N+n]\!]$. It follows that $f = z\varphi_l$ on $\mathbb{Z}_+$. Conversely, there is no difficulty in checking that $D_\infty[\varphi_l] = -l\varphi_l$ for $l \in \mathbb{Z}_+$.

■

Next, similarly to the beginning of this subsection, we define $\mathcal{K}_\infty$ and $\mathcal{L}_\infty$ as convex sets of Markov kernels from $\mathbb{Z}_+$ to $\mathbb{R}$. With the help of Lemma 13, the proof of Proposition 10 is still valid and shows that

$$\mathcal{L}_\infty = \{\Lambda_a : a \in \mathcal{A}_\infty\}$$

where Λ_a and $\mathcal{A}_\infty$ are defined as above, but with $[\![0, N]\!]$ replaced by $\mathbb{Z}_+$ in (26), (27) and (28). The problem is that it is not clear that $\mathcal{A}_\infty$ is not reduced to the singleton $\{(\delta_{0,n})_{n\in\mathbb{Z}_+}\}$, which is equivalent to say there is non-trivial intertwinings from D_∞ to L.

The compatibility property (30) leads to the following characterization of $\mathcal{L}_\infty$:

$$\forall\ \Lambda \in \mathcal{K}_\infty, \qquad \Lambda \in \mathcal{L}_\infty \iff \forall\ N \in \mathbb{Z}_+,\ I_N\Lambda \in \mathcal{L}_N$$

Remark 14 The latter property with $N = 1$ and thus $\mathcal{L}_1 = \{\gamma\}$ may seem to contradict the first part of Theorem 3.4 in Biane [2], providing an intertwining from the Yule process to the Ornstein-Uhlenbeck process with a Markov kernel $\Lambda \in \mathcal{L}_\infty$ satisfying

$$\Lambda(1, dx) = h_1^2(x)\,\gamma(dx) = x^2\,\gamma(dx)$$

As explained to us by Philippe Biane, there is no contradiction, because he is considering the Ornstein-Uhlenbeck process generated by $\widetilde{L} := (1/8)\partial^2 - (x/2)\partial$ and not by L. The spectrum of $-\widetilde{L}$ is $\mathbb{Z}_+/2$ and his Markov kernel Λ kills all the eigenvalues of the form $k/2$ with $k \in \mathbb{N}$ odd. As a consequence, since the even Hermite polynomial are not reduced to zero by Λ, his relation can be seen as a intertwining from the Yule process to the square of an Ornstein-Uhlenbeck process.

□

We are left wondering if Conjecture 1 is true, i.e. whether $\mathcal{L}_\infty$ is reduced to a singleton.

Remark 15 Let $(a_{2n})_{n\in\mathbb{Z}_+}$ be a sequence of non-negative numbers, with $a_0 = 1$. Assume there is only a finite number of elements of this sequence that are non-zero. Then for the mapping

$$\mathbb{R} \ni x \mapsto \sum_{n\in\mathbb{Z}_+} \frac{a_{2n}}{(2n)!}\binom{y}{n} h_{2n}(x)$$

to remain non-negative for all $y \in \mathbb{Z}_+$, we must have $a_{2n} = 0$ for all $n \in \mathbb{N}$. Indeed, suppose on the contrary that there exists $N \in \mathbb{N}$ with $a_{2N} > 0$ and consider N the largest such positive integer. There exists some $x_0 \in \mathbb{R}$ with $h_{2N}(x_0) < 0$ (since $\int h_{2N}\,d\gamma = 0$). The quantity $\sum_{n\in\mathbb{Z}_+} \frac{a_{2n}}{(2n)!}\binom{y}{n} h_{2n}(x_0) = \sum_{n\in[\![0,N]\!]} \frac{a_{2n}}{(2n)!}\binom{y}{n} h_{2n}(x_0)$ behaves like $\frac{a_{2N}}{(2N)!}\binom{y}{N} h_{2n}(x_0) < 0$ for large $y \in \mathbb{Z}_+$, because $\binom{y}{N} \sim y^N/(N!)$. This is a contradiction justifying the above assertion.

Thus any sequence $a \in \mathcal{A}_\infty \backslash \{0\}$ must have an infinite number of non-zero elements.

□

Let us discuss the probabilistic implications of the previous considerations. Fix $N \in \mathbb{Z}_+$. For $\Lambda \in \mathcal{L}_N$, consider $\mathfrak{P}_N(\Lambda)$ the convex hull generated by $\{\Lambda(y, \cdot) : y \in [\![0, N]\!]\}$. Choose some $\mu_0 \in \mathfrak{P}_N(\Lambda)$ and m_0 a probability measure on $[\![0, N]\!]$ such that $m_0\Lambda = \mu_0$. Let $X := (X_t)_{t\geqslant 0}$ be an Ornstein-Uhlenbeck diffusion with initial law μ_0 and let $Y := (Y_t)_{t\geqslant 0}$ be a Yule jump process with generator D_N and starting from m_0. From the general theory of Markov intertwinings developed by Diaconis and Fill [3] (see [9] for an example of technical extension to a one-dimensional diffusion context), it is possible to construct a coupling of X and Y such that (2) is satisfied. As a consequence, the stopping time

$$\tau \quad := \quad \inf\{t \geqslant 0 \,:\, Y_t = 0\}$$

is a **strong stationary time** for X, namely it is a finite randomized stopping time for X such that τ and X_τ are independent and X_τ is distributed according to γ.

Note that the law of τ is a mixture of the distributions $\mathcal{E}(1) * \mathcal{E}(2) * \cdots * \mathcal{E}(n)$, for $n \in [\![0, N]\!]$, where $\mathcal{E}(n)$ stands for the exponential law of parameter n and $*$ for the convolution. In particular the law of τ is stochastically dominated by $\mathcal{E}(1) * \mathcal{E}(2) * \cdots * \mathcal{E}(N)$. Recall that the **separation** $\mathfrak{s}(\mu, \mu')$ between two probability distributions on a same measurable space is defined by

$$\mathfrak{s}(\mu, \mu') \quad := \quad \operatorname*{ess\,sup}_{\mu'} 1 - \frac{d\mu}{d\mu'}$$

where $d\mu/d\mu'$ is the Radon-Nykodim density of μ with respect to μ'. The above observations (see Diaconis and Fill [3] for the general argument) lead to the following quantitative estimates on the convergence to equilibrium for the Ornstein-Uhlenbeck diffusion in the total variation and separation senses:

$$\begin{aligned} \forall\, t \geqslant 0, \qquad \|\mathcal{L}(X_t) - \gamma\|_{\mathrm{tv}} &\leqslant \mathfrak{s}(\mathcal{L}(X_t), \gamma) \\ &\leqslant \mathbb{P}[\tau > t] \\ &\leqslant \mathcal{E}(1) * \mathcal{E}(2) * \cdots * \mathcal{E}(N)((t, +\infty)) \end{aligned}$$

under the condition that the initial law $\mathcal{L}(X_0)$ belongs to

$$\mathfrak{P}_N \quad := \quad \bigcup_{\Lambda \in \mathcal{L}_N} \mathfrak{P}_N(\Lambda)$$

Considering

$$\mathfrak{P} \quad := \quad \bigcup_{n \in \mathbb{Z}_+} \mathfrak{P}_N$$

we get that when the initial distribution of X belongs to $\mathfrak{P}$, there exists a strong stationary time for X. It was proven in [9] this is not true for all initial distributions, in particular for those with compact support. Nevertheless, one can wonder if $\mathfrak{P}$ would not be dense, e.g. in the total variation sense, in the set of probability measures on $\mathbb{R}$ absolutely continuous with respect to γ. Other natural questions are:

- is the set $\mathfrak{P}_N$ convex and in this case what are the extremal points?
- is the sequence $(\mathfrak{P}_N)_{N \in \mathbb{Z}_+}$ non-decreasing?
- what is the link between $\mathfrak{P}$ and $\mathfrak{P}_\infty := \bigcup_{\Lambda \in \mathcal{L}_\infty} \mathfrak{P}_\infty(\Lambda)$?, where for $\Lambda \in \mathcal{L}_\infty$, $\mathfrak{P}_\infty(\Lambda)$ is the set of $m\Lambda$, with m a probability measure on $\mathbb{Z}_+$.

3.3 On the reverse Yule family

Under the questions at the end of last subsection lies the interrogation: is it always possible to slightly modify the initial condition of an Ornstein-Uhlenbeck diffusion to insure the existence of a strong stationary time? A reverse question is: are there strong times which bring the diffusion close to its equilibrium? Recall that a **strong time** is a finite randomized stopping time τ for X such that τ and X_τ are independent. The reverse Yule family seems more appropriate to construct strong times that deal first with the low lying eigenvalues and next with the high lying eigenvalues: the corresponding jump processes begin by encountering the rate 1, then the rate 2, etc. It was the opposite with the Yule processes of the previous sections. Thus in the spirit of the motivations described in the introduction, reverse Yule processes should be preferable. Note this point of view can also be found in Fill [5] and in [8], but in finite settings. Unfortunately, due to (f) of the introduction, this direction ends up being not so relevant, except if Conjecture 2 was to be true.

Let us try to extend the analysis done for the Yule family to the reverse Yule family. Our first task is to compute the eigenvectors of the reverse Yule family.

Lemma 16 *Fix $N \in \mathbb{Z}_+$. For all $n \in [\![N]\!] := [\![1, N]\!]$, the eigenspace associated to the eigenvalue $-n$ of $\widetilde{D}_N$, defined in (5), is generated by the function $\widetilde{\varphi}_n$ given by*

$$\forall\, y \in [\![0, N]\!], \qquad \widetilde{\varphi}_n(y) \;:=\; (-1)^y \binom{n-1}{-y}$$

For $n = 0$, it is sufficient to take $\widetilde{\varphi}_0 = \mathbb{1}$.

Proof

For $n = 0$, it is obvious that $\widetilde{D}_N[\mathbb{1}] = 0$. For $n \in [\![N]\!]$, the relation $\widetilde{D}_N[\widetilde{\varphi}_n] = -n\widetilde{\varphi}_n$ leads to an iteration on the values of $\widetilde{\varphi}_n$:

$$\forall\, y \in [\![-(N-1), 0]\!], \qquad \widetilde{\varphi}_n(y-1) \;=\; \frac{1-n-y}{1-y}\widetilde{\varphi}_n(y)$$

which gives the announced function, when we take $\widetilde{\varphi}_n(0) = 1$.

■

Remark that $\widetilde{\varphi}_n(y) = 0$ as soon as $y \leqslant -n$, in particular all $\widetilde{\varphi}_n$, for $n \in [\![N]\!]$, vanish at $-N$. This comes from the fact that $-N$ is absorbing for $\widetilde{D}_N$. Due to this property, the restriction of $\widetilde{D}_\infty$ to $[\![-N, 0]\!]$ is different from $\widetilde{D}_N$, but only at the entry $(-N, -N)$. So the proof by restriction of Lemma 13 cannot be applied directly, nevertheless we get a similar result (either by a direct proof or by taking into account that $\widetilde{\varphi}_n(-N) = 0$ for all $n \in [\![N]\!]$): the spectrum of $-\widetilde{D}_\infty$ is $\mathbb{Z}_+$, $\mathbb{1}$ spans the eigenspace associated to 0 and for $n \in \mathbb{N}$, the eigenspace associated to n is generated by the extension to $\mathbb{Z}_-$ of the previous function $\widetilde{\varphi}_n$, namely:

$$\forall\, y \in \mathbb{Z}_-, \qquad \widetilde{\varphi}_n(y) \;:=\; (-1)^y \binom{n-1}{-y}$$

As in the beginning of Subsection 3.2, consider $\widetilde{\mathcal{K}}_N$ the set of Markov kernels from $[\![-N, 0]\!]$ to $\mathbb{R}$ such that for all $y \in [\![-N, 0]\!]$, $\Lambda(y, \cdot) \in \mathcal{M}$, and

$$\widetilde{\mathcal{L}}_N \;:=\; \{\Lambda \in \widetilde{\mathcal{K}}_N \,:\, \widetilde{D}_N\Lambda = \Lambda L\}$$

Introduce for any $a := (a_n)_{n\in[\![0,N]\!]}$, the mapping $\widetilde{\lambda}_a \,:\, [\![-N, 0]\!] \times \mathbb{R} \to \mathbb{R}$ and the signed kernel $\widetilde{\Lambda}_a$ from $[\![-N, 0]\!]$ to $\mathbb{R}$ via

$$\forall\, y \in [\![-N, 0]\!],\, \forall\, x \in \mathbb{R}, \qquad \widetilde{\lambda}_a(y, x) \;:=\; \sum_{n\in[\![-y+1,N]\!]} \frac{a_n}{n!}\widetilde{\varphi}_n(y)h_n(x) \tag{31}$$

$$\forall\, y \in [\![-N, 0]\!], \qquad \widetilde{\Lambda}_a(y, dx) \;:=\; \widetilde{\lambda}_a(y, x)\,\gamma(dx) \tag{32}$$

Finally, write $\widetilde{\mathcal{A}}_N$ for the set of $a := (a_n)_{n\in[\![0,N]\!]} \in \mathbb{R}^{N+1}$ with $a_0 = 1$ such that

$$\forall\, y \in [\![-N, 0]\!],\, \forall\, x \in \mathbb{R}, \qquad \widetilde{\lambda}_a(y, x) \;\geqslant\; 0 \tag{33}$$

The proof of Proposition 10 leads to

$$\widetilde{\mathcal{L}}_N \;=\; \{\widetilde{\Lambda}_a \,:\, a \in \widetilde{\mathcal{A}}_N\}$$

Thus Assertion (f) of the introduction is a consequence of:

Lemma 17 *We have*

$$\widetilde{\mathcal{A}}_N = \{(1,0,0,...,0)\}$$

Proof

Note that for fixed $y \in [\![-N,0]\!]$, the polynomial $x \mapsto \widetilde{\lambda}_a(y,x)$ has no component on h_0 in the basis $(h_n)_{n\in\mathbb{Z}_+}$. Thus by orthogonality in $\mathbb{L}^2(\gamma)$, we have $\int \widetilde{\lambda}_a(y,x)\,\gamma(dx) = 0$. Since for $a \in \widetilde{\mathcal{A}}_N$ the integrand must be non-negative, we deduce that $\widetilde{\lambda}_a(y,\cdot) \equiv 0$ and it follows that $a_n = 0$ for $n \in [\![N]\!]$, by considering $y = 0$.

■

The case $N = \infty$ is not so clear, since now (31) is an infinite sum and we don't know precisely in which sense it should converge, probably in some weak sense so that (32) has still a meaning (maybe with the $\widetilde{\Lambda}_a(y,\cdot)$ no longer absolutely continuous with respect γ). The set $\widetilde{\mathcal{A}}_\infty$ should be the set of all $a := (a_n)_{n\in\mathbb{Z}_+} \in \mathbb{R}^{\mathbb{Z}_+}$ with $a_0 = 1$ such that $\widetilde{\Lambda}_a$ is a non-negative kernel and we would end up with the conclusion that $\widetilde{\mathcal{L}}_\infty = \{\widetilde{\Lambda}_a : a \in \widetilde{\mathcal{A}}_\infty\}$. The above arguments are no longer sufficient to deduce that $\widetilde{\mathcal{L}}_\infty$ is a singleton. Nevertheless, if we assume that in (31), $\sum_{n\geqslant -y+1} \frac{a_n}{n!}\widetilde{\varphi}_n(y)h_n(x)$ holds in $\mathbb{L}^2(\gamma)$ for x (for any fixed $y \in \mathbb{Z}_-$), then the above proof still implies that $a_n = 0$ for $n \geqslant 1$. Thus any nontrivial intertwining between the reverse Yule and the Ornstein-Uhlenbeck processes has to be quite singular.

References

[1] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93(5):333–348, 1986.

[2] Philippe Biane. Intertwining of Markov semi-groups, some examples. In *Séminaire de Probabilités, XXIX*, volume 1613 of *Lecture Notes in Math.*, pages 30–36. Springer, Berlin, 1995.

[3] Persi Diaconis and James Allen Fill. Strong stationary times via a new form of duality. *Ann. Probab.*, 18(4):1483–1522, 1990.

[4] James Allen Fill. Strong stationary duality for continuous-time Markov chains. I. Theory. *J. Theoret. Probab.*, 5(1):45–70, 1992.

[5] James Allen Fill. The passage time distribution for a birth-and-death chain: strong stationary duality gives a first stochastic proof. *J. Theoret. Probab.*, 22(3):543–557, 2009.

[6] Robert Griffiths. Multivariate Krawtchouk polynomials and composition birth and death processes. *Symmetry*, 8(5):Art. 33, 19, 2016.

[7] Laurent Miclo. Comportement de spectres d'opérateurs de Schrödinger à basse température. *Bull. Sci. Math.*, 119(6):529–553, 1995.

[8] Laurent Miclo. On absorption times and Dirichlet eigenvalues. *ESAIM Probab. Stat.*, 14:117–150, 2010.

[9] Laurent Miclo. Strong stationary times for one-dimensional diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(2):957–996, 2017.

[10] Laurent Miclo. On the construction of set-valued dual processes. *e-prints HAL*, page `hal-01911989`, November 2018.

[11] Laurent Miclo. On the Markovian similarity. In *Séminaire de Probabilités XLIX*, volume 2215 of *Lecture Notes in Math.*, pages 375–403. Springer, Cham, 2018.

[12] Laurent Miclo and Pierre Patie. On a gateway between continuous and discrete Bessel and Laguerre processes. *Annales Henri Lebesgue*, online, 41p., 2019.

[13] Gábor Szegö. *Orthogonal polynomials.* American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.

[14] Michael E. Taylor. *Partial differential equations II. Qualitative studies of linear equations*, volume 116 of *Applied Mathematical Sciences.* Springer, New York, second edition, 2011.

† miclo@math.cnrs.fr

Toulouse School of Economics,
Manufacture des Tabacs, 21, Allée de Brienne
31015 Toulouse cedex 6, France

Institut de Mathématiques de Toulouse
Université Paul Sabatier, 118, route de Narbonne
31062 Toulouse cedex 9, France

‡ pp396@cornell.edu

School of Operations Research and Information Engineering
Cornell University
Ithaca, NY 14853, USA