Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2011, Vol. 47, No. 3, 875–916
DOI: 10.1214/10-AIHP391

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

(Homogeneous) Markovian bridges

Vincent Vigon

IRMA, Université de Strasbourg. E-mail: vigon@math.u-strasbg.fr

Received 31 March 2009; revised 16 September 2010; accepted 28 September 2010

Abstract. (Homogeneous) Markov bridges are (time homogeneous) Markov chains which begin at a given point and end at a given point. The price to pay for preserving the homogeneity is to work with processes with a random life-span.

Bridges are studied both for themselves and for their use in describing the transformations of Markov chains: restriction on a random interval, time reversal, time change, various conditionings comprising the confinement in some part of the state space.

These bridges lead us to look at Markov chains from an unusual point of view: we will work, no longer with only one transition matrix, but with a class of matrices which can be deduced one from the other by Doob transformations. This way of proceeding has the advantage of better describing the "past $\leftrightarrow$ future symmetries": The symmetry of conditional independence (well known) and the symmetry of homogeneity (less well known).

Résumé. Les ponts markoviens (homogènes) sont des chaines de Markov (homogènes) qui démarrent à un point donné et meurent à un point donné. Pour préserver l'homogénéité, une telle chaine de Markov a nécessairement une durée de vie aléatoire.

Nous étudions les ponts pour eux mêmes et pour leur utilité à décrire les transformations d'une chaine de Markov : restriction à un intervalle aléatoire, renversement temporel, changement de temps, conditonnements variés : notament le confinement dans une partie de l'espace d'état.

Ces ponts nous conduisent à considérer les chaines de Markov d'un point de vue inhabituel : nous ne travaillons plus avec une seule matrice de transition comme à l'accoutumée, mais avec une classe de matrices qui se déduisent les unes des autres par transformation de Doob. Cette méthode a l'avantage de mieux décrire les symétries passé $\leftrightarrow$ futur : symétrie de l'indépendance conditionnelle (bien connue) et symétrie de l'homogénéité (moins bien connue).

MSC: 60J10; 60J45; 47A68; 15A23; 60J50

Keywords: Markov chains; Random walks; LU-factorization; Path-decomposition; Fluctuation theory; Probabilistic potential theory; Infinite matrices; Martin boundary

0. Introduction

0.1. *Motivations*

To start with, let us enumerate some transformations of (good) Markov processes:

- The translation of the origin to a stopping time which preserves the law of the process up to a change of the initial distribution (strong Markov property).
- The translation of the origin at the "splitting times" introduced by Jacobsen [18]. This family of times contains stopping times, return times, co-terminal times. As a remarkable example, the global minimizer of a trajectory is a splitting time (see Millar [24]).
- The killing of the trajectory at a splitting time which is the symmetric transformation of the previous one.
- The confinement of trajectories to a part of the state space by conditioning. Such confinement, mixed with some specific translation or killing, preserves the Markovian character of the process (see Jacobsen–Pitman [19]).

- The time reversal (see e.g. Nagasawa [25]) which preserves the Markovian character (and which emphasizes the past–future symmetry).
- The time change (or subordination) which allows us to restrict processes to some interesting random closed set (e.g. the left minima of Lévy processes, see Bertoin [1], Chapter 7). Subordination also allows us to compare Markov processes (see Simon [30]).
- The restriction of the trajectory to the complementary of a regenerative set which leads to the Itô [17] excursion theory.

To describe the law of transformed Markov processes many authors use bridges, the most famous one being the Brownian bridge. For example Markovian bridges with deterministic life-spans are used by Getoor and Sharpe [13] to describe excursions outside a regenerative set (i.e. set of type $\{t\colon\ X_t \in A\}$), while Fitzsimmons [8], for the same purpose, used bridges with random life-spans which have the advantage of being *homogeneous* Markov processes. These same homogeneous bridges were used by Fourati [11] to generalize the Veervat [32] transform from the Brownian case to the more general Lévy processes case. We think that homogeneous bridges are a powerful tool, perhaps not well enough known.

Most of the results appearing in the articles previously cited are profound and rather hard to establish. One of the great difficulties comes from the continuous time setting. Discrete Markov chains, which will occupy us, are really much easier to work with. Meanwhile, the presented results offer a taste of the corresponding results for Markov processes. Homogeneous bridges will be our leading thread: we will study them for themselves and will explain how they are usuful for understanding the transformations of Markov chains. Technically, it is more practical to considere the family of bridges indexed by all the possible beginnings and ends. Such families will be called Markov-bridge-kernels.

0.2. *Detailed summary*

The global scheme of this article is the following: Section 1: Markov-bridge-kernels are defined by axioms. Section 2: We define the class of matrices which will parametrize Markov-bridge-kernels (just as transition matrices parametrize Markov chains). Section 3: We construct Markov-bridge-kernels by Doob transformations. Sections 4–9: We show how bridges interact with classical Markov chains. Sections 10–12: We extend the definition of bridges in different directions. Sections 13–14: We ask some questions about existence and unicity relative to bridges.

We will now write summaries of each section. To keep them as short as possible, we will skip some hypotheses and bibliographic comments. All of these will be specified in the body of the text.

A trajectory $\omega \in \Omega$ is a function from $\mathbb{N}$ to $F = E \cup \{\dagger\}$ such that $\omega_t = \dagger \Rightarrow \omega_{t+1} = \dagger$. Probabilities on Ω are denoted by $\mathbf{E}$ or $\mathbf{E}\{X \in \bullet\}$ (X being the canonical process i.e. the identity on Ω). Expectations are denoted by $\mathbf{E}[\mathfrak{f}]$ or $\mathbf{E}[\mathfrak{f}(X)]$ (brackets allow the distinction). For example, the law under $\mathbf{E}$ of the canonical process killed at a random time T is denoted by $\mathbf{E}\{X_{[0,T]} \in \bullet\} = \mathbf{E}[1_{\{X_{[0,T]} \in \bullet\}}]$.

Section 1: Parallel definition of Markov chains and Markov bridges

From an axiomatic point of view, a Markov-bridge-kernel is a family $\mathbf{E}_{\bullet \rhd \bullet} = (\mathbf{E}_{x \rhd z})_{x,z \in E}$ of probabilities on the set of trajectories such that, almost-surely under $\mathbf{E}_{x \rhd z}$, the canonical process begins at x and ends at z at a random time ζ, and such that:

$$\mathbf{E}_{x \rhd z}\Big[\sum_t \mathfrak{f}(X_{[0,t]}) 1_{\{X_t = y\}} \mathfrak{g}(X_{[t,\zeta]})\Big]$$
$$= \mathbf{E}_{x \rhd y}\big[\mathfrak{f}(X)\big] \mathbf{E}_{x \rhd z}\Big[\sum_t 1_{\{X_t = y\}}\Big] \mathbf{E}_{y \rhd z}\big[\mathfrak{g}(X)\big],$$

where $\mathfrak{f}$, $\mathfrak{g}$ are any test functions on Ω. Remark: in the left and right factors of the second line, the t disappears, which is the sign of the time homogeneity. The above formula, called *past–future extraction* is the formula that makes bridges so practical and beautiful.

To understand this past–future extraction, let us consider $\mathbf{E}_\bullet = (\mathbf{E}_x)_{x\in E}$ a transient *Markov-chain-kernel* which is simply the law of a transient Markov chain started from all possible $x \in E$. The Markov property, summed at all times t gives:

$$\mathbf{E}_x\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Big] = \mathbf{E}_x\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\Big]\mathbf{E}_y\big[\mathfrak{g}(X)\big].$$

The past–future extraction can be seen as a symmetrization of the above formula. Comparing the two previous formulae, we can see that a Markov-bridge-kernel $\mathbf{E}_{\bullet\rhd\bullet}$ is a family of Markov-chain-kernels $(\mathbf{E}_{\bullet\rhd z})_{z\in E}$, but this family is coherent and the two indices x, z play a symmetric role.

Anticipating Section 5, $\mathbf{E}_{x\rhd z}$, can be interpreted as the law of X under $\mathbf{E}_x$, killed the last time it visits z. But we will not adopt this technique to construct Markov-bridge-kernels. Following Doob [5], and Fitzsimmons [8], we will use Doob transformations.

Section 2: *Matrices and their Doob classes*
A Markov chain is described by its starting point x and its transition matrix P. Henceforth, Markov-chain-kernels will now be denoted by $\mathbf{E}^P_\bullet = (\mathbf{E}^P_x)_{x\in E}$. A Markov-bridge-kernel is more naturally described by a *transient D-class of matrices*:

Let P be a non-negative matrix (i.e. a matrix with non-negative entries) and let $U := \sum_n P^n$ be its potential matrix. P is said to be transient if there exists $g > 0$ satisfying $U_{[P]}g < \infty$. The transience is stable by transposition.

Let $h > 0$ be a function. The Doob transform of P is $D_h P(x, y) = \frac{h(y)}{h(x)} P(x, y)$. The D-class of P is $DP = \{D_h P\colon\ h > 0\}$. We see that P is transient if and only if any matrix of DP is transient.

We also give some foundations for discrete potential theory, describing excessive functions ($h \ge 0$, $Ph \le h$), invariant functions ($h \ge 0$, $Ph = h$), and potential functions $h = Ug$, $g \ge 0$. To excessive functions, we also associate a Doob transform $\tilde{D}_h P$ which is simply $D_h P$ restricted to the set where h is positive. It is immediate that $\tilde{D}_h P$ is sub-stochastic.

From now on, our basic data will be a non-negative matrix P, or equivalently a D-class DP with a chosen representative P. Troughout the article, except in Section 12, P will be assumed transient. Each time we have to write $\mathbf{E}^P_x$, we also require that P be sub-stochastic, but the bridge itself is defined for any transient P. This allows us to replace P by $P^\top$ at any moment (doing this, we play with the time reversal).

Section 3: *Construction of Markov-bridge-kernels*
The function $U(\bullet, z) = \sum_n P^n(\bullet, z)$ is P-excessive so the matrix $P_{\rhd z} := \tilde{D}_{U(\bullet,z)}P$ is sub-stochastic. A Markov chain driven by $P_{\rhd z}$ dies at z with probability one. Moreover, for any $h > 0$, we have $P_{\rhd z} = (D_h P)_{\rhd z}$. So we can denote $\mathbf{E}^{P_{\rhd z}}_x$ by $\mathbf{E}^{DP}_{x\rhd z}$. We establish the "past–future extraction" under $\mathbf{E}^P_x$ i.e.

$$\mathbf{E}^P_x\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Big] = \mathbf{E}^{DP}_{x\rhd y}\big[\mathfrak{f}(X)\big]\mathbf{E}^P_x\Big[\sum_t \mathfrak{f}(X_{[0,t]})\Big]\mathbf{E}^P_y\big[\mathfrak{g}(X)\big].$$

We will use often this formula to describe transformations of Markov chains. Replacing $\mathbf{E}^P_x$ by $\mathbf{E}^{DP}_{x\rhd z}$ in the past–future extraction shows that the constructed family $\mathbf{E}^{DP}_{\bullet\rhd\bullet}$ is a Markov-bridge-kernel as defined in the axiomatic Section 1.

Section 4: *Time reversal*
We show that the law of the reversed path under the bridge i.e. $\mathbf{E}^{DP}_{x\rhd z}\{(X_{\overleftarrow{[0,\zeta]}}) \in \bullet\}$ is equal to $\mathbf{E}^{DP^\top}_{z\rhd x}\{X \in \bullet\}$. Note that we have defined bridges for any transient matrix P (not necessarily sub-stochastic) so $\mathbf{E}^{DP^\top}_{z\rhd x}$ is well defined.

Section 5: *Initial and final conditionings*
Providing $\mathbf{E}^P_x\{\exists t\colon\ X_t = z\} > 0$ we show that $\mathbf{E}^{DP}_{x\rhd z}\{X \in \bullet\} = \mathbf{E}^P_x\{X_{\tau_z} \in \bullet / \tau_z \in [0,\infty[\}$ where τ_z is the last visit at z.

As a consequence, we show that any Markov chain that begins at x and ends at z has law $\mathbf{E}^{DP}_{x\rhd z}$ for a certain DP.

Section 6: *Confinement*
Let $\mathbb{A}$ be a part of $E \times E$ representing some chosen transitions. Let $\{XX \subset \mathbb{A}\}$ be the event "all transitions of X belong to $\mathbb{A}$." For example if $E = \mathbb{N}$ and $\mathbb{A} = \{(x, y)\colon\ x < y\}$ then $\{XX \subset \mathbb{A}\}$ is "X increases." Let $S_{\mathbb{A}}$ be the first time where X makes a transition outside $\mathbb{A}$. We explain the link between the conditioning by $\{XX \subset \mathbb{A}\}$, the killing at $S_{\mathbb{A}}$ and making a final conditioning. For example we show that

$$\mathbf{E}_x^P\big\{(X_{[0,S_{\mathbb{A}}]})_{[0,\tau_z]} \in \bullet/\tau_z \in [0,\infty[\big\} = \mathbf{E}_{x\rhd z}^{DP}\big\{X \in \bullet/\{XX \subset \mathbb{A}\}\big\} = \mathbf{E}_{x\rhd z}^{DP_{\mathbb{A}}}\{X \in \bullet\},$$

where $P_{\mathbb{A}}(a, b) = P(a, b)1_{\{(a,b)\in\mathbb{A}\}}$.

Section 7: *Splitting trajectory*
Following Jacobsen [18], we define splitting times as random times S such that, on $\{t \le \zeta\}$, $\{S = t\} = \{X_{[0,t]} \in \mathcal{C}_S\} \cap \{X_{[t,\zeta]} \in \mathcal{D}_S\}$ where $\mathcal{C}_S, \mathcal{D}_S$ are some subsets of Ω. Such splitting times generalize stopping times (the ones with $\mathcal{D}_S = \Omega$) and return times (the ones with $\mathcal{C}_S = \Omega$). An immediate consequence of the past–future extraction is that, under $\mathbf{E}_{x\rhd z}^P\{X \in \bullet/X_S = y\}$, the pieces of trajectory $X_{[0,S]}, X_{[S,\zeta]}$ are independent and have laws $\mathbf{E}_{x\rhd y}^{DP}\{X \in \bullet/\mathcal{D}_S\}$ and $\mathbf{E}_{y\rhd z}^{DP}\{X \in \bullet/\mathcal{C}_S\}$.

Then we define "markers" which are families of times that generalize splitting times. These markers will be used further on to describe excursions outside random sets.

Section 8: *Time-change and excursion*
Let T be any stopping time taking values in $[1, \zeta] \cup \{+\infty\}$. We define the time-change $\daleth$ by iterating this stopping time: $\daleth_0 = 0$, $\daleth_1 = T, \ldots, \daleth_{n+1} = \daleth_n + T(X_{[\daleth_n,\zeta]})$. It is easy to verify that $n \mapsto X_{\daleth_n}$ is also a Markov chain under $\mathbf{E}_x^P$. Let us write $\daleth_f$ the latest finite $\daleth_n$; using the past–future extraction we obtain:

$$\begin{aligned}\mathbf{E}_{x\rhd z}^P\{X_{\daleth_f} = y\}U(x, z) &= \mathbf{E}_x^P\Big[\sum_n 1_{\{X_{\daleth_n}=y\}}\Big]\mathbf{E}_y^P\Big[\sum_{t<T} 1_{\{X_t=z\}}\Big] \\ &=: V_T(x, y)W_T(y, z).\end{aligned}\tag{1}$$

Summing over all y we get the factorization $U = V_T W_T$ which we will meet again further on.

The set $\Lambda = \{t\colon\ \exists n\colon\ \daleth_n = t\}$, called past-spotted set, has some interesting properties and can be seen as a generalization of what is usually called a regenerative set. An excursion outside this set is a piece of trajectory of type $X_{[g,d]}$ where $]g, d[$ is a connected part of $[0, \zeta] \setminus \Lambda$. Each excursion which we can spot by events: before g, after d, inside $[g, d]$, or by intersections of three such events, can be described using the marker technique. We show that more complicated excursions cannot be described so easily.

Section 9: *Fluctuation theory*
This theory is concerned with extrema of trajectories. The first minimizer ρ is a splitting time, and is also the end of a past-spotted set. The associated factorization $U = V_T W_T$ is then the classical LU-factorization, also called Wiener–Hopf factorization in the framework of random-walks. Bridges allow us to write an equivalent of the Veervat [32] transformation, namely:

$$\mathbf{E}_{x\rhd x}^{DP}\big[\mathfrak{f}(X_{[\rho,\zeta]} \wr X_{[0,\rho]})/X_\rho = y\big] = \mathbf{E}_{y\rhd y}^{DP}\big[\mathfrak{f}(X)/\exists t\colon\ X_t = x, \forall t\ X_t \ge y\big],$$

where $\wr$ stands for the "biting" concatenation. Such a generalization of the Veervat transform has been made by Fourati [11] in the Lévy processes setting.

Section 10: *Extension of the universe*
Until now, we have worked on the canonical space which implies that all random elements are functions of X. In this section we add an independent U[0, 1] random variable which allows us to define randomized versions of splitting-time and time-change. For example the factorization (1) is still true when T is a randomized stopping time and $\daleth_0, \daleth_1, \daleth_2, \ldots$ are its randomized iterations.

We see also that an inequality $P \ge Q$ can be interpreted by the fact that the Q-Markov chain is equal to the P-Markov chain killed at a randomized stopping time. While $U_{[P]} \ge QU_{[P]}$ can be interpreted by the fact that the Q-Markov chain is equal to the P-Markov chain time-changed by a randomized time-change.

Section 11: *Extension to the Martin boundary*
In this section we add the hypothesis that the graph of P is locally finite and irreducible. It is a common assumption when constructing the Martin boundary. Roughly speaking, the minimal Martin boundary $\mathcal{M}_{\min}$ is the set of all possible limits ξ for a Markov chain X driven by any one of sub-stochastic matrices $Q \in DP$. This Martin boundary can also be identified with all the P-invariant minimal functions. It is very natural to extend the bridge $(\mathbf{E}^{DP}_{x \rhd z})_{x,z \in E}$ to $(\mathbf{E}^{DP}_{x \rhd \eta})_{x \in E, \eta \in E \cup \mathcal{M}_{\min}}$.

We explain the link between confinement (as in Section 6) and final conditioning, but now "final conditioning" also included the conditioning by: "X reaches the Martin boundary at $\xi \in \mathcal{M}_{\min}$" (in an infinite length of time).

Section 12: *Un-normalized bridges*
When P is sub-stochastic and transient, an expression for the bridge can be:

$$\mathbf{E}^{DP}_{x \rhd y} = \frac{\mathbf{E}^P_x[\sum_t 1_{\{X_{[0,t]} \in \bullet\}} 1_{\{X_t = y\}}]}{U_{[P]}(x,y)},$$

the transience hypothesis ensuring $U_{[P]}(x,y) < \infty$. Then we define the un-normalized bridge $\mathbf{F}^P_{x \rhd z}$ as the numerator of the above ratio. This notion extends to any P with finite spectral radius. We see that the past–future extraction and the time-reversal formula are still valid. We compare advantages of normalized and un-normalized bridges.

Section 13: *When two bridges partially coincide*
A set $K \subset E$ is said to be P-closed when, in the directed graph of P, every trajectory which begins and ends in K, lies entirely in K. E.g. P-absorbing sets are P-closed. Let K be a P-closed and P'-closed set. Writing $P_K(x,y) := P(x,y)1_{\{x,y \in K\}}$ and $DP_K := D(P_K)$ we have:

$$DP_K = DP'_K \quad \Rightarrow \quad \forall x, z \in K \quad \mathbf{E}^{DP}_{x \rhd z} = \mathbf{E}^{DP'}_{x \rhd z}.$$

Assuming the existence of some "directed-spanning-trees" in graphs of P, P', we show that the converse of this implication is also true. The existence of such a tree is always true if you suppose P, P' irreducible.

Applying this, we give a necessary and sufficient condition to have

$$\mathbf{E}^P_x\{X_{[0,t]} \in \bullet / X_t = z\} = \mathbf{E}^{P'}_x\{X_{[0,t]} \in \bullet / X_t = z\}.$$

Section 14: *All axiomatic bridges can be constructed*
For any Markov-bridge-kernel $\mathbf{E}_{\bullet \rhd \bullet}$ (in the sense of the axiomatic definition given in Section 1), we show that there exists a D-class DP such that $\mathbf{E}_{\bullet \rhd \bullet} = \mathbf{E}^{DP}_{\bullet \rhd \bullet}$.

Remark. Why we do not assume irreducibility? *This assumption, which is almost necessary for recurrent theory, is not really appropriate for our subject because some killings, conditionings and time-changings do not keep the irreducible character. Moreover, Markov chains are sometimes called "random dynamic system." To merit this name, we have to include cases where trajectories starting from different points do not necessarily intersect each other.*

1. Parallel definition of Markov chains and Markov bridges

1.1. *Definition of a Markov chain*

Let F be a denumerable set. Elements of $F^{\mathbb{N}}$ are called trajectories. The set $F^{\mathbb{N}}$ is endowed with the cylindrical σ-field. We can restrict a trajectory: $\omega_{[0,t]} = (\omega_0\omega_1 \cdots \omega_t)$, translate it: $\omega_{[t,\infty[} = (\omega_t\omega_{t+1}\omega_{t+2}\cdots)$, mix the two operations: $\omega_{[t,t+s]} = (\omega_{[t,\infty[})_{[0,s]} = (\omega_t\omega_{t+1}\cdots\omega_{t+s})$, and even reverse the time: $\omega_{\underset{[0,t]}{\leftarrow}} = \omega_t\omega_{t-1}\cdots\omega_0$.

An inhomogeneous Markov chain is a random element $X : (\Omega, \mathcal{F}, \mathbf{E}) \mapsto F^{\mathbb{N}}$ such that, for all $t \in \mathbb{N}$, $X_{[0,t]}$ and $X_{[t,\infty[}$ are independent conditionally to X_t. A Markov chain is an inhomogeneous Markov chain which satisfies moreover $\forall x \in F$, $\forall s, t \in \mathbb{N}$, $\mathbf{E}\{X_{[s,\infty[} \in \bullet / X_s = x\} = \mathbf{E}\{X_{[t,\infty[} \in \bullet / X_t = x\}$, whenever the two conditionings are well defined.

Our purpose is to give general considerations about Markov chains (and not to talk about a particular one, constructed in a particular setting), so it is more convenient to work in the canonical framework: from now on, we consider $X : F^{\mathbb{N}} \mapsto F^{\mathbb{N}}$ the identity application. A probability $\mathbf{E}$ on $F^{\mathbb{N}}$ is a *Markov-chain-law* when, under $\mathbf{E}$, X is a Markov chain. Actually, it is more practical to consider a family of Markov-chain-laws rather that one:

Definition 1.1. *A* Markov-chain-kernel $\mathbf{E}_\bullet = (\mathbf{E}_x)_{x\in F}$ *is a family of probabilities on* $F^{\mathbb{N}}$ *satisfying the following axioms*:

1. "Support": *We have* $\mathbf{E}_x\{X_0 = x\} = 1$ *for all* $x \in F$.
2. "Markov property": *We have*

$$\mathbf{E}_x\big[\mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\infty[})\big] = \mathbf{E}_x\big[\mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\big]\mathbf{E}_y\big[\mathfrak{g}(X)\big]$$

for all states $x, y \in F$, *times* $t \in \mathbb{N}$, *test functions* $\mathfrak{f} : F^{[0,t]} \mapsto \mathbb{R}_+$ *and* $\mathfrak{g} : F^{\mathbb{N}} \mapsto \mathbb{R}_+$.

Clearly, when $\mathbf{E}_\bullet$ is a Markov-chain-kernel and when α is a probability on F then $\mathbf{E}_\alpha := \sum_a \alpha(a)\mathbf{E}_a$ is a Markov-chain-law.

Reciprocally, when $\mathbf{E}$ is a Markov-chain-law, then, for all $x \in F$, we can define: $\mathbf{E}_x := \mathbf{E}\{X_{[T_x,\infty[} \in \bullet / T_x < \infty\}$ where $T_x = \inf\{t \geq 0\colon\ X_t = x\}$, if the conditioning is well defined (and anything if not). The kernel $(\mathbf{E}_a)_{a\in F}$ that we went to construct is a Markov-chain-kernel and we have $\mathbf{E} = \mathbf{E}_\alpha$ with $\alpha = \mathbf{E}\{X_0 \in \bullet\}$.

1.2. *Transient case*

Now we distinguish a state $\dagger \in F$, called *cemetery point*, and put $E = F \setminus \{\dagger\}$. We denote by Ω the event $\{\forall t : X_t = \dagger \Rightarrow X_{t+1} = \dagger\}$ (i.e. the event "$\dagger$ is absorbing"). On Ω we define the random time $\zeta = \sup\{t\colon\ X_t \in E\}$. Thus the event $\{\zeta < \infty\} \subset \Omega$ represents all the trajectories that meet $\dagger$. From now on, a restricted trajectory $\omega_{[0,t]}$ is identified with $\omega_0\omega_1 \cdots \omega_t \dagger\dagger \cdots$ so that $\bigcup_{s\in\mathbb{N}} E^{[0,s]} \cup E^{\mathbb{N}}$ is identified with Ω. As a consequence $X_{[t,\infty[} = X_{[t,\zeta]}$.

Definition 1.2. *A* transient Markov-chain-kernel $\mathbf{E}_\bullet = (\mathbf{E}_x)_{x\in E}$ *is a family of probabilities on* $F^{\mathbb{N}}$ *satisfying the following axioms*:

1. "Support": *We have* $\mathbf{E}_x\{X_0 = x \cap \Omega\} = 1$ *for all* $x \in E$.
2. "Transience": *We have* $\mathbf{E}_x[\sum_t 1_{\{X_t\in K\}}] < \infty$ *for all* $x \in E$ *and finite* $K \subset E$.
3. "Future extraction": *We have*

$$\mathbf{E}_x\left[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\right] = \mathbf{E}_x\left[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\right]\mathbf{E}_y\big[\mathfrak{g}(X)\big]$$

for all states $x, y \in E$, *and test functions* $\mathfrak{f}, \mathfrak{g} : \Omega \mapsto \mathbb{R}^+$.

Remark 1.3.

- *In the above formula, the summation* $\sum_t$ *can be replaced by* $\sum_{t\leq\zeta}$ *thanks to the term* $1_{\{X_t=y\}}$ *with* $y \in E$.
- *The family* $\mathbf{E}_\bullet = (\mathbf{E}_x)_{x\in E}$ *can be extended to* $\mathbf{E}_\bullet = (\mathbf{E}_x)_{x\in F}$ *by setting* $\mathbf{E}_\dagger$ *to be the Dirac measure on the trajectory* $\dagger\dagger\dagger\cdots$.
- *Let us consider any process* $t \mapsto Z_t$ *which is "past-adapted" i.e.* $\forall t$ Z_t *is* $\sigma(X_{[0,t]})$*-measurable. Using the fact that* $X_{[0,t]} = X_{[0,t]} \circ X_{[0,t]}$ *we see that* $Z_t = Z_t(X_{[0,t]})$. *Using the fact that, on* $\{t \leq \zeta\}$, $t = \zeta \circ X_{[0,t]}$, *we get*: $Z_t 1_{\{t\leq\zeta\}} = (Z_\zeta)(X_{[0,t]})1_{\{t\leq\zeta\}}$. *So processes of type* $t \mapsto \mathfrak{f}(X_{[0,t]})$ *represent the most general class of past-adapted process* (*whenever we work on the life interval* $[0, \zeta]$).
- *The future extraction is a reformulation of the Markov property* (*second axiom of Definition* 1.1): *you go from the Markov property to the future extraction by summing over all* $t \in \mathbb{N}$. *You go from the future extraction to the Markov property replacing* $\mathfrak{f}(X)$ *by* $\mathfrak{f}(X)1_{\{\zeta=t\}}$. *So clearly a transient Markov-chain-kernel* $(\mathbf{E}_x)_{x\in E}$ *is a Markov-chain-kernel*.

Definition 1.4. *A* mortal Markov-chain-kernel $\mathbf{E}_\bullet = (\mathbf{E}_x)_{x\in E}$ *is a family of probabilities on* $F^{\mathbb{N}}$ *which satisfies* $\forall x \in E$ $\mathbf{E}_x\{X_0 = x, \zeta < \infty\} = 1$ *and the axiom "future extraction" of Definition* 1.2.

It is a classical exercise to check that a mortal Markov-chain-kernel also verifies the second axiom (transience) of Definition 1.2. So the mortality is a particular case of the transience.

1.3. *Definition of a bridge*

Definition 1.5. *A* Markov-bridge-kernel $\mathbf{E}_{\bullet\rhd\bullet} = (\mathbf{E}_{x\rhd z})_{x,z\in E}$ *is a family of probabilities on* $F^{\mathbb{N}}$ *satisfying the following axioms for all* $x, y, z \in E$:

1. "Degeneracy": *If* $\mathbf{E}_{x\rhd z}\{\zeta = 0\} = 1$ *then* $\mathbf{E}_{x\rhd z}$ *is the Dirac measure on the trajectory* $x\dagger\dagger\cdots$. *In this case we say that* $\mathbf{E}_{x\rhd z}$ *is degenerated.*
2. "Support": *If* $\mathbf{E}_{x\rhd z}$ *is non-degenerated then* $\mathbf{E}_{x\rhd z}\{X_0 = x, X_\zeta = z, \zeta < \infty\} = 1$.
3. "Cohesion": *If* $\mathbf{E}_{x\rhd y}, \mathbf{E}_{y\rhd z}$ *are non-degenerated then* $\mathbf{E}_{x\rhd z}\{\exists t\colon X_t = y\} > 0$.
4. "Past-future extraction": *We have*

$$\mathbf{E}_{x\rhd z}\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Big] = \mathbf{E}_{x\rhd y}\big[\mathfrak{f}(X)\big]\mathbf{E}_{x\rhd z}\Big[\sum_t 1_{\{X_t=y\}}\Big]\mathbf{E}_{y\rhd z}\big[\mathfrak{g}(X)\big]$$

for all test functions $\mathfrak{f}, \mathfrak{g} : \Omega \mapsto \mathbb{R}^+$.

Remark 1.6. *The past–future extraction can be seen as a symmetrization of the future extraction. To see the symmetry, we can define* $(\mathbf{E}^{\top}_{x\rhd z})_{x,z\in E}$ *by* $\mathbf{E}^{\top}_{x\rhd z} = \mathbf{E}_{z\rhd x}\{(X_{\overleftarrow{[0,\zeta]}}) \in \bullet\}$ *when* $\mathbf{E}_{z\rhd x}$ *is non-degenerated, and* $\mathbf{E}^{\top}_{x\rhd z} = \delta_{x\dagger\dagger\cdots}$ *when* $\mathbf{E}_{z\rhd x}$ *is degenerated. The reader is invited to check that this new family* $\mathbf{E}^{\top}_{\bullet\rhd\bullet}$ *is a Markov-bridge-kernel.*

Remark 1.7. *The axiom "support" is partially included in the axiom "past–future extraction": If you weaken "support" by: "If* $\mathbf{E}_{x\rhd z}$ *is not degenerated then* $\mathbf{E}_{x\rhd z}\{\exists t\colon X_t = x\} > 0, \mathbf{E}_{x\rhd z}\{\exists t\colon X_t = z\} > 0$*" then, applying "past–future extraction" with* $\mathfrak{f} := 1_{\{X_\zeta = z\}}$, $y := z$, $\mathfrak{g} := 1_\Omega$ *gives you* $\mathbf{E}_{x\rhd z}\{X_\zeta = z\} = 1$ *while applying "past–future extraction" with* $\mathfrak{f} := 1_\Omega$, $y := x$, $\mathfrak{g} := 1_{\{X_0=x\}}$ *gives you* $\mathbf{E}_{x\rhd z}\{X_0 = x\} = 1$.

Proposition 1.8 (Past and future extractions). *Let* $\mathbf{E}_{\bullet\rhd\bullet}$ *be a Markov-bridge-kernel. For every states* $x, y \in E$ *and test functions* $\mathfrak{f}, \mathfrak{g} : \Omega \mapsto \mathbb{R}^+$ *we have*:

- *The* future extraction:

$$\mathbf{E}_{x\rhd z}\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Big] = \mathbf{E}_{x\rhd z}\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\Big]\mathbf{E}_{y\rhd z}\big[\mathfrak{g}(X)\big].$$

- *The* past extraction:

$$\mathbf{E}_{x\rhd z}\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Big] = \mathbf{E}_{x\rhd y}\big[\mathfrak{f}(X)\big]\mathbf{E}_{x\rhd z}\Big[\sum_t 1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Big].$$

Proof. Applying the past–future extraction with the pair $(\mathfrak{f}, 1_\Omega)$ gives:

$$\mathbf{E}_{x\rhd z}\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\Big] = \mathbf{E}_{x\rhd y}\big[\mathfrak{f}(X)\big]\mathbf{E}_{x\rhd z}\Big[\sum_t 1_{\{X_t=y\}}\Big].$$

Substituing this equality in the past–future extraction with a pair $(\mathfrak{f}, \mathfrak{g})$ gives the future extraction. Past extraction is proven similarly. □

Corollary 1.9. *Let* $\mathbf{E}_{\bullet\rhd\bullet}$ *be a Markov-bridge-kernel. For every* $z \in E$, $\mathbf{E}_{\bullet\rhd z}$ *is a mortal Markov-chain-kernel.*

Thus, a Markov-bridge-kernel may be considerated as a coherent family of Markov-chain-kernels.

1.4. *From Markov-chain-kernel to matrices*

To a Markov-chain-kernel $\mathbf{E}_\bullet$, we can associate its *transition matrix* $\forall x, y \in F\ \overline{P}(x, y) := \mathbf{E}_x\{X_1 = y\}$ which is stochastic i.e. $\sum_{a\in F} \overline{P}(x, a) = 1$. Reciprocally, to each stochastic matrix corresponds a unique Markov-chain-kernel as explained below:

Let $\overline{P}$ be any stochastic matrix on F. For each $x \in F$, there exists a unique probability $\mathbf{E}_x^{\overline{P}}$ satisfying

$$\mathbf{E}_x^{\overline{P}}\{X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n\} = \delta_x(x_0)\overline{P}(x_0, x_1)\cdots\overline{P}(x_{n-1}, x_n)$$

for all times $t \in \mathbb{N}$ and families of states $x_0, \ldots, x_n \in F$. (The unicity comes from the monotone class theorem, the existence from the Kolmogorov theorem.) The family $\mathbf{E}_\bullet^{\overline{P}}$ constructed by this way is a Markov-chain-kernel whose transition matrix is $\overline{P}$.

A Markov-chain-kernel $\mathbf{E}_\bullet^{\overline{P}}$ supported by Ω is also characterized by the restriction of its transition matrix to E. Indeed, once we get $P := \overline{P}_{|E}$, we can reconstruct $\overline{P}$ on F by $\overline{P}(x, \dagger) = 1 - \sum_{y\in E} P(x, y)$ and $\overline{P}(\dagger, \dagger) = 1$. In this case, we will prefer to write $\mathbf{E}_\bullet^{P}$ instead of $\mathbf{E}_\bullet^{\overline{P}}$. Remark also that the restricted matrix P is sub-stochastic i.e. $\forall x\ \sum_y P(x, y) \le 1$.

From now on, it is clear that the data of a sub-stochastic matrix P on E or a Markov-chain-kernel $\mathbf{E}_\bullet^{P}$ supported by Ω are equivalent. We prefer to use the matrix as our first data.

2. Matrices and their Doob classes

The main purpose of this section is to introduce and study the "transient D-classes of matrices" which will parametrize bridges.

Data 2.1. *Throughout this article, P will represent a non-negative matrix indexed by E (non-negative means $\forall x, y \in E\ P(x, y) \ge 0$). Assumptions of transience (see below), sub-stochasticity and irreducibility will be added further on.*

2.1. *Notations*

By default, functions f, g, h, measures α, β, μ, ν and matrices P, Q, are defined on E or $E \times E$ and take values in $\mathbb{R}_+$ (never in $\mathbb{R}_+ \cup \{+\infty\}$). We write briefly $h > 0$ to indicate that the function h is everywhere positive. The difference between a function and a measure is purely of a lexical order. The destiny of functions is to be multiplied on the left of matrices, and so they can be drawn as column vectors. The destiny of measures is to be multiplied on the right of matrices, and so they can be drawn as row vectors. E.g.

$$\mu P(y) = \sum_a \mu(a)P(a, y), \qquad Ph(x) = \sum_b P(x, b)h(x).$$

In this paper, a directed graph $(B, \to)$ is the data of a denumerable set B (often $B = E$) and a relation $\to$ between elements of B. A vertex b is an element of B, an directed edge $(a \to c)$ is a related pair of vertices. For example, the directed graph P is $(E, \overset{P}{\to})$ with $a \overset{P}{\to} c \Leftrightarrow P(a, c) > 0$.

We write $a \rightsquigarrow c$ when there exists an directed chain $a \to b_1 \to \cdots \to b_n \to c$. We accept degenerated chains, so we always have $a \rightsquigarrow a$. For example, when P is sub-stochastic: $a \overset{P}{\rightsquigarrow} c \Leftrightarrow \mathbf{E}_a^P\{\exists t\colon X_t = c\} > 0$.

For A, C sets of vertices, we write $\{\bullet \rightsquigarrow C\} = \{b\colon \exists c \in C\colon b \rightsquigarrow c\}$, $\{A \rightsquigarrow \bullet\} = \{b\colon \exists a \in A\colon a \rightsquigarrow b\}$, and $\{A \rightsquigarrow \bullet \rightsquigarrow C\} = \{b\colon \exists a \in A, \exists c \in C\colon a \rightsquigarrow b \rightsquigarrow c\}$.

Because P will come back very often, we will sometimes drop the mention of P in sentences like: "Let h be a P-excessive function.…" We also leave out P in many notations like "$U_{[P]}$" (the potential matrix relative to P) or "$x \overset{P}{\rightsquigarrow} y$."

2.2. *Transient and mortal matrices*

We denote by:

$$U_{[P]}(x,y)=\sum_t P^t(x,y)\in[0,\infty[,\quad \text{where } P^0=I \text{ is identity.}$$

$U_{[P]}$ is called the potential matrix associated to P.

Definition 2.2 (Transience). *A* weighting function *for P if a function $g>0$ such that $\forall x\in E$: $U_{[P]}g(x)<\infty$. We say that P is* transient *if such a weighting function exists for P. Transient matrices are matrices with spectral radius a bit less than one, in a sense we will explain in Section 12.1.*

We insist on the fact that, for us, a transient matrix is not necessarily sub-stochastic; however in this case, we have a nice characterization of transient matrices:

Proposition 2.3. *Suppose that P is sub-stochastic. The following points are equivalent*:

1. *P is transient.*
2. *For all $x,y\in E$, $U_{[P]}(x,y)<\infty$.*
3. *The Markov-chain-kernel $\mathbf{E}^P_\bullet$ is transient.*

Proof. $1\Rightarrow 2$ is obvious.
$2\Leftrightarrow 3$ comes from the fact that we can interpret $U_{[P]}(x,y)$ as $\mathbf{E}^P_x[\sum_t 1_{\{X_t=y\}}]$.
$3\Rightarrow 1$. Applying the Markov property at time $T_a=\inf\{t\colon\ X_t=a\}$ gives:

$$U(x,a)=\mathbf{E}^P_x\left[\sum_{t\geq T_a}1_{\{X_t=a\}}\right]=\mathbf{E}^P_x\{T_a<\infty\}U(a,a)$$

so that $\frac{U(\bullet,a)}{U(a,a)}$ is bounded by 1. Pick a probability $\alpha>0$ and set $g(a)=\frac{\alpha(a)}{U(a,a)}>0$. Therefore we obtain $Ug\leq 1$, so g is a weighting function for P. □

Proposition 2.4. *P is transient if and only if $P^\top$ is transient.*

Proof. Let g be a weighting function for P. We have $Ug\geq g>0$. We take a probability $\alpha>0$. We define the measure $\mu(x)=\frac{\alpha(x)}{Ug(x)}$. Then we have $\mu Ug=1$, so $\forall x\ \mu U(x)$ is finite and $\mu^\top$ can be taken as a weighting function for $P^\top$. □

We appreciate the transience because of its stability under transposition (which is not the case for the sub-stochasticity). In particular, working with such matrices avoids having to make different statements for excessive functions and excessive measures.

Definition 2.5 (Mortality). *We say that P is mortal when it is sub-stochastic and when moreover $\forall x\in E$: $\lim_n P^n 1_E(x)=0$.*

Because $P^n 1_E(x)=\mathbf{E}^P_x\{\zeta\geq n\}$, the mortality of P is equivalent to the mortality of $\mathbf{E}^P_\bullet$. So the mortal matrices are sub-stochastic and transient. We can compare mortal matrices with the stochastic ones, which could also be qualified of immortal.

2.3. *Excessive, invariant, potential functions*

A function $h\geq 0$ is said to be: P-excessive if $Ph\leq h$, P-invariant if $Ph=h$, P-potential if there exists a function $g\geq 0$, called the charge, such that $h=U_{[P]}g$. A (non-negative) measure μ is said to be P-excessive, P-invariant, P-potential if its transposition $\mu^\top$ is $P^\top$-excessive, $P^\top$-invariant, $P^\top$-potential.

An excessive function h is a potential function if and only if $P^n h$ tends to 0 point-wise. The Riesz decomposition says that every excessive function can be written as the sum of an invariant function and a potential one. For more details about this subject see e.g. Dellacherie–Meyer [4].

2.4. *Doob transformations*

For any $h > 0$ we note $D_h P(x, y) = \frac{h(y)}{h(x)} P(x, y)$. The application $P \mapsto D_h P$ is invertible and we easily check that $U_{[D_h P]} = D_h U_{[P]}$.

When h is a P-excessive function (which might vanish), we write:

$$\tilde{D}_h P(x, y) = 1_{\{h(x)>0\}} \frac{h(y)}{h(x)} P(x, y)$$
$$= \frac{h(y)P(x, y)}{h(x)} \quad \text{with the convention } \frac{0}{0} = 0.$$

The equality between the two lines comes from the fact that $Ph \le h$ imposes $h(x) = 0 \Rightarrow h(y)P(x, y) = 0$.

If you prefer to add the assumption that P is irreducible, then all excessive functions are everywhere positive, and you can replace $\tilde{D}$ by D.

Proposition 2.6. *Let h be a function everywhere positive. We have*:

- *$D_h P$ is sub-stochastic if and only if h is P-excessive.*
- *$D_h P$ is sub-stochastic and mortal if and only if h is P-potential.*
- *$D_h P$ is stochastic if and only if h is P-invariant.*

The proof is easy and left to the reader.

2.5. *Doob classes* (D-*classes*)

Let P', P'' be two non-negative matrices. We say that they are D-equivalent if there exists a function $h > 0$ such that $P'' = D_h P'$. The equivalence class of P is denoted by DP. If P is transient, then, all matrices of DP are transient (because $U_{[D_h P]} = D_h U_{[P]}$).

As we will see in the proof of the following proposition, properties relative to a transient matrix P can be often proved using a sub-stochastic matrix P' in DP.

Proposition 2.7. *P is transient if and only if there exists a function $h > 0$ which is P-potential.*

Proof. Suppose P transient. Let $g > 0$ be a weighting function for P. Then the function $h = U_{[P]} g$ is potential and everywhere positive.

Conversely, suppose the existence of a function $h = U_{[P]} f > 0$. We set $P' = D_h P$ which is sub-stochastic. The transience of P is equivalent to the transience of P'. We also remark that $1_E = U_{[P']} \frac{f}{h}$.

We pick $q \in \,]0, 1[$. The quantity $g(x) = U_{[P']}(U_{[qP']} \frac{f}{h})(x) = U_{[qP']} 1_E(x) \le \frac{1}{1-q}$ is everywhere finite. In order to use g as a weighting function for P, it remains to prove the positivity of g.

Let $T = \inf\{t\colon X_t \in \{\frac{f}{h} > 0\}\} = \inf\{t\colon X_t \in \{f > 0\}\}$. The positivity of $U_{[P']} \frac{f}{h}(x)$ implies (and is equivalent to) $\mathbf{E}_x^{P'}\{T < \infty\} > 0$. So, denoting by μ_x the sub-probability on $\mathbb{N}$ defined by $\mu_x(t) = \mathbf{E}_x^{P'}\{T = t\}$, this sub-probability is non-identically zero. On another hand, the positivity of $U_{[qP']} \frac{f}{h}(x)$ is equivalent to the positivity of $\mathbf{E}_x^{P'}[q^T 1_{\{T<\infty\}}] = \sum_t q^t \mu_x(t)$ which is true because the function $t \mapsto q^t$ is everywhere positive on $\mathbb{N}$. □

Corollary 2.8. *P is transient if and only if, among matrices of DP, there exists a least one matrix which is sub-stochastic and mortal.*

Proof. Suppose P transient. Let g be a weighting function for P. Put $h = Ug > 0$ and $Q = D_h P \in DP$. Because h is a P-potential function, Q is sub-stochastic and mortal. Reciprocally, suppose that there exists Q in DP which is mortal and sub-stochastic. But because Q is mortal, from Proposition 2.6, h is P-potential. From Proposition 2.7, P is transient. $\square$

Remark 2.9. *At this point comes a natural question: when P is transient, does there exist a matrix Q in DP which is stochastic, or equivalently: does there exist a P-invariant function $h > 0$? When P is irreducible the necessary a sufficient condition is that E is infinite: it is a consequence of the Martin Boundary theory and this will be explained in Section 11. For a reducible matrix, the natural condition would be the existence, for each state x, of an infinite path starting from x and leaving every finite subset of E. But we do not find a reference for this....*

2.6. *Absorbing sets*

A set $K \subset E$ is said absorbing when $K = \{K \overset{P}{\rightsquigarrow} \bullet\}$. Every set of the form $\{A \overset{P}{\rightsquigarrow} \bullet\}$ are P-absorbing.

A short analysis shows us that K is $P^\top$-absorbing if and only if 1_K is a P-excessive function. Moreover, the support of any P-excessive function is $P^\top$-absorbing. For a P-potential function $h = Ug$ we have $\{h > 0\} = \{\bullet \overset{P}{\rightsquigarrow} \{g > 0\}\}$.

Looking at the definition of the transformation $\tilde{D}$, we see that the graph of $\tilde{D}_h P$ is the restriction of the graph of P to $\{h > 0\}$. In particular the graph of $\tilde{D}_{U(\bullet,z)} P$ is the restriction of the graph of P to $\{\bullet \overset{P}{\rightsquigarrow} z\}$. In the next section, we will see how such a matrix $\tilde{D}_{U(\bullet,z)} P$ allows us to construct the bridge ending at z.

3. Construction of Markov-bridge-kernels

The use of Doob transformations to produce Markovian trajectories ending at one specified point, comes back to Doob [5]. The idea of considering homogeneous bridges as a whole family for describing some pieces of trajectory is due to Fitzsimmons [8]. The notion of bridges we develop here is the normalized and discrete version of the one developed by Fitzsimmons.

Data 3.1. *From now on, P is a transient matrix.*

3.1. *A killing*

Lemma 3.2. *Let Q be a sub-stochastic matrix. We have $\mathbf{E}_x^Q\{X_\zeta = y\} = U_{[Q]}(x,y)Q(y,\dagger)$ with $Q(y,\dagger) = 1 - \sum_a Q(y,a)$.*

Proof. $\mathbf{E}_x^Q\{X_\zeta = y\} = \mathbf{E}_x^Q \sum_t 1_{\{X_t = y\}} 1_{\{X_{t+1} = \dagger\}} = U_{[Q]}(x,y)Q(y,\dagger)$. $\square$

Proposition 3.3. *Let $U_{[P]}g = Ug$ be a P-potential function. The Markov-chain-kernel $\mathbf{E}_\bullet^{\tilde{D}_{Ug}P}$ is mortal and satisfies:*

$$\forall x, y \in E \quad \mathbf{E}_x^{\tilde{D}_{Ug}P}\{X_\zeta = y\} = \frac{U_{[P]}(x,y)g(y)}{U_{[P]}g(y)} \quad \textit{with } \frac{0}{0} = 0.$$

Proof. The mortality of $\tilde{D}_{Ug}P$ ($\Leftrightarrow$ the mortality of $\mathbf{E}_\bullet^{\tilde{D}_{Ug}P}$) comes from the fact that Ug is a potential function. Let us compute the law of the position at ζ. From the previous lemma, we need to compute $\tilde{D}_{Ug}P(x,\dagger)$. We make the next computation, using the convention $\frac{0}{0} = 0$:

$$\tilde{D}_{Ug}P(x,\dagger) = 1 - \sum_a \frac{Ug(a)}{Ug(x)} P(x,a) = 1 - \frac{PUg(x)}{Ug(x)} = 1 - \frac{Ug(x) - g(x)}{Ug(x)} = \frac{g(x)}{Ug(x)}.$$

From the previous lemma: $\mathbf{E}_x^{\tilde{D}_{Ug}P}\{X_\zeta = y\} = U_{[\tilde{D}_{Ug}P]}(x,y)\frac{g(y)}{Ug(y)} = \frac{U(x,y)g(y)}{Ug(y)}$. $\square$

3.2. *A candidate for the bridge*

For any transient matrix Q we write $Q_{\triangleright z}$ for $\tilde{D}_{U_{[Q]}(\bullet,z)}Q$.

Lemma 3.4 (Fundamental). *Fix $z \in E$. For any function $h > 0$ we have $(D_h P)_{\triangleright z} = P_{\triangleright z}$. For any P-excessive function $h \geq 0$ such that $z \in \{h > 0\}$ we have $(\tilde{D}_h P)_{\triangleright z} = P_{\triangleright z}$.*

Proof. We write $P' = D_h P$ and $U' = U_{[P']}$. Clearly $\{U(\bullet, z) > 0\} = \{U'(\bullet, z) > 0\}$. When a, b belong to this set:

$$\tilde{P}'_{\triangleright z}(a,b) = \frac{U'(b,z)}{U'(a,z)} P'(a,b) = \frac{(h(z)/h(b))U(b,z)}{(h(z)/h(a))U(a,z)} \frac{h(b)}{h(a)} P(a,b) = P_{\triangleright z}(a,b), \tag{2}$$

when a or b does not belong to this set then clearly $P'_{\triangleright z}(a,b) = P_{\triangleright z}(a,b) = 0$.

Suppose now that $P' = \tilde{D}_h P$ with h excessive and $z \in \{h > 0\}$. Because $\{h > 0\}$ is $P^\top$-absorbing we have $\{h > 0\} \supset \{\bullet \overset{P}{\rightsquigarrow} z\}$. In particular $\{U(\bullet, z) > 0\} = \{U'(\bullet, z) > 0\}$. The rest of the proof is similar. $\square$

If $x \not\rightsquigarrow z$ then $P_{\triangleright z}(x,z) = 0$ so that $\mathbf{E}_x^{P_{\triangleright z}}$ is the Dirac measure on the trajectory $x\dagger\dagger\cdots$. If $x \rightsquigarrow z$ then Proposition 3.3 indicates that $\mathbf{E}_x^{P_{\triangleright z}}\{X_\zeta = y\} = I(y,z)$. So the family $(\mathbf{E}_x^{P_{\triangleright z}})_{x,z\in E}$ is the good candidate to be a Markov-bridge-kernel.

On the other hand, the first part of the fundamental Lemma 3.4 says that: if P' is a matrix of DP, then $P'_{\triangleright z} = P_{\triangleright z}$. This allows us to state:

$$\mathbf{E}_{x\triangleright z}^{DP} := \mathbf{E}_x^{P_{\triangleright z}} = \mathbf{E}_x^{P'_{\triangleright z}}.$$

This notation emphasizes that $\mathbf{E}_{x\triangleright z}^{DP}$ is a function of the class DP. This notation also emphasizes a symmetry between the initial point x and the final point z.

3.3. *Past–future extraction*

The following theorem gives formulae which explain the appearance of $\mathbf{E}_{x\triangleright z}^{DP}$ appear in many contexts (see Section 7 for concrete applications). Moreover, these formulae will say that $\mathbf{E}_{\bullet\triangleright\bullet}^{DP}$ is a Markov-bridge-kernel as defined in the axiomatic Definition 1.5.

Theorem 3.5. *We fix $\mathfrak{f}, \mathfrak{g} : \Omega \mapsto \mathbb{R}_+$ and $x, y, z \in E$.*

1. *Suppose that P is sub-stochastic and transient. $\mathbf{E}_x^P$ satisfies the* past–future extraction *property i.e.*:

$$\begin{aligned}\mathbf{E}_x^P\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Big] &= \mathbf{E}_{x\triangleright y}^{DP}\big[\mathfrak{f}(X)\big]\mathbf{E}_x^P\Big[\sum_t 1_{\{X_t=y\}}\Big]\mathbf{E}_y^P\big[\mathfrak{g}(X)\big] \\ &= \mathbf{E}_{x\triangleright y}^{DP}\big[\mathfrak{f}(X)\big]U(x,y)\mathbf{E}_y^P\big[\mathfrak{g}(X)\big].\end{aligned}$$

2. *Suppose that P is just transient. $\mathbf{E}_{x\triangleright z}^{DP}$ satisfies the* past–future extraction *property i.e.*:

$$\begin{aligned}&\mathbf{E}_{x\triangleright z}^{DP}\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Big] \\ &\quad= \mathbf{E}_{x\triangleright y}^{DP}\big[\mathfrak{f}(X)\big]\mathbf{E}_{x\triangleright z}^{DP}\Big[\sum_t 1_{\{X_t=y\}}\Big]\mathbf{E}_{y\triangleright z}^{DP}\big[\mathfrak{g}(X)\big] \\ &\quad= \mathbf{E}_{x\triangleright y}^{DP}\big[\mathfrak{f}(X)\big]\frac{U(x,y)U(y,z)}{U(x,z)}\mathbf{E}_{y\triangleright z}^{DP}\big[\mathfrak{g}(X)\big]\end{aligned}$$

(*with the convention* $\frac{0}{0} = 0$).

Proof. 1. If $x \not\rightsquigarrow y$ then the equation becomes $0 = 0$. We assume now that $x \rightsquigarrow y$. Applying the Markov property at each t (or the "future extraction"):

$$\mathbf{E}_x^P\left[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\right] = \mathbf{E}_x^P\left[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\right]\mathbf{E}_y^P\big[\mathfrak{g}(X)\big].$$

So, to prove the past–future extraction it is sufficient to establish:

$$\mathbf{E}_x^P\left[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\right] = \mathbf{E}_{x\rhd y}^{DP}\big[\mathfrak{f}(X)\big]U(x,y). \tag{3}$$

By monotone class theorem, it is sufficient to establish this equality for all functions of type $\mathfrak{f} = 1_{\{X_0=x_0,\dots,X_n=x_n\}}$. Replacing $\mathfrak{f}$ in the left-hand side of (3) gives:

$$\begin{aligned}
&\mathbf{E}_x^P\left[\sum_t 1_{\{X_0=x_0,\dots,X_n=x_n\}}1_{\{X_t=y\}}\right]\\
&\quad= \sum_{t\ge n} I(x,x_0)P(x_0,x_1)\cdots P(x_{n-1},x_n)P^{t-n}(x_n,y)\\
&\quad= I(x,x_0)P(x_0,x_1)\cdots P(x_{n-1},x_n)U(x_n,y).
\end{aligned}$$

Replacing $\mathfrak{f}$ in the right-hand side of (3) gives:

$$\mathbf{E}_{x\rhd y}^{DP}\big[\mathfrak{f}(X)\big]U(x,y) = I(x,x_0)P_{\rhd y}(x_0,x_1)\cdots P_{\rhd y}(x_{n-1},x_n)U(x,y).$$

To identify, we just have to recall that $P_{\rhd y} = \tilde{D}_{U(\bullet,y)}P$.

2. We can apply the first part of this theorem, substituting P by $P_{\rhd z}$ (which is sub-stochastic):

$$\mathbf{E}_x^{P_{\rhd z}}\left[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\right] = \mathbf{E}_{x\rhd y}^{DP}\big[\mathfrak{f}(X)\big]U_{[P_{\rhd z}]}(x,y)\mathbf{E}_y^{P_{\rhd z}}\big[\mathfrak{g}(X)\big].$$

But we have $U_{[P_{\rhd z}]}(x,y) = \frac{U(x,y)U(y,z)}{U(x,z)}$, with $\frac{0}{0} = 0$. $\square$

Corollary 3.6. *The family $\mathbf{E}_\bullet^{DP} = (\mathbf{E}_{x\rhd z}^{DP})_{x,z\in E}$ is a Markov-bridge-kernel in the sense of Definition 1.5.*

Remark 3.7. *Let $h \ge 0$ be a P-excessive function. Using the second part of the fundamental Lemma 3.4, we see that the past extraction can be applied to $\tilde{D}_h P$ as follows*

$$\forall x \in \{h>0\}\quad \mathbf{E}_x^{\tilde{D}_hP}\left[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\right] = \mathbf{E}_{x\rhd y}^{DP}\big[\mathfrak{f}(X)\big]\tilde{D}_hU(x,y)$$

while the previous theorem gives us only $\mathbf{E}_{x\rhd y}^{D(\tilde{D}_hP)}[\mathfrak{f}(X)]\tilde{D}_hU(x,y)$.

4. Time reversal

Data 4.1. *From now on, P is a transient and sub-stochastic matrix.*

Here is a very old question: At what kind of times can a Markov chain be reversed to produce a Markov chain? More precisely: for which $L : \Omega \mapsto \mathbb{N} \cup \{\bowtie\}$ ($\bowtie$ is a refuge) is $\mathbf{E}_\alpha^P\{X_{\overleftarrow{[0,L]}} \in \bullet / L \ne \bowtie\}$ still a Markov chain law? Let us call such times L "reversers."

Kolmogorov [21] explained that a Markov chain started with an invariant distribution can be reversed at a deterministic time t to produce a t-steps Markov chain (i.e. a chain killed at t). But remark that a Markov chain killed at a deterministic time is inhomogeneous, so deterministic times are not reversers.

Hunt [15] shows that last entrance times are reversers. Nagasawa [25], in a continuous-time setting, shows that "return times" (see Section 7.1) are reversers. Proposition 4.2 below is a discrete time version of Nagasawa's theorem.

But the more synthetic answer is: Reversers are random times L such that $X_{[0,L]}$ is a Markov chain: The action of reversing the time is completely inoffensive for Markov chains. The real danger is the killing! So a better term for reversers is: "death times," and the most general class of such times has been described by Jacobsen (see Section 7.2). It is worth noting that there is death times (and so reversers) that are not return times; so the terminology "return times" is a bit of an overstatement.

In this section, we consider mortal Markov chains, so we are not really preoccupied by the killing. We see how the transition matrix is changed by time reversal, and we see how this becomes simple when the mortal Markov chain is a bridge.

4.1. *To reverse a mortal Markov chain*

Proposition 4.2. *Let α be a probability on E. Denote by $\widehat{P} = \tilde{D}_{\alpha U} P^\top$ and by $\beta = \mathbf{E}^P_\alpha\{X_\zeta = y\}$. We have*:

$$\mathbf{E}^P_\alpha\{X_{\underset{[0,\zeta]}{\leftarrow}} \in \bullet\} = \mathbf{E}^{\widehat{P}}_\beta\{X \in \bullet\}.$$

Proof. It is enough to test this equality with cylindrical functions $\mathfrak{f} = 1_{\{X_0=y_0,\dots,X_n=y_n\}}$:

$$\begin{aligned}
&\mathbf{E}^P_\alpha\big[\mathfrak{f}(X_{\underset{[0,\zeta]}{\leftarrow}})\big] \\
&\quad= \sum_{t\ge n} \mathbf{E}^P_\alpha\{\zeta = t, X_t = y_0, X_{t-1} = y_1, \dots, X_{t-n+1} = y_{n-1}, X_{t-n} = y_n\} \\
&\quad= \sum_{t\ge n} \mathbf{E}^P_\alpha\{X_{t-n} = y_n, X_{t-n+1} = y_{n-1}, \dots, X_{t-1} = y_1, X_t = y_0, X_{t+1=\dagger}\} \\
&\quad= \sum_{t\ge n} \alpha P^{t-n}(y_n) P(y_n, y_{n-1}) \cdots P(y_1, y_0) P(y_0, \dagger) \\
&\quad= P(y_0, \dagger) P^\top(y_0, y_1) \cdots P^\top(y_{n-1}, y_n) \alpha U(y_n) \\
&\quad= \alpha U(y_0) P(y_0, \dagger) \tilde{D}_{\alpha U} P^\top(y_0, y_1) \cdots \tilde{D}_{\alpha U} P^\top(y_{n-1}, y_n) \\
&\quad= \beta(y_0) \widehat{P}(y_0, y_1) \cdots \widehat{P}(y_{n-1}, y_n),
\end{aligned}$$

to make β appear in the last line, we used Lemma 3.2. □

4.2. *To reverse a bridge*

Theorem 4.3. *Suppose $x \overset{P}{\rightsquigarrow} z$. We have*:

$$\mathbf{E}^{DP}_{x\rhd z}\{X_{\underset{[0,\zeta]}{\leftarrow}} \in \bullet\} = \mathbf{E}^{DP^\top}_{z\rhd x}\{X \in \bullet\}.$$

Proof. Let us consider the probability $\mathbf{E}^{DP}_{x\rhd z}\{X_{\underset{[0,\zeta]}{\leftarrow}} \in \bullet\} = \mathbf{E}^{P_{\rhd z}}_x\{X_{\underset{[0,\zeta]}{\leftarrow}} \in \bullet\}$. According to the previous lemma, it is a Markov chain whose initial law is δ_z and whose transition matrix is:

$$\widehat{P} = \tilde{D}_{\delta_x U_{[P_{\rhd z}]}} (P_{\rhd z})^\top = (P^\top)_{\rhd x}.$$

So that $\mathbf{E}^{DP}_{x\rhd z}\{X_{\underset{[0,\zeta]}{\leftarrow}} \in \bullet\} = \mathbf{E}^{DP^\top}_{z\rhd x}\{X \in \bullet\}$. □

5. Initial and final conditionings

We give a recipe to produce a bridge by mixing a killing and a conditionning.

5.1. *Notations*

We denote by $T_y = \inf\{t\colon\ X_t = y\}$. We call initial conditioning the transformation from $\mathbf{E}\{X \in \bullet\}$ into $\mathbf{E}\{X_{[T_y,\zeta]} \in \bullet / T_y < \infty\}$ which is written symbolically by an arrow as follows:

$$\mathbf{E}\{X \in \bullet\} \xrightarrow{y\rhd} \mathbf{E}\{X_{[T_y,\zeta]} \in \bullet / T_y < \infty\}.$$

If we apply this transformation to $\mathbf{E}_x^P$ (assuming $x \rightsquigarrow y$) then the strong Markov property applied to T_y gives immediately $\mathbf{E}_x^P \xrightarrow{y\rhd} \mathbf{E}_y^P$.

Symmetrically, we denote by $\tau_y = \sup\{t\colon\ X_t = y\}$. We call final conditioning the transformation from $\mathbf{E}\{X \in \bullet\}$ into $\mathbf{E}\{X_{[0,\tau_y]} \in \bullet / \tau_y \in [0,\infty[\}$ which is written symbolically:

$$\mathbf{E}\{X \in \bullet\} \xrightarrow{\rhd y} \mathbf{E}\big\{X_{[0,\tau_y]} \in \bullet / \tau_y \in [0,\infty[\big\}.$$

5.2. *Final conditioning produces a bridge*

Theorem 5.1.

1. *Suppose $x \rightsquigarrow y$. The final conditioning transforms a Markov chain into a bridge*:

$$\mathbf{E}_x^P \xrightarrow{\rhd y} \mathbf{E}_{x\rhd y}^{DP}.$$

2. *Suppose $x \rightsquigarrow y \rightsquigarrow z$. The final conditioning transforms a bridge into another bridge*:

$$\mathbf{E}_{x\rhd z}^{DP} \xrightarrow{\rhd y} \mathbf{E}_{x\rhd y}^{DP}.$$

3. *Suppose $x \rightsquigarrow y \rightsquigarrow z$. The initial conditioning transforms a bridge into another bridge*:

$$\mathbf{E}_{x\rhd z}^{DP} \xrightarrow{y\rhd} \mathbf{E}_{y\rhd z}^{DP}.$$

Proof. 1. We have $\{\tau_y \in [0,\infty[\} = \{T_y < \infty\} = \{\exists t\colon\ X_t = y\}$ and in virtue of our hypothesis $x \rightsquigarrow y$, this event as non-zero probability under $\mathbf{E}_x^P$. Applying the past–future extraction (Theorem 3.5, item 1) gives:

$$\begin{aligned}\mathbf{E}_x^P\big[\mathfrak{f}(X_{[0,\tau_y]})1_{\{\exists t X_t = y\}}\big] &= \mathbf{E}_x^P\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}1_{\{\forall s\geq 1, X_s\neq y\}} \circ X_{[t,\zeta]}\Big]\\ &= \mathbf{E}_{x\rhd y}^{DP}\big[\mathfrak{f}(X)\big]U(x,y)\mathbf{E}_y^P\{\forall s \geq 1, X_s \neq y\}.\end{aligned}$$

Denote $R(\mathfrak{f})$ the quantity above. Computing $R(\mathfrak{f})/R(1_\Omega)$ gives the first point of the theorem. The second comes by doing the same calculus, but starting from $\mathbf{E}_{x\rhd z}^P$ instead of $\mathbf{E}_x^P$. The third point comes directly by applying the strong Markov property at the time T_y under $\mathbf{E}_{x\rhd z}^{DP}$. □

Corollary 5.2. *The initial and final conditioning commute*:

$$\mathbf{E}_x^P \xrightarrow{a\rhd} \mathbf{E}_a^P \xrightarrow{\rhd b} \mathbf{E}_{a\rhd b}^{DP},$$
$$\mathbf{E}_x^P \xrightarrow{\rhd b} \mathbf{E}_{x\rhd b}^{DP} \xrightarrow{a\rhd} \mathbf{E}_{a\rhd b}^{DP}.$$

5.3. *Application*: *Characterizations of bridges*

The following proposition indicates that a bridge is a just a Markov chain that dies at an unique point.

Proposition 5.3. *Suppose $x \rightsquigarrow z$. The following points are equivalent*:

1. $\mathbf{E}_x^P = \mathbf{E}_{x \rhd z}^{DP}$.
2. $\mathbf{E}_x^P\{X_\zeta = z\} = 1$.

Proof. $1 \Rightarrow 2$. Obvious.
$2 \Rightarrow 1$. From Theorem 5.1 we know that final conditioning produces bridge i.e:

$$\mathbf{E}_{x \rhd z}^{DP}\{X \in \bullet\} = \mathbf{E}_x^P\left\{X_{[0,\tau_z]} \in \bullet / \tau_z \in [0,\infty[\right\}$$

but point 2 implies that $\tau_z = \zeta$ a.s. and $\mathbf{E}_x^P\{\tau_z \in [0,\infty[\} = 1$, so the right-hand term of the above formula is equal to $\mathbf{E}_x^P\{X \in \bullet\}$. □

Proposition 5.4. *Let z be a fixed state. The following points are equivalent*:

1. $\forall x\ \mathbf{E}_x^P = \mathbf{E}_{x \rhd z}^{DP}$.
2. *The function $x \mapsto U_{[P]}(x,z)$ is constant and $P(z,\dagger) = 1 - \sum_a P(z,a) > 0$.*
3. *The function $x \mapsto \mathbf{E}_x^P\{X_\zeta = z\}$ is constant and non-zero.*
4. *The function $x \mapsto \mathbf{E}_x^P\{T_z < \infty\}$ is constant and non-zero.*

When it is the case, *the two last functions are equal to* 1.

Proof. If there exists a such that $a \not\rightsquigarrow z$, then, all points 1, 2, 3 and 4 are false. So we can suppose without loss of generality that $E = \{a : a \rightsquigarrow z\}$, so that the function $a \mapsto U(a,z)$ is everywhere positive.

$1 \Rightarrow 2$. We see that 1 is also equivalent to $P = P_{\rhd z}$. We have:

$$\forall x \quad U_{[P]}(x,z) = U_{[P_{\rhd z}]}(x,z) = D_{U(\bullet,z)}U(x,z) = U(z,z).$$

$2 \Leftrightarrow 3$. Comes directly from Lemma 3.2.
$2 \Leftrightarrow 4$. Comes from the formula $\mathbf{E}_x^P\{T_z < \infty\} = \frac{U(x,z)}{U(z,z)}$.
$3 \Rightarrow 1$. Suppose 3. We have:

$$\mathbf{E}_x^P\{X_\zeta = z\} = \mathbf{E}_x^P\{T_z < \infty, X_\zeta \circ X_{[T_z,\zeta]} = z\} = \mathbf{E}_x^P\{T_z < \infty\}\mathbf{E}_z^P\{T_z < \infty\}$$

so the function $x \mapsto \mathbf{E}_x^P\{T_z < \infty\}$ is necessarily equal to 1. This forces the canonical process to go into z from any state. Because we have excluded the recurrent case (P is transient), we deduce that the canonical process necessarily dies at z with probability one. Thus $\forall x\ \mathbf{E}_x^P = \mathbf{E}_{x \rhd z}^P$. □

6. Confinement

6.1. *Restriction of possible transitions*

Let $\mathbb{A}$ be a subset of $E \times E$ which represents some "selected transitions." Pair of states will be often written xy instead of (x,y). We define the following event:

$$\{XX \subset \mathbb{A}\} := \{\forall t < \zeta\colon\ X_t X_{t+1} \in \mathbb{A}\}.$$

For example, if $K \subset E$ and $\mathbb{A} = K \times K$ then $\{XX \subset \mathbb{A}\}$ means "X stays all its life inside K" and this event is also denoted by $X \subset K$. As a second example, if $E = \mathbb{Z}$ and $\mathbb{A} = \{xy\colon\ x < y\}$ then $\{XX \subset \mathbb{A}\}$ is the set trajectories that increase (until ζ).

We denote by: $P_{\mathbb{A}}(x,y)=P(x,y)1_{xy\in\mathbb{A}}$ and by $S_{\mathbb{A}}=\inf\{t\colon X_tX_{t+1}\notin\mathbb{A}\}$, so $X_{[0,S_{\mathbb{A}}]}$ belongs to $\{XX\subset\mathbb{A}\}$. If $\mathbf{E}$ is any probability supported by Ω, we define the following two transformations:

$$\mathbf{E}\{X\in\bullet\}\xrightarrow{X_{[0,S_{\mathbb{A}}]}}\mathbf{E}\{X_{[0,S_{\mathbb{A}}]}\in\bullet\},$$
$$\mathbf{E}\{X\in\bullet\}\xrightarrow{/XX\subset\mathbb{A}}\mathbf{E}\{X\in\bullet/XX\subset\mathbb{A}\}$$

for the second transformation, we require $\mathbf{E}\{XX\subset\mathbb{A}\}>0$. These two transformations are substantially different (see Section 6.3) but, according to the next theorem, these two transformations give the same probability once we compound them with the final conditioning. We recall that the symbol $\xrightarrow{\rhd z}$ stands for the "final conditioning at z," which we can also call "bridgification" (see Theorem 5.1).

6.2. *Confinement and final conditioning*

Theorem 6.1.

1. *Suppose* $x\overset{P_{\mathbb{A}}}{\rightsquigarrow}z$, *we have*:

$$\mathbf{E}^P_x\xrightarrow{X_{[0,S_{\mathbb{A}}]}}\xrightarrow{\rhd z}\mathbf{E}^{DP_{\mathbb{A}}}_{x\rhd z}.$$

2. *Suppose* $x\overset{P_{\mathbb{A}}}{\rightsquigarrow}z$ *and* $\mathbf{E}^P_x\{XX\subset\mathbb{A}\}>0$, *we have*:

$$\mathbf{E}^P_x\xrightarrow{/XX\subset\mathbb{A}}\xrightarrow{\rhd z}\mathbf{E}^{DP_{\mathbb{A}}}_{x\rhd z}.$$

3. *Suppose* $x\overset{P_{\mathbb{A}}}{\rightsquigarrow}z$ *and* $\mathbf{E}^{DP}_{x\rhd z}\{XX\subset\mathbb{A}\}>0$, *we have*:

$$\mathbf{E}^P_x\xrightarrow{\rhd z}\xrightarrow{/XX\subset\mathbb{A}}\mathbf{E}^{DP_{\mathbb{A}}}_{x\rhd z}.$$

Proof. In all this proof x and z are fixed and we suppose $x\overset{P_{\mathbb{A}}}{\rightsquigarrow}z$. The monotone class theorem allows us to work with test functions of type $\mathfrak{f}=1_{\{X_0=x_0,\dots,X_n=x_n\}}$.

Proof of 1. Let us characterize $\mathbf{E}^P_x\{X_{[0,S_{\mathbb{A}}]}\in\bullet\}$.

$$\begin{aligned}\mathbf{E}^P_x\big[\mathfrak{f}(X_{[0,S_{\mathbb{A}}]})\big]&=\mathbf{E}^P_x[1_{\{X_0=x_0,\dots,X_n=x_n\}}1_{\{n\le S_{\mathbb{A}}\}}]\\&=\mathbf{E}^P_x[1_{\{X_0=x_0,\dots,X_n=x_n\}}1_{\{x_0x_1\in\mathbb{A},\dots,x_{n-1}x_n\in\mathbb{A}\}}]\\&=I(x,x_0)P(x_0,x_1)\cdots P(x_{n-1},x_n)1_{\{x_0x_1\in\mathbb{A},\dots,x_{n-1}x_n\in\mathbb{A}\}}\\&=I(x,x_0)P_{\mathbb{A}}(x_0,x_1)\cdots P_{\mathbb{A}}(x_{n-1},x_n).\end{aligned}$$

So the probability $\mathbf{E}^P_x\{X_{[0,S_{\mathbb{A}}]}\in\bullet\}$ is a Markov chain with initial distribution δ_x and matrix transition $P_{\mathbb{A}}$. If you transform it by the final conditioning at z, you get $\mathbf{E}^{DP_{\mathbb{A}}}_{x\rhd z}$.

Proof of 2. Let us study $\mathbf{E}^P_x\{X\in\bullet/XX\subset\mathbb{A}\}$.

$$\begin{aligned}&\mathbf{E}^P_x\big[\mathfrak{f}(X)1_{\{XX\subset\mathbb{A}\}}\big]\\&\quad=\mathbf{E}^P_x[1_{\{X_0=x_0,\dots,X_n=x_n\}}1_{\{x_0x_1\in\mathbb{A},\dots,x_{n-1}x_n\in\mathbb{A}\}}1_{\{XX\subset\mathbb{A}\}}\circ X_{[n,\zeta]}]\\&\quad=I(x,x_0)P(x_0,x_1)\cdots P(x_{n-1},x_n)1_{\{x_0x_1\in\mathbb{A},\dots,x_{n-1}x_n\in\mathbb{A}\}}\mathbf{E}^P_{x_n}\{XX\subset\mathbb{A}\}\\&\quad=I(x,x_0)P_{\mathbb{A}}(x_0,x_1)\cdots P_{\mathbb{A}}(x_{n-1},x_n)\mathbf{E}^P_{x_n}\{XX\subset\mathbb{A}\}.\end{aligned}$$

We set $h(a) := \mathbf{E}_a^P\{XX \subset \mathbb{A}\}$ which is a $P_{\mathbb{A}}$-excessive function. We have:

$$\mathbf{E}_x^P\big[\mathfrak{f}(X)/\{XX \subset \mathbb{A}\}\big] = I(x, x_0)P_{\mathbb{A}}(x_0, x_1)\cdots P_{\mathbb{A}}(x_{n-1}, x_n)\frac{h(x_n)}{h(x)}$$
$$= I(x, x_0)\tilde{D}_h P_{\mathbb{A}}(x_0, x_1)\cdots \tilde{D}_h P_{\mathbb{A}}(x_{n-1}, x_n).$$

So $\mathbf{E}_x^P\{X \in \bullet/XX \subset \mathbb{A}\}$ is a Markov chain law with transition matrix $\tilde{D}_h P_{\mathbb{A}}$. The fundamental Lemma 3.4 indicates that $(\tilde{D}_h P_{\mathbb{A}})_{\rhd z} = (P_{\mathbb{A}})_{\rhd z}$ and as conclusion:

$$\mathbf{E}_x^P \xrightarrow{/XX\subset\mathbb{A}} \mathbf{E}_x^{\tilde{D}_h P_{\mathbb{A}}} \xrightarrow{\rhd z} \mathbf{E}_{x\rhd z}^{DP_{\mathbb{A}}}.$$

Proof of 3. Let us study $\mathbf{E}_{x\rhd z}^{DP}\{X \in \bullet/XX \subset \mathbb{A}\}$. Because $\mathbf{E}_{x\rhd z}^{DP}$ is simply a Markov chain with transition matrix $P_{\rhd z}$, we can apply the same analyze as previously to show that $\mathbf{E}_{x\rhd z}^{DP}\{X \in \bullet/XX \subset \mathbb{A}\}$ is a Markov chain whose transition matrix is $\tilde{D}_{h'}(P_{\rhd z})_{\mathbb{A}}$ where $h'(a) = \mathbf{E}_a^{P_{\rhd z}}\{xX \subset \mathbb{A}\}$. The computation shows also that:

$$\tilde{D}_{h'}(P_{\rhd z})_{\mathbb{A}} = \tilde{D}_{h''}P_{\mathbb{A}} \quad \text{with } h''(a) = h'(a)U_{[P_{\mathbb{A}}]}(a, z).$$

At this point we have shown:

$$\mathbf{E}_x \xrightarrow{\rhd z} \mathbf{E}_x^{P_{\rhd z}} \xrightarrow{/XX\subset\mathbb{A}} \mathbf{E}_x^{\tilde{D}_{h''}P_{\mathbb{A}}}.$$

The crucial point is to realize that, under $\mathbf{E}_{x\rhd z}^{DP}\{X \in \bullet/XX \subset \mathbb{A}\}$, the canonical process dies at z with probability one. Thus, this Markov chain is equal to its bridge (Proposition 5.3), so the right-hand term of the above formula can also be denoted by $\mathbf{E}_{x\rhd z}^{DP_{\mathbb{A}}}$. □

6.3. *Comments on simulations*

To appreciate the difference between the operations of killing and conditioning we can have a look at simulation methods. To simulate a process following the law $\mathbf{E}_x^P\{X_{[0,S_{\mathbb{A}}]} \in \bullet\}$, we just have to pick one trajectory under $\mathbf{E}_x^P$ and to kill this trajectory at time $S_{\mathbb{A}}$. To simulate a process following the law $\mathbf{E}_x^P\{X \in \bullet/XX \subset \mathbb{A}\}$, we can use the rejection method: we simulate an i.i.d. sequence of trajectories following the law $\mathbf{E}_x^P$, and then we take the first trajectory which lies entirely in the set $\{XX \subset \mathbb{A}\}$. This second procedure can be rather long, especially if the set $\{XX \subset \mathbb{A}\}$ is "slim." Let us give some other means to simulate $\mathbf{E}_x^P\{X \in \bullet/XX \subset \mathbb{A}\}$:

Firstly, if we have a method to compute the function $h(x) = \mathbf{E}_x^P\{XX \subset \mathbb{A}\}$, then we just have to simulate a Markov chain with transitions $D_h P_{\mathbb{A}}$. But the computation of this function h seems as difficult as the simulation itself.

Secondly, by an indirect method, we can imagine that we are able to simulate a random variable with the law $z \mapsto \mathbf{E}_x^P\{X_\zeta = z/XX \subset \mathbb{A}\}$. Suppose that Z is such a random variable, then it is now enough to simulate a trajectory following the law $\mathbf{E}_{x\rhd Z}^{DP_{\mathbb{A}}}\{X \in \bullet\}$, which can be done by computing the function $U_{P_{\mathbb{A}}}(\bullet, Z)$ (it is a classical problem e.g. this can be done by the relaxation method). Perhaps, for the reader, it is not so clear that $\mathbf{E}_{x\rhd Z}^{DP_{\mathbb{A}}}\{X \in \bullet\} = \mathbf{E}_x^P\{X \in \bullet/XX \subset \mathbb{A}\}$, so let us explain this fact in detail. Of course, we suppose that, under $\mathbf{E}_x^P\{X \in \bullet/XX \subset \mathbb{A}\}$, the canonical process dies (it is not possible to simulate infinite trajectories). Then we have:

$$\mathbf{E}_x^P\{X \in \bullet/XX \subset \mathbb{A}\} = \sum_z \mathbf{E}_x^P\{X \in \bullet/X_\zeta = z/XX \subset \mathbb{A}\}\mathbf{E}_x^P\{X_\zeta = z/XX \subset \mathbb{A}\}. \tag{4}$$

But, for every z such that $\mathbf{E}_x^P\{X_\zeta = z/XX \subset \mathbb{A}\} > 0$, the double conditioning appearing above can also be seen as the result of the two following transformations:

$$\mathbf{E}_x^P\{X \in \bullet\} \xrightarrow{/XX\subset\mathbb{A}}\xrightarrow{\rhd z} \mathbf{E}_x^P\{X \in \bullet/X_\zeta = z/XX \subset \mathbb{A}\}.$$

According to Theorem 6.1, the right-hand term is equal to $\mathbf{E}^{DP_{\mathbb{A}}}_{x\rhd z}$ and thus the formula (4) can be rewritten:

$$\mathbf{E}^P_x\{X\in\bullet/XX\subset\mathbb{A}\} = \sum_z \mathbf{E}^{DP_{\mathbb{A}}}_{x\rhd z}\{X\in\bullet\}\mathbf{E}^P_x\{X_\zeta=z/XX\subset\mathbb{A}\}$$
$$\stackrel{\text{law}}{=} \mathbf{E}^{DP_{\mathbb{A}}}_{x\rhd Z}\{X\in\bullet\}.$$

7. Splitting trajectory

The definition of splitting times and their property of splitting the Markovian path into conditionally independent components (Theorem 7.3) are due to Jacobsen [18]. He worked both on a the continuous time setting [18] and on a discrete time setting in collaboration with Pitman [19]. Jacobsen's intentions were to generalize: (1) the famous Williams decomposition of the Brownian path (see Section 9), (2) the splitting at coterminal times (see Section 8).

7.1. *Splitting, stopping, return times*

Definition 7.1. *Let S be an application from Ω to $\mathbb{N}\cup\{\bowtie\}$, where $\bowtie$ is a refuge which can be $+\infty$, $-\infty$ or otherwise. We say that S is a splitting time if there exist $\mathcal{C}_S$ and $\mathcal{D}_S$, two measurable subsets of Ω such that for all t:*

$$\text{on } \{t\le\zeta\}\quad \{S=t\}=\{X_{[0,t]}\in\mathcal{C}_S\}\cap\{X_{[t,\zeta]}\in\mathcal{D}_S\}$$

which can also be written:

$$\text{on } \{t\le\zeta\}\quad 1_{\{S=t\}}=(1_{\mathcal{C}_S}\circ X_{[0,t]})(1_{\mathcal{D}_S}\circ X_{[t,\zeta]}).$$

Sets $\mathcal{C}_S$ and $\mathcal{D}_S$ are called first and second parameters of the splitting time S. Perhaps it will help your intuition to consider these sets as two dictionaries of trajectories; to say "$\{S(\omega)=t\}$," you have to check that the past $X_{[0,t]}(\omega)$ belongs to $\mathcal{C}_S$, and that the future $X_{[t,\zeta]}(\omega)$ belongs to $\mathcal{D}_S$.

Remark 7.2. *To any $\mathcal{C},\mathcal{D}\subset\Omega$ we cannot associate a splitting time because the cardinality of $\{t\colon X_{[0,t]}\in\mathcal{C}, X_{[t,\zeta]}\in\mathcal{D}\}$ can be greater than one. On the other hand, a splitting time can admit several pairs of parameters, but there is a "canonical way" to choose them. We prefer present these facts in detail in another paper.*

Recall that a *stopping time* is an application $T:\Omega\mapsto\mathbb{N}\cup\{+\infty\}$ such that, for every t, $\{T=t\}$ is in the σ-field $\sigma(X_{[0,t]})$. This implies that, on $\{t\le\zeta\}$, $1_{\{T=t\}}=1_{\{T=\zeta\}}\circ X_{[0,t]}$. Thus we deduce that stopping times are exactly splitting time whose second parameter is Ω.

Symmetrically, we call *return time* any splitting time whose first parameter is Ω (or $\{\zeta<\infty\}$). As example $\tau_y=\sup\{t\colon X_t=y\}$, $\sup\varnothing=-\infty$ is a splitting time with $\bowtie=-\infty$ and parameters: $\mathcal{C}_{\tau_y}=\Omega$, $\mathcal{D}_{\tau_y}=\{\forall t\ge 1\colon X_t\ne y\}$.

Theorem 7.3 (Splitting theorem). *Let S be a splitting time with parameters $\mathcal{C}_S,\mathcal{D}_S$. Suppose $\mathbf{E}^P_x\{X_S=y\}>0$. Under $\mathbf{E}^P_x\{X\in\bullet/X_S=y\}$, pieces of trajectory $X_{[0,S]}$ and $X_{[S,\zeta]}$ are independent. Moreover:*

$$\mathbf{E}^P_x\{X_{[0,S]}\in\bullet/X_S=y\}=\mathbf{E}^{DP}_{x\rhd y}\{X\in\bullet/\mathcal{C}_S\},$$
$$\mathbf{E}^P_x\{X_{[S,\zeta]}\in\bullet/X_S=y\}=\mathbf{E}^P_y\{X\in\bullet/\mathcal{D}_S\}.$$

Proof. Take $\mathfrak{f},\mathfrak{g}$ two positive test functions. The property of the splitting times, and then, the past–future extraction (Theorem 3.5, item 1) gives:

$$\mathbf{E}^P_x\big[\mathfrak{f}(X_{[0,S]})1_{\{X_S=y\}}\mathfrak{g}(X_{[S,\zeta]})\big]$$
$$=\mathbf{E}^P_x\bigg[\sum_t\big((\mathfrak{f}1_{\mathcal{C}_S})\circ X_{[0,t]}\big)1_{\{X_t=y\}}\big((\mathfrak{g}1_{\mathcal{D}_S})\circ X_{[t,\zeta]}\big)\bigg]$$
$$=\mathbf{E}^{DP}_{x\rhd y}[\mathfrak{f}1_{\mathcal{C}_S}]U(x,y)\mathbf{E}^P_y[\mathfrak{g}1_{\mathcal{D}_S}]. \tag{5}$$

To be brief, denote by $R(\mathfrak{f},\mathfrak{g})$ the quantity above. The fact that $R(\mathfrak{f},\mathfrak{g})$ can be factorized into a term depending only on $(\mathfrak{f},y)$ and a term depending only on $(\mathfrak{g},y)$ implies conditional independence. Then, by computing $R(1_\Omega,1_\Omega)$, we see that this quantity is not zero. So we can compute $R(\mathfrak{f},1_\Omega)/R(1_\Omega,1_\Omega)$ and $R(1_\Omega,\mathfrak{g})/R(1_\Omega,1_\Omega)$ which gives us the two equations. □

Remark 7.4. *For the pleasure of the symmetry, we can restate formulae of the theorem, replacing $\mathbf{E}_x^P$ by $\mathbf{E}_{x\rhd z}^{DP}$:*

$$\mathbf{E}_{x\rhd z}^{DP}\{X_{[0,S]}\in\bullet/X_S=y\}=\mathbf{E}_{x\rhd y}^{DP}\{X\in\bullet/\mathcal{C}_S\},$$
$$\mathbf{E}_{x\rhd z}^{DP}\{X_{[S,\zeta]}\in\bullet/X_S=y\}=\mathbf{E}_{y\rhd z}^{DP}\{X\in\bullet/\mathcal{D}_S\}.$$

Applying the splitting theorem with a deterministic (stopping) time, we make the connection between inhomogeneous and homogeneous bridges:

Corollary 7.5 (Inhomogeneous bridge). *Let $x,z\in E$ and $t\in\mathbb{N}$ such that $\mathbf{E}_x^P\{X_t=z\}>0$. We have*:

$$\mathbf{E}_x^P\{X_{[0,t]}\in\bullet/X_t=z\}=\mathbf{E}_{x\rhd z}^P\{X\in\bullet/\zeta=t\}.$$

In continuous time setting, the inhomogeneous bridge is really more famous than the homogeneous one, see e.g. Fitzsimmons, Pitman and Yor [10].

Remark 7.6. *It could be a nice convention to require that stopping, return, splitting times take their values in $[0,\zeta]\cup\bowtie$: Because we only evaluate quantities of type $f(X_S),\mathfrak{f}(X_{[0,S]}),\mathfrak{f}(X_{[S,\zeta]})$, this convention does not imply any loss of generality.*

7.2. *Death and birth times*

We call death time a splitting time S whose first parameter is of the form

$$\mathcal{C}_S=\{X_0\in A_0\}\cap\{XX\subset\mathbb{A}\}\cap\{X_\zeta\in A_\zeta\}\tag{6}$$

for $A_0,A_\zeta\subset E$ and $\mathbb{A}\subset E\times E$ (see Section 6) and whose second parameter is any measurable part $\mathcal{D}_S\subset\Omega$.

Proposition 7.7. *Let S be a death time as defined above. Let α be a probability on E. Then we have*:

$$\mathbf{E}_\alpha^P\{X_{[0,S]}\in\bullet/S\in[0,\infty[\}=\mathbf{E}_{\alpha_h}^{D_hP_{\mathbb{A}}}\{X\in\bullet\},$$

where $h(x)=\mathbf{E}_x^P\{S\in[0,\infty[\}$ and $\alpha_h(x)=\frac{\alpha(x)h(x)}{\sum_a\alpha(a)h(a)}$.

Proof. Using the computation (5) we get:

$$\mathbf{E}_x^P\big[\mathfrak{f}(X_{[0,S]})1_{\{X_S=y\}}\big]=\mathbf{E}_{x\rhd y}^{DP}[\mathfrak{f}1_{\mathcal{C}_S}]U(x,y)\mathbf{E}_y^P\{\mathcal{D}_S\}.\tag{7}$$

So we have

$$h(x)=\mathbf{E}_x^P\{S\in[0,\infty[\}=\sum_y\mathbf{E}_{x\rhd y}^{DP}\{\mathcal{C}_S\}U(x,y)\mathbf{E}_y^P\{\mathcal{D}_S\}.$$

Now we choose $\mathfrak{f}=1_{\{X_0=x_0,\dots,X_n=x_n\}}$ and we apply (7):

$$\begin{aligned}&\mathbf{E}_x^P\big[\mathfrak{f}(X_{[0,S]})1_{\{S\in[0,\infty[\}}\big]\\&\quad=\sum_y\mathbf{E}_{x\rhd y}^{DP}\big\{X_0=x_0\in A_0,\dots,X_n=x_n,(XX\subset\mathbb{A},X_\zeta\in A_\zeta)\circ X_{[0,n]}\big\}U(x,y)\mathbf{E}_y^P\{\mathcal{D}_S\}\end{aligned}$$

$$= \sum_{y} I(x,x_0)P_{\mathbb{A}}(x_0,x_1)\cdots P_{\mathbb{A}}(x_{n-1},x_n)\mathbf{E}^{DP}_{x_n\rhd y}\{XX\subset\mathbb{A}, X_\zeta\subset A_\zeta\}U(x_n,y)\mathbf{E}^P_y\{\mathcal{D}_S\}$$

$$= I(x,x_0)P_{\mathbb{A}}(x_0,x_1)\cdots P_{\mathbb{A}}(x_{n-1},x_n)h(x_n) = h(x)\mathbf{E}^{D_h P_{\mathbb{A}}}_x\big[\mathfrak{f}(X)\big].$$

Thus:

$$\mathbf{E}^P_\alpha\big[\mathfrak{f}(X_{[0,S]})1_{\{S\in[0,\infty[\}}\big] = \sum_x \alpha(x)h(x)\mathbf{E}^{D_h P_{\mathbb{A}}}_x\big[\mathfrak{f}(X)\big].$$

Dividing this line by the same equation taken with $\mathfrak{f}=1_\Omega$ gives the result. □

A death time S has two special properties:

1. Under $\mathbf{E}_\alpha\{\bullet/X_S=y\}$, pieces of trajectory $X_{[0,S]}, X_{[S,\zeta]}$ are independent (this is true for any splitting time).
2. Under $\mathbf{E}_\alpha$, $X_{[0,S]}$ is still a Markov chain (see the above proposition).

Jacobsen and Pitman [19], in our discrete time setting, proved that every random time with these two properties are almost-surely equal to a death time. We advice that this is true because we work on the canonical space. More general death times can be defined by using some extra randomness, see Section 10.

Jacobsen and Pitman also established a similar statement for the "birth times": A "birth time" S is a splitting time whose first parameter is any part of Ω and second parameter is of the form $\{X_0\in A_0\}\cap\{XX\subset\mathbb{A}\}\cap\{X_\zeta\in A_\zeta\}$ with $A_0\subset E$, $\mathbb{A}\subset E\times E$ and $A_\zeta\subset E\cup\mathcal{M}$ where $\mathcal{M}$ is the Martin boundary (see Section 11). Suppose that the function $h(x):=1_{x\in A_0}\mathbf{E}_x\{XX\subset\mathbb{A}, X_\zeta\in A_\zeta\}$ is not identically zero then, under $\mathbf{E}_\alpha$, the process $X_{[S,\zeta]}$ is still a Markov chain with transition matrix $\tilde{D}_h P_{\mathbb{A}}$ (proof is left to the reader).

For a study of birth and death times in the continuous time setting, see Meyer, Smythe and Walsh [23].

7.3. *Markers*

A n-uplet of markers is a non-decreasing family of random times $\urcorner_1,\dots,\urcorner_n$ taking values in $[0,\zeta]\cup+\infty$ such that there exist $n+1$ measurable subsets of Ω: $\mathcal{A}_0,\mathcal{A}_1,\dots,\mathcal{A}_n$, called parameters, such that:

$$\text{on } \{t_n\le\zeta\}\quad \{\urcorner_1=t_1\le\urcorner_2=t_2\le\cdots\le\urcorner_n=t_n\} = \{X_{[0,t_1]}\in\mathcal{A}_0, X_{[t_1,t_2]}\in\mathcal{A}_1,\dots,X_{[t_n,\zeta]}\in\mathcal{A}_n\}. \tag{8}$$

For example a splitting-time is a 1-uplet of Markers. Remark also that the enlarged family $0,\urcorner_1,\dots,\urcorner_n,\zeta$ is a $n+2$-uplet of markers with parameters $\{\zeta=0\},\mathcal{A}_0,\dots,\mathcal{A}_n,\{\zeta=0\}$. Markers are particularly useful to describe excursions outside a random set (see Section 8.2).

Proposition 7.8. *Let $\urcorner_1,\dots,\urcorner_n$ be a n-uplet of Markers with parameters $\mathcal{A}_0,\mathcal{A}_1,\dots,\mathcal{A}_n$. Pick $a_0,a_1,\dots,a_n\in E$. Under $\mathbf{E}^P_{a_0}\{X\in\bullet/X_{\urcorner_1}=a_1,\dots,X_{\urcorner_n}=a_n\}$, the pieces of trajectory $X_{[0,\urcorner_1]}, X_{[\urcorner_1,\urcorner_2]},\dots,X_{[\urcorner_n,\zeta]}$ are independent and have law $\mathbf{E}^P_{a_0\rhd a_1}\{X\in\bullet/\mathcal{A}_0\}, \mathbf{E}^P_{a_1\rhd a_2}\{X\in\bullet/\mathcal{A}_1\},\dots,\mathbf{E}^P_{a_n}\{X\in\bullet/\mathcal{A}_n\}$.*

Proof. We do the proof for a pair $(\urcorner_1,\urcorner_2)$ of Markers. The general case is similar. Using (8) we make the following computation:

$$\begin{aligned}
&\mathbf{E}^P_{a_0}\big[\mathfrak{f}(X_{[0,\urcorner_1]})\mathfrak{g}(X_{[\urcorner_1,\urcorner_2]})\mathfrak{h}(X_{[\urcorner_2,\zeta]})\big]\\
&\quad=\sum_{t_1\le t_2}\mathbf{E}^P_{a_0}\big[(\mathfrak{f}1_{\mathcal{A}_0})\circ X_{[0,t_1]}(\mathfrak{g}1_{\mathcal{A}_1})\circ X_{[t_1,t_2]}(\mathfrak{h}1_{\mathcal{A}_2})\circ X_{[t_2,\zeta]}\big]\\
&\quad=\sum_{t_1\le t_2}\mathbf{E}^P_{a_0}\big[(\mathfrak{f}1_{\mathcal{A}_0})\circ X_{[0,t_1]}\big((\mathfrak{g}1_{\mathcal{A}_1})\circ X_{[0,t_2-t_1]}(\mathfrak{h}1_{\mathcal{A}_2})\circ X_{[t_2-t_1,\zeta]}\big)\circ X_{[t_1,\zeta]}\big]\\
&\quad=\sum_{t_1}\mathbf{E}^P_{a_0}\left[(\mathfrak{f}1_{\mathcal{A}_0})\circ X_{[0,t_1]}\left(\sum_u(\mathfrak{g}1_{\mathcal{A}_1})\circ X_{[0,u]}(\mathfrak{h}1_{\mathcal{A}_2})\circ X_{[u,\zeta]}\right)\circ X_{[t_1,\zeta]}\right].
\end{aligned}$$

Substituting $\mathfrak{f} := \mathfrak{f}1_{\{X_\zeta = a_1\}}$, $\mathfrak{g} := \mathfrak{g}1_{\{X_\zeta = a_2\}}$ and applying twice the past–future extraction we get:

$$\begin{aligned}
&\mathbf{E}^P_{a_0}\big[\mathfrak{f}(X_{[0,\urcorner_1]})1_{\{X_{\urcorner_1}=a_1\}}\mathfrak{g}(X_{[\urcorner_1,\urcorner_2]})1_{\{X_{\urcorner_2}=a_2\}}\mathfrak{h}(X_{[\urcorner_2,\zeta]})\big] \\
&\quad = \mathbf{E}^P_{a_0\rhd a_1}[\mathfrak{f}1_{\mathcal{A}_0}]U(a_0,a_1)\mathbf{E}_{a_1\rhd a_2}[\mathfrak{g}1_{\mathcal{A}_1}]U(a_1,a_2)\mathbf{E}_{a_2}[\mathfrak{h}1_{\mathcal{A}_2}].
\end{aligned}$$

The rest of the proof is the same routine as the end of the proof of the splitting Theorem 7.3. □

8. Time-change and excursion

8.1. *Factorization*

Let T be any stopping time taking values in $[1,\zeta] \cup \{+\infty\}$. We define the iterations of T by $\urcorner_0 = 0$, $\urcorner_1 = T, \ldots, \urcorner_{n+1} = \urcorner_n + T(X_{[\urcorner_n,\zeta]})$. We see that $\urcorner_n < \infty \Rightarrow \urcorner_n < \urcorner_{n+1}$. In particular, on $\{\zeta < \infty\}$, the number of finite $\urcorner_n$ is finite and we write $\urcorner_f = \max\{\urcorner_n \colon \urcorner_n < \infty\}$.

Applying the strong Markov property at the stopping times $\urcorner_n$, we check easily that, under $\mathbf{E}^P_x$, the process $n \mapsto X_{\urcorner_n}$ is a Markov chain. Let us write $V_T(x,y) = \mathbf{E}^P_x[\sum_n 1_{\{X_{\urcorner_n}=y\}}]$ its potential matrix and $W_T(x,y) := \mathbf{E}^P_x[\sum_{t<T} 1_{\{X_t=y\}}]$. The application of the strong Markov property leads easily to the factorization $U = V_T W_T$. The next theorem gives an original interpretation of this factorization (and also a less elementary second proof of it).

Theorem 8.1. *For $x, y, z \in E$ we have*:

$$\mathbf{E}^P_{x\rhd z}\{X_{\urcorner_f} = y\}U(x,z) = \mathbf{E}^P_x\Big[\sum_n 1_{\{X_{\urcorner_n}=y\}}\Big]\mathbf{E}^P_y\Big[\sum_{t<T} 1_{\{X_t=z\}}\Big].$$

Remark 8.2. *By summing on all y we get $U = V_T W_T$.*

Proof of Theorem 8.1. Firstly, by the past–future (Theorem 3.5) extraction applied to $\mathbf{E}^P_{x\rhd z}$:

$$\begin{aligned}
\mathbf{E}^P_{x\rhd z}\{X_{\urcorner_f} = y\} &= \sum_n \mathbf{E}^P_{x\rhd z}\{X_{\urcorner_n} = y, \urcorner_{n+1} = \infty\} \\
&= \sum_t \sum_n \mathbf{E}^P_{x\rhd z}\big\{\urcorner_n = t, X_t = y, t + T(X_{[t,\zeta]}) = \infty\big\} \\
&= \sum_t \sum_n \mathbf{E}^P_{x\rhd z}\big\{\urcorner_n(X_{[0,t]}) = t, X_t = y, T(X_{[t,\zeta]}) = \infty\big\} \\
&= \sum_n \mathbf{E}^P_{x\rhd y}\{X_{\urcorner_n} = y\}\frac{U(x,y)U(y,z)}{U(x,z)}\mathbf{E}^P_{y\rhd z}\{T = \infty\}.
\end{aligned} \tag{9}$$

Secondly, by the past extraction applied to $\mathbf{E}^P_x$:

$$\begin{aligned}
\sum_n \mathbf{E}^P_x\{X_{\urcorner_n} = y\} &= \sum_t \sum_n \mathbf{E}^P_x\big\{\urcorner_n(X_{[0,t]}) = t, X_t = y\big\} \\
&= \sum_n \mathbf{E}^P_{x\rhd y}\{X_{\urcorner_n} = y\}U(x,y).
\end{aligned} \tag{10}$$

Thirdly, because T is a stopping time taking values in $[0,\zeta] \cup \{+\infty\}$, on $\{\zeta < t\}$ we have $1_{\{T>t\}} = 1_{\{T=\infty\}} \circ X_{[0,t]}$. Then, by the past extraction applied to $\mathbf{E}^P_y$:

$$\mathbf{E}^P_y\Big[\sum_{t<T} 1_{\{X_t=z\}}\Big] = \mathbf{E}^P_y \sum_t (1_{\{T=\infty\}} \circ X_{[0,t]})1_{\{X_t=z\}} = \mathbf{E}^P_{y\rhd z}\{T = \infty\}U(y,z). \tag{11}$$

To gather formulae (9), (10), (11), gives the result. □

8.2. *Past-spotted sets*

We write $\Lambda = \{t\colon \exists n\colon \daleth_n = t\}$. This random set satisfies:

$$\text{“support”} \qquad 0 \in \Lambda \quad \text{and} \quad \Lambda \subset [0, \zeta], \tag{12}$$

$$\text{“past adaptation”} \qquad \forall t \in [0, \zeta] \quad (1_\Lambda)_{[0,t]} = (1_\Lambda) \circ X_{[0,t]}, \tag{13}$$

$$\text{“future homogeneity”} \qquad \forall t \in \Lambda \quad (1_\Lambda)_{[t,\infty[} = (1_\Lambda) \circ X_{[t,\zeta]}. \tag{14}$$

Reciprocally, if Λ is any random subset of $\mathbb{N}$ satisfying the above three conditions, then it is equal to the reunion of the graphs of the iteration of $T = \inf\{t \in \Lambda \cap [1, \infty[\}$ (the proof of this fact is left to the reader as an amusing exercise). We call such a sets Λ a *past-spotted set.*

Remark 8.3. *Even if the “future homogeneity” looks like as a dual axiom of “past adaptation,” the notion of past-spotted set is not symmetric for the time-reversion. In another paper, we will present the symmetrized notion.*

We can compare past-spotted sets with sets of type $\Gamma_A := \{t\colon X_t \in A\}$, $A \subset E$, which satisfy:

$$\text{“support}'\text{”} \qquad \Gamma_A \subset [0, \zeta], \tag{15}$$

$$\text{“past adaptation”} \qquad \forall t \in [0, \zeta] \quad (1_{\Gamma_A})_{[0,t]} = (1_{\Gamma_A}) \circ X_{[0,t]}, \tag{16}$$

$$\text{“future homogeneity}'\text{”} \qquad t \in [0, \zeta] \quad (1_{\Gamma_A})_{[t,\infty[} = (1_{\Gamma_A}) \circ X_{[t,\zeta]}. \tag{17}$$

Sets of type Γ_A are usually called *regenerative sets*. Of course $\{0\} \cup \Gamma_A$ is the past-spotted set constructed from the iterations of $T_A^* = \inf\{t \geq 1\colon X_t \in A\}$.

We denote by:

$$G_t = \max\bigl(\Lambda \cap [0, t]\bigr), \quad \max\varnothing = 0, \qquad D_t = \min\bigl(\Lambda \cap]t, \infty[\bigr), \quad \min\varnothing = \zeta.$$

An *excursion interval* is a random interval $[g, d]$ of type $[G_L, D_L]$ for L random time. The first one is $[0, \daleth_1]$, last one is $[\daleth_f, \zeta]$. An excursion is a piece of trajectory of type $X_{[g,d]}$.

8.3. *Some excursions are easy to describe*

When extremities g, d of an excursion interval are markers with parameters $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2$ (see Section 7.3), then Proposition 7.8 indicates that, under $\mathbf{E}_\alpha^P\{X \in \bullet / X_g = a, X_d = b\}$, the pieces of trajectory $X_{[0,g]}, X_{[g,d]}, X_{[d,\zeta]}$ are independent and moreover:

$$\mathbf{E}_\alpha^P\{X_{[g,d]} \in \bullet / X_g = a, X_d = b\} = \mathbf{E}_{a \triangleright b}^P\{X \in \bullet / \mathcal{A}_1\}.$$

We give examples of such good situations:

- The nth first excursion interval $[\daleth_n, \daleth_{n+1}]$ for $n \geq 0$. Parameters are $\{\daleth_n = \zeta\}, \{T = \zeta\}, \Omega$.
- The latest excursion interval $[\daleth_f, \zeta]$. Parameters are $\{\zeta \in \Lambda\}, \{T = \infty\}, \{\zeta = 0\}$.
- The nth latest excursion interval $[\daleth_{f-n}, \daleth_{f-n+1}]$ for $n \geq 1$. Parameters are $\{\zeta \in \Lambda\}, \{T = \zeta\}, \{\daleth_n < \infty, \daleth_{n+1} = \infty\}$.
- The first excursion interval where X enters in $B \subset E$. Parameters are $\{\zeta \in \Lambda, \forall t < \zeta \ X_t \notin B\}, \{T = \zeta, \exists t\colon X_t \in B\}, \Omega$.
- The excursion interval with length $d - q = 4$ and such as $\zeta - d = 7$. Parameters are $\{\zeta \in \Lambda\}, \{T = \zeta = 4\}, \{\zeta = 7\}$.

Morality: each excursion interval which we can spot by some events before its left extremity, after its right extremity, between the two extremities, and by intersections of these types of events, can be described using the markers technique.

8.4. *Some excursions are difficult to describe*

Notation 8.4. *When ω, w are trajectories such that $\omega_\zeta = w_0$, we denote by $\omega \wr w$ their biting-concatenation*: $\omega_0\omega_1 \cdots \omega_\zeta w_1 w_2 \cdots$ (*remark that we have suppressed w_0 in this special concatenation*).

We considerate L any stopping time. The following calculus shows us that, conditionally to X_{G_L}, pieces of trajectory $X_{[0,G_L]}$, $X_{[G_L,D_L]}$ are not independent in general. For shortness we write $\mathfrak{f}_a = \mathfrak{f} 1_{\{X_\zeta = a\}}$. Let us start the calculus:

$$\begin{aligned}
&\mathbf{E}^P_\alpha\big[\mathfrak{f}(X_{[0,G_L]})1_{\{X_{G_L}=a\}}\mathfrak{g}(X_{[G_L,D_L]})\big] \\
&\quad= \mathbf{E}^P_\alpha\Bigg[\sum_{t\in\Lambda\cap[0,L]} 1_{\{L\le D_t\}}\mathfrak{f}_a(X_{[0,t]})\mathfrak{g}(X_{[t,D_t]})\Bigg] \\
&\quad= \mathbf{E}^P_\alpha\Bigg[\sum_{t\in\Lambda\cap[0,L]} 1_{\{L\le\zeta\}}\circ(X_{[0,t]}\wr X_{[t,D_t]})\mathfrak{f}_a(X_{[0,t]})\mathfrak{g}(X_{[t,D_t]})\Bigg].
\end{aligned}$$

Using $t \in \Lambda \Rightarrow X_{[t,D_t]} = X_{[0,T]} \circ X_{[t,\zeta]}$, and using the past extraction we get:

$$= \int_{\Omega\times\Omega} 1_{\{\zeta\in\Lambda\cap[0,L]\}}(\omega)1_{\{L\le\zeta\}}\big(\omega \wr X_{[0,T]}(w)\big)\mathfrak{f}_a(\omega)\mathfrak{g}\big(X_{[0,T]}(w)\big)\mathbf{E}^{DP}_{\alpha\rhd a}\{X\in \mathrm{d}\omega\}\mathbf{E}^P_a\{X\in \mathrm{d}w\}.$$

Because of the term $1_{\{L\le\zeta\}}(\omega \wr X_{[0,T]}(w))$ this expression cannot be factorized (in general) and thus the conditional independence is not always satisfied.

Meanwhile, in the case $L = T_B$, for $t \le T_B$ we have $1_{\{T_B\le\zeta\}} = 1_{T_B\le\zeta} \circ X_{[t,\zeta]}$ which allows the factorization. But this case has already been described in the previous subsection.

8.5. *Further comments about excursions*

Let us consider $\Gamma_x = \{t\colon X_t = x\}$. In this particular case, the process of successive excursions $n \mapsto X_{[\urcorner_n, \urcorner_{n+1}]}$, under $\mathbf{E}^P_x$, can be described as an i.i.d. sequence of random elements of law $\mathbf{E}_x\{X_{[0,T^*_x]} \in \bullet\}$, killed in a geometrical time with parameter $\mathbf{E}_x\{T^*_x = \infty\}$.

The Itô [17] theory of excursions concerns the set Γ_x in continuous time. The situation is much more complicated because excursions accumulate. The process of excursions is described by a Poisson point process, while the law of each excursion is described by a Markovian measure with infinite total mass. The set Γ_x itself is described as the range of a subordinator (an increasing process with stationary and independent increments). For a recent survey see Pitman–Yor [26].

In many papers, the word "regenerative set" designates sets of type Γ_x. Here, following Hoffmann–Jorgensen [14], Maisonneuve [22], we have called "regenerative" the random sets of type $\Gamma_A = \{t\colon X_t \in A\}$. Such sets are often described by axioms without mentioning any subjacent Markov process. Descriptions of excursions outside such regenerative sets can be found in Fitzsimmons [8] and Getoor–Sharpe [13]. This last paper is the closest to ours: Fitzsimmons used non-normalized (homogeneous) bridges to describe excursions.

The study of regenerative sets is inseparable from that of co-terminal times i.e. times of type $\sup \Gamma_A$ (while terminal times are times of type $\inf \Gamma_A$). See Pittenger–Shih [27], Getoor–Sharpe [12]. Of course, in our context, co-terminal times can be described as splitting times with parameters $\mathcal{C} = \{X_\zeta \in A\}$, $\mathcal{D} = \{T^*_A = \infty\}$.

The recurrent regenerative sets also have many interesting aspects: we can compute asymptotes, and use the Palm duality to obtain some stationary versions of these sets. For a complete treatment, both in discrete and continuous time, we sent the reader to the book by Thorison [31].

9. Fluctuation theory

Let ρ be the minimizer of the Brownian motion B on the random interval $[0,\sigma]$ where $\sigma = \sup\{t \le T_{-1}\colon B_t = 0\}$ (and as usual $T_{-1} = \inf\{t \ge 0\colon B_t = -1\}$). Williams' decomposition [34] indicates that components $B_{[0,\rho]}$ and $B_{[\rho,\sigma]}$ are independent conditionally to B_ρ (Williams also describes the law of the two components). But actually, we can consider $B_{[0,\sigma]}$ as simply a Markov process and expect that the result is generalizable: For a càdlàg Markov process X, taking values in $\mathbb{R}$, with a finite minimizer ρ, are $X_{[0,\rho]}$ and $X_{[\rho,\zeta]}$ independent conditionally to $(X_\rho, X_{\rho-})$? Millar [24] gives an affirmative answer to this question providing that the germ σ-field at ρ is degenerate. In particular, this last condition is fulfilled by all Lévy processes. Is this case, the path decomposition gives rise to the famous Wiener–Hopf factorization: see Bertoin [1], Chapter 6, Vigon [33], and, for random walks, Feller [7], Chapter 12.

Now we come back to our discrete world.

Data 9.1. *Throughout this section, we take $E = \mathbb{Z}$.*

9.1. LU-*factorization*

We keep all the notations from the previous section, but we choose as stopping time: $T = \inf\{t\colon\ X_t < X_0\}$. In this situation, the past-spotted set Λ is the set where the process reaches its left minima. Its end $\urcorner_f = \sup \Lambda := \rho$ is the first time where the process X reaches its global minima. The time-changed process $n \mapsto X_{\urcorner_n}$ is strictly decreasing and its potential matrix V_T is triangular inferior with 1 on the diagonal. On the other hand, $W_T(x, y) = \mathbf{E}_x^P \sum_{t<T} 1_{\{X_t=y\}}$ is triangular superior. Now it is clear that the factorization $U = V_T W_T$ is one of the LU-factorizations of U (also called Wiener–Hopf factorization in the context of random-walks). Another LU-factorization can be obtained by starting our procedure with $T' = \inf\{t\colon\ X_t \le X_0\}$, in this case $\rho' = T'_f$ is the last minimizer.

So the first minimizer ρ is a splitting time, whose parameters are $\mathcal{C}_\rho = \{\forall s < \zeta\colon\ X_s > X_\zeta\}$, $\mathcal{D}_\rho = \{\forall s\colon\ X_s \ge X_0\}$. According to the splitting Theorem 7.3, the processes before and after ρ are independent conditionally to X_ρ and:

$$\mathbf{E}_x^P\{X_{[0,\rho]} \in \bullet / X_\rho = y\} = \mathbf{E}_{x \triangleright y}^{DP}\{X \in \bullet / \forall s < \zeta\ X_s > X_\zeta\}, \tag{18}$$

$$\mathbf{E}_x^P\{X_{[\rho,\zeta]} \in \bullet / X_\rho = y\} = \mathbf{E}_y^P\{X \in \bullet / \forall s\ X_s \ge X_0\}. \tag{19}$$

These formulae indicate that pre- and post-minimizer trajectories are really ordinary pieces of trajectory: you simply condition them by the natural constraint. This fact is well illustrated by the result of the next subsection.

9.2. *Vervaat transformation*

The next theorem is a Markov chain version of the Vervatt transform for Brownian motion (cf. Vervaat [32]). The generalization of this transformation to Lévy processes was made by Fourati [11]. In the last paragraph, Fourati also explain how to extend this transformation to Markov processes. Recall that $\wr$ stands for the biting-concatenation (see Notation 8.4).

Theorem 9.2. *Let $x \ge y$. We have*:

$$\mathbf{E}_{x \triangleright x}^{DP}\big[\mathfrak{f}(X_{[\rho,\zeta]} \wr X_{[0,\rho]}) / X_\rho = y\big] = \mathbf{E}_{y \triangleright y}^{DP}\big[\mathfrak{f}(X) / \exists t\colon\ X_t = x, \forall s\ X_s \ge y\big]$$

for all states x, y, all test functions $\mathfrak{f}\colon \Omega \mapsto \mathbb{R}_+$.

The next figure, Fig. 1, illustrates the previous theorem.

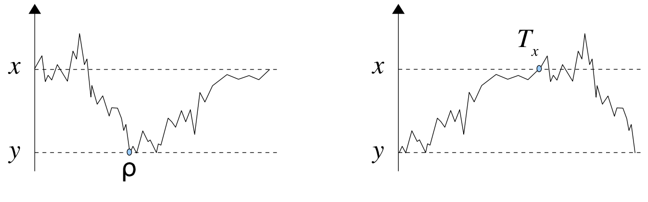

Fig. 1. On the left, a typical trajectory of $(X_{[\rho,\zeta]} \wr X_{[0,\rho]})$ under the probability $\mathbf{E}_{x \triangleright x}^{DP}\{\bullet / X_\rho = y\}$. On the right, a typical trajectory of X under the probability $\mathbf{E}_{y \triangleright y}^{DP}\{\bullet / \exists t\colon X_t = x, \forall s\colon X_s \ge y\}$.

Proof. We consider S the splitting time with parameters

$$\mathcal{C}_S = \{\forall t\ X_t \geq y, X_\zeta = x\}, \qquad \mathcal{D}_S = \{\forall t < \zeta\ X_t > y, X_0 = x\}.$$

Clearly, under $\mathbf{E}^P_{y \rhd y}$, the three events $\{\exists t\colon\ X_t = x, \forall s\ X_s \geq y\}$, $\{S < \infty\}$, $\{X_S = x\}$ almost surely coincide. From the splitting Theorem 7.3:

$$\begin{aligned}
&\mathbf{E}^{DP}_{y \rhd y}\big[\mathfrak{f}(X)/\exists t\colon\ X_t = x, \forall t\ X_t \geq y\big] = \mathbf{E}^{DP}_{y \rhd y}\big[\mathfrak{f}(X)/X_S = x\big] \\
&\quad = \int_{\Omega^2} \mathfrak{f}(\omega \wr w)\mathbf{E}^{DP}_{y \rhd x}[X \in \mathrm{d}\omega/\forall t\ X_t \geq y, X_\zeta = x]\mathbf{E}^{DP}_{x \rhd y}[X \in \mathrm{d}w/\forall t < \zeta\ X_t > y, X_0 = x] \\
&\quad = \int_{\Omega^2} \mathfrak{f}(\omega \wr w)\mathbf{E}^{DP}_{y \rhd x}[X \in \mathrm{d}\omega/\forall t\ X_t \geq y]\mathbf{E}^{DP}_{x \rhd y}[X \in \mathrm{d}w/\forall t < \zeta\ X_t > y] \\
&\quad = \int_{\Omega^2} \mathfrak{f}(\omega \wr w)\mathbf{E}^{DP}_{y \rhd x}[X_{[\rho,\zeta]} \in \mathrm{d}\omega]\mathbf{E}^{DP}_{x \rhd y}[X_{[0,\rho]} \in \mathrm{d}w]
\end{aligned}$$

for the last step we used the bridge versions of (18) and (19). $\square$

10. Extension of the universe

10.1. *Adding an independent variable*

Markov chains can also be defined on some probability space $(\overline{\Omega}, \overline{\mathbf{E}})$ which is not the canonical one. This makes it possible to define some random elements independent of the Markov chain. By some classical coupling arguments (see Kallenberg [20], Chapter 5) the more general situation can described by the addition of only one independent random variable with a continuous law.

Let $\overline{\Omega} = \Omega \times [0, 1]$, the two canonical projections are denoted by X, ϑ. On $\overline{\Omega}$ we define the probability $\overline{\mathbf{E}}^P_x = \mathbf{E}^P_x \otimes \mathrm{d}v$, where $\mathrm{d}v$ is the Lebesgue measure on $[0, 1]$. We define similarly $\overline{\mathbf{E}}^{DP}_{x \rhd z}$. Thus, under $\overline{\mathbf{E}}^P_x$, X is a Markov chain, ϑ is a $\mathcal{U}[0, 1]$ variable, and X, ϑ are independent. ϑ represents all the extra randomness we need.

10.2. *Randomized time*

Random elements defined on $\overline{\Omega}$ are called *randomized variables*. They can be written $Z = \dot{Z}(X, \vartheta)$ (the "dot" just indicates that $\dot{Z}$ is an application that helps to understand Z). Let us consider a randomized time $S = \dot{S}(X, \vartheta) : \overline{\Omega} \mapsto \mathbb{N} \cup \{\bowtie\}$. We have:

$$\overline{\mathbf{E}}^P_\alpha\big[\mathfrak{f}(X_{[0,S]})1_{\{X_S = y\}}\mathfrak{g}(X_{[S,\zeta]})\big] = \mathbf{E}^P_\alpha\Bigg[\sum_t \ell_t(X)\mathfrak{f}(X_{[0,s]})1_{\{X_s = y\}}\mathfrak{g}(X_{[s,\zeta]})\Bigg],$$

where $\ell_t(X) = \int_0^1 1_{\{\dot{S}(X,v) = t\}}\,\mathrm{d}v$. We see that an estimation made at a randomized time can be interpreted by a "weighted estimation," but without leaving the canonical space.

Reciprocally, every weighted estimation made with weight $\ell_t(X)$ such that $\sum_t \ell_t(X) \leq 1$ can also be interpreted by an estimation made at a randomized time defined by $\dot{S}(X, v) = \inf\{t\colon\ \sum_{s \leq t} \ell_t(X) \geq v\}$, $\inf \varnothing = \bowtie$.

The weighted estimations, and particularly the ones made by "additive functionals," are very important in the continuous time setting, see Blumenthal–Getoor [3], Chapter IV.

A randomized time S is a randomized splitting time such that there exist $\overline{\mathcal{C}}_S, \overline{\mathcal{D}}_S$, two measurable parts of $\overline{\Omega}$, such that:

$$\text{on } \{t \leq \zeta\} \quad \{S = t\} = \big\{(X_{[0,t]}, \vartheta_1) \in \overline{\mathcal{C}}_S\big\} \cap \big\{(X_{[t,\zeta]}, \vartheta_2) \in \overline{\mathcal{D}}_S\big\},$$

where ϑ_1, ϑ_2 are two independent uniform variables, which are functions of ϑ (thus independent of X). The randomized stopping times are splitting times with second parameters $\overline{\mathcal{D}} = \overline{\Omega}$.

We can easily see that the splitting theorem is still valid in this context.

10.3. *Randomized time-change*

Let $T = \dot{T}(X, \vartheta)$ be a randomized stopping time. Let (ϑ_n) be an i.i.d. sequence of U[0, 1] constructed from ϑ (e.g. by extracting different sequences of decimal of the number ϑ). We "iterate" T as follows:

$$
\begin{aligned}
&\daleth_0 = 0,\\
&\daleth_1 = \dot{\daleth}_1(X, \vartheta_1) = \dot{T}(X, \vartheta_1),\\
&\daleth_2 = \dot{\daleth}_2(X, \vartheta_1, \vartheta_2) = \daleth_1 + \dot{T}(X_{[\daleth_1,\zeta]}, \vartheta_2),\\
&\daleth_3 = \dot{\daleth}_3(X, \vartheta_1, \vartheta_2, \vartheta_3) = \daleth_2 + \dot{T}(X_{[\daleth_2,\zeta]}, \vartheta_3), \quad \text{etc.}
\end{aligned}
$$

We see that under $\overline{\mathbf{E}}_x^P$, the process $n \mapsto X_{\daleth_n}$ is a Markov chain. We add the supplementary assumption that $\forall x$ $\overline{\mathbf{E}}_x^P\{T \geq 1\} > 0$ which implies that a.s. the trajectories $n \mapsto \daleth_n$ cannot be stopped at an integer $k < \infty$. In particular, on $\{\zeta < \infty\}$, the number $\daleth_n$ is finite and we denote by $\daleth_f$ the latest finite $\daleth_n$.

Proposition 10.1. *Fix $x, y, z \in E$. We have*:

$$
\overline{\mathbf{E}}_{x \rhd z}^P\{X_{\daleth_f} = y\} U(x, z) = \overline{\mathbf{E}}_x^P\Big[\sum_n 1_{\{X_{\daleth_n} = y\}}\Big] \overline{\mathbf{E}}_x^P\Big[\sum_{t<T} 1_{\{X_t = y\}}\Big].
$$

Proof. By construction $\dot{\daleth}(\bullet, \vartheta_1, \ldots, \vartheta_n)$ are stopping times so that $\{\daleth_n = t\} = \{\dot{\daleth}(X_{[0,t]}, \vartheta_1, \ldots, \vartheta_n) = t\}$. Using the past extraction and the independence between $(\vartheta_1, \ldots, \vartheta_n)$ and ϑ_{n+1} we get:

$$
\begin{aligned}
&\overline{\mathbf{E}}_{x \rhd z}^P\{X_{\daleth_f} = y\}\\
&\quad = \sum_n \overline{\mathbf{E}}_{x \rhd z}^P\{X_{\daleth_n} = y, \daleth_{n+1} = \infty\}\\
&\quad = \sum_n \sum_t \overline{\mathbf{E}}_{x \rhd z}^P\big\{\dot{\daleth}_n(X_{[0,t]}, \vartheta_1 \cdots \vartheta_n) = t, X_t = y, \dot{T}(X_{[t,\zeta]}, \vartheta_{n+1}) = \infty\big\}\\
&\quad = \sum_n \sum_t \overline{\mathbf{E}}_{x \rhd z}^P\big\{\dot{\daleth}_n(X_{[0,t]}, \vartheta_1 \cdots \vartheta_n) = t, X_t = y\big\} \overline{\mathbf{E}}_y^P\big\{\dot{T}(X, \vartheta) = \infty\big\}.
\end{aligned}
$$

The rest of the proof follows the same way as the end of the proof of Theorem 8.1. $\square$

10.4. *The comparison problem*

Proposition 10.2. *Let $\mathbb{A}: E \times E \mapsto [0, 1]$. Let $\vartheta_0, \vartheta_1, \ldots$ an i.i.d. sequence of* U[0, 1], *constructed from ϑ. Let $S_{\mathbb{A}} = \inf\{t\colon \vartheta_t > \mathbb{A}(X_t, X_{t+1})\}$ We have*

$$
\overline{\mathbf{E}}_x^P\{X_{[0,S_{\mathbb{A}}]} \in \bullet\} = \mathbf{E}_x^{P_{\mathbb{A}}}\{X \in \bullet\},
$$

where $P_{\mathbb{A}}(x, y) = P(x, y)\mathbb{A}(x, y)$.

Proof. It is sufficient to check the equality for function of type $\mathfrak{f} = 1_{\{X_0 = x_0, \ldots, X_n = x_n\}}$:

$$
\begin{aligned}
&\overline{\mathbf{E}}_x^P\big[\mathfrak{f}(X_{[0,S_{\mathbb{A}}]})\big]\\
&\quad = \overline{\mathbf{E}}_x^P\{X_0 = x_0, \ldots, X_n = x_n, n \leq S_{\mathbb{A}}\}\\
&\quad = \overline{\mathbf{E}}_x^P\big\{X_0 = x_0, \ldots, X_n = x_n, \vartheta_1 \leq \mathbb{A}(x_0, x_1), \ldots, \vartheta_n \leq \mathbb{A}(x_{n-1}, x_n)\big\}\\
&\quad = P(x_0, x_1) \cdots P(x_{n-1}, x_n)\mathbb{A}(x_0, x_1) \cdots \mathbb{A}(x_{n-1}, x_n).
\end{aligned}
$$

$\square$

Remark 10.3. *Of course the murder $X \mapsto X_{[0,S_{\mathbb{A}}]}$ is a generalization of the one introduced in Section 6. Such a murder can be assimilated with murder by multiplicative functional, see Blumental and Getoor [3], Chapter* III. *The reader can also extend Theorem 6.1 in the extended universe.*

Corollary 10.4. *Let P, Q be two sub-stochastic matrices. $P \geq Q$ if and only if there exists a randomized time S such that $\overline{\mathbf{E}}_x^P\{X_{[0,S]} \in \bullet\} = \mathbf{E}_x^Q\{X \in \bullet\}$.*

Proof. When $P \geq Q$ then we can find $\mathbb{A}$: $E \times E \in [0,1]$ such that $P_{\mathbb{A}} = Q$, so the time $S = S_{\mathbb{A}}$ works.
Reciprocally, $Q(x,y) = \mathbf{E}_x^Q\{X_1 = y\} = \overline{\mathbf{E}}_x^P\{X_1 = y, 1 \leq S\} \leq P(x,y)$. □

Comments. *Consider P, Q two sub-stochastic matrices. We see easily that:*

$$P \geq Q \quad \Rightarrow \quad U_{[P]} \geq QU_{[P]}.$$

We saw that the left-hand side inequality can be interpreted by a murder.

Let us suppose now the right-hand side inequality: Rost [28] shows that this inequality implies the existence of a randomized stopping time T such that $Q(x,y) = \overline{\mathbf{E}}_x^P\{X_T = y\}$. Denoting by $\urcorner$ the iteration of this stopping time, we get the factorization $U_{[P]} = U_{[Q]}W$ where $W(x,y) = \overline{\mathbf{E}}_x^P[\sum_{t<T} 1_{\{X_t = y\}}]$. Of course, this factorization can also be directly deduced by putting $W = U_{[P]} - QU_{[P]} \geq 0$ and then:

$$U_{[P]} = W + QU_{[P]} = W + QW + Q^2U_{[P]} = \cdots = U_{[Q]}W.$$

To summarize, we have presented two ways to compare Markov chains: the first one is interpreted by a murder, the second one by a time-change. Of course, murder is a very particular case of time-change.

In a continuous-time setting, the interpretation of $U_{[P]} = U_{[Q]}W$ in terms of time-change as been done by Simon [30] for Lévy processes. The case of Markov processes has not been treated yet.

11. Extension to the Martin boundary

Data 11.1. *Throughout this section, E is infinite and the undirected graph of P is irreducible and locally finite (i.e. there is a finite number of edges attached to each vertex).*

11.1. *A few reminders*

Here we will recall definitions and basic results about the Martin Boundary. For a complete exposition, we send the reader to Dynkin [6], Hunt [16] or to the book of Woess [35]; our notations and hypotheses are the same as the ones in this book.

We choose any reference point $o \in E$. The Martin kernel is defined by:

$$K(x,y) = K_{[P,o]}(x,y) = \frac{U_{[P]}(x,y)}{U_{[P]}(o,y)} = \frac{\mathbf{E}_x^P\{T_y < \infty\}}{\mathbf{E}_o^P\{T_y < \infty\}}.$$

A sequence (y_n) is a Martin–Cauchy sequence if it leaves every finite subset of E, and if, for all x, the sequence $n \mapsto K(x,y_n)$ converges. Two Martin–Cauchy sequences (y_n) and (y'_n) are equivalent if for all x, we have $\lim_n K(x,y_n) = \lim_n K(x,y'_n)$. The Martin boundary does not depend on the reference point. Because $K_{[D_hP]}(x,y) = \frac{h(o)}{h(x)} K_{[P]}(x,y)$, the Martin boundary is the same for every matrix in DP. So we can denote the Martin boundary by $\mathcal{M} = \mathcal{M}_{[DP]}$.

Elements of $\mathcal{M}$ will be generically denoted by ξ. Elements of $E \cup \mathcal{M}$ will be generically denoted by η. On $E \cup \mathcal{M}$, we put the natural topology (e.g. y_n converges to $\xi \in \mathcal{M}$ if and only if y_n is an element of the equivalence class ξ) which makes $E \cup \mathcal{M}$ a compact and metrizable space. The Martin kernel K is prolonged by continuity on $\mathcal{M}$. It is easy to verify that the excessive function $x \mapsto K(x,\eta)$ (defined on E) is potential if $\eta \in E$ and invariant if $\eta \in \mathcal{M}$.

Remark 11.2. *As a consequence, under our hypothesis, there always exists an invariant function (i.e. the Martin boundary has at least one point).*

An excessive function h is said to be minimal if, for every excessive function h' (non-identically zero), the inequality $h' \le h$ implies that h' and h are proportional. An excessive function h is said to be normalized if $h(o) = 1$. The set of points $\xi \in \mathcal{M}$ such that $K(\bullet, \xi)$ is a minimal invariant function is denoted by $\mathcal{M}_{\min}$. This set is Borelian in $\mathcal{M}$. In some classical cases (e.g. Markov chain whose undirected graph is a tree, or random walks in $\mathbb{Z}^d$) the sets $\mathcal{M}$ and $\mathcal{M}_{\min}$ are equal. The set $\mathcal{M}_{\min}$ is more important for us, because it is the "target" where the canonical process finishes its life. This is the subject of the following theorem:

Theorem 11.3 (Convergence to the boundary). *Under $\mathbf{E}_x^P$, almost surely on $\{\zeta = \infty\}$, the canonical process $t \mapsto X_t$ converges to a random variable X_∞ taking values in $\mathcal{M}_{\min}$. Moreover, we have*:

$$\mathbf{E}_x^P\{X_\zeta = y, \zeta < \infty\} = K(x, y)\mathbf{E}_o^P\{X_\zeta = y, \zeta < \infty\},$$
$$\mathbf{E}_x^P\{X_\zeta \in \mathrm{d}\xi, \zeta = \infty\} = K(x, \xi)\mathbf{E}_o^P\{X_\zeta \in \mathrm{d}\xi, \zeta = \infty\}.$$

Remark 11.4. *The first equation of the theorem is an immediate consequence of the equality $\mathbf{E}_x^P\{X_\zeta = y, \zeta < \infty\} = U(x, y)P(y, \dagger)$ established in Lemma 3.2. The second one can be seen as a prolongation of the first one. Both can be summarized as follows*:

$$\mathbf{E}_x^P\{X_\zeta \in \mathrm{d}\eta\} = K(x, \eta)\mathbf{E}_o^P\{X_\zeta \in \mathrm{d}\eta\} \tag{20}$$

with the natural convention: $X_\zeta = X_\zeta 1_{(\zeta<\infty)} + X_\infty 1_{(\zeta=\infty)}$.

The proof of the previous theorem, as well as the proof of the next one, can be readed in Woess [35], *Chapter* IV, *Section* 24.

Theorem 11.5 (Representation with unicity). *For every excessive function h, there exists a unique finite measure μ, supported by $E \cup \mathcal{M}_{\min}$ such that*:

$$h(x) = \int_{E\cup\mathcal{M}} K(x, \eta)\mu(\mathrm{d}\eta). \tag{21}$$

This measure is given by

$$\mu(\mathrm{d}\eta) = \mathbf{E}_o^{D_h P}\{X_\zeta \in \mathrm{d}\eta\}.$$

Corollary 11.6. *The set of normalized minimal invariant functions is exactly $\{K(\bullet, \xi)\colon \xi \in \mathcal{M}_{\min}\}$.*

Remark 11.7. *Of course, the equation* (21) *can be split into*:

$$h(x) = \sum_{a\in E} K(x, a)\mu(a) + \int_{\mathcal{M}} K(x, \xi)\mu(\mathrm{d}\xi),$$

which gives the Riesz decomposition of h.

11.2. *Prolongation of the bridge*

Let us take $\xi \in \mathcal{M}_{\min}$. Let us consider the normalized minimal invariant function $K(\bullet, \xi)$. Because of its invariant property, under $\mathbf{E}^{D_{K(\bullet,\xi)}P}$, the canonical process X has an infinite life. According to Theorem 11.3, this process converges a.s. to $X_\infty \in \mathcal{M}$. Moreover we have obviously:

$$K(x, \xi) = \int_{E\cup\mathcal{M}} K(x, \eta)\delta_\xi(\mathrm{d}\eta).$$

From the unicity of the measure in Theorem 11.5, we have:

$$\mathbf{E}_o^{D_{K(\bullet,\xi)}P}\{X_\infty \in \mathrm{d}\eta\} = \delta_\xi(\mathrm{d}\eta).$$

We can summarize this as follows: Every normalized minimal invariant function $K(\bullet,\xi)$ leads the canonical process to a unique point $\xi \in \mathcal{M}_{\min}$ (in an infinite time); as well as, every normalized minimal potential function $K(\bullet,z)$ leads the canonical process to a unique state $z \in E$ (in a finite time).

Here is the extension of the fundamental Lemma 3.4. The proof is similar and even simpler because now, the irreducible assumption, allows to replace $\tilde{D}$ by D.

Lemma 11.8. *Fix $\eta \in E \cup \mathcal{M}_{\min}$. Let P' be a matrix of DP. We have*:

$$D_{K(\bullet,\eta)}P = D_{K'(\bullet,\eta)}P',$$

where $K' = K_{[P',o]}$ and $K = K_{[P,o]}$.

All these facts suggest to us the following definition:

Definition 11.9. *For every $\eta \in E \cup \mathcal{M}_{\min}$ we denote by $P_{\rhd\eta} = D_{K(\bullet,\eta)}P$ and $\mathbf{E}^{DP}_{x\rhd\eta} = \mathbf{E}^{P_{\rhd\eta}}_x$.*

Because $D_{K(\bullet,z)}P = D_{U(\bullet,z)}P$, this definition is an extension of the one we have been using from the beginning.

11.3. *Reconstruction of a Markov chain from its bridge*

Theorem 11.10. *We have the following identity between measure on $\Omega \times E \cup \mathcal{M}_{\min}$*:

$$\mathbf{E}^P_x\{X \in \mathrm{d}\omega, X_\zeta \in \mathrm{d}\eta\} = \mathbf{E}^{DP}_{x\rhd\eta}\{X \in \mathrm{d}\omega\}\mathbf{E}^P_x\{X_\zeta \in \mathrm{d}\eta\}.$$

Proof. It is sufficient to verify this equation on cylindrical functionals:

$$\begin{aligned}
&\mathbf{E}^P_x\{X_0 = x_0, \ldots, X_n = x_n, X_\zeta \in \mathrm{d}\eta\}\\
&\quad = \mathbf{E}^P_x\bigl\{X_0 = x_0, \ldots, X_n = x_n, (X_\zeta \circ X_{[n,\zeta]}) \in \mathrm{d}\eta\bigr\}\\
&\quad = \mathbf{E}^P_x\{X_0 = x_0, \ldots, X_n = x_n\}\mathbf{E}^P_{x_n}\{X_\zeta \in \mathrm{d}\eta\}\\
&\quad = \mathbf{E}^P_x\{X_0 = x_0, \ldots, X_n = x_n\}K(x_n,\eta)\mathbf{E}^P_o\{X_\zeta \in \mathrm{d}\eta\}\\
&\quad = \mathbf{E}^P_x\{X_0 = x_0, \ldots, X_n = x_n\}\frac{K(x_n,\eta)}{K(x,\eta)}\mathbf{E}^P_x\{X_\zeta \in \mathrm{d}\eta\}\\
&\quad = \mathbf{E}^{DP}_{x\rhd\eta}\{X_0 = x_0, \ldots, X_n = x_n\}\mathbf{E}^P_x\{X_\zeta \in \mathrm{d}\eta\}.
\end{aligned}$$

□

Remark 11.11. *Let $Q = D_h P$ be a sub-stochastic matrix and α be a probability on E. The previous theorem gives*:

$$\mathbf{E}^{D_hP}_\alpha\{X \in \bullet\} = \sum_{x\in E}\int_{\eta\in E\cup\mathcal{M}} \alpha(x)\mathbf{E}^{DP}_{x\rhd\eta}\{X \in \bullet\}\mathbf{E}^{D_hP}_x\{X_\zeta \in \mathrm{d}\eta\}.$$

This equation perfectly illustrates the following fact: *A Markov chain can be parameterized by*:

- *An initial distribution α on E.*
- *A mechanism of evolution given by the class of matrices DP.*
- *A control given by a P-excessive function h.*

11.4. *Confinement*

When η is such that $\mathbf{E}^P_x\{X_\zeta = \eta\} > 0$ then, from Theorem 11.10, we have

$$\mathbf{E}^P_x\{X \in \bullet / X_\zeta = \eta\} = \mathbf{E}^{DP}_{x\rhd\eta}\{X \in \bullet\}$$

and thus the bridge can be seen as a final conditioning exactly as we treated it in Section 5. When $\mathbf{E}_x^P\{X_\zeta = \eta\} = 0$ the right-hand side of the above equation has no sense, but, for convenience, we keep the notation $\xrightarrow{\rhd\eta}$ to designate the "bridgification" i.e. the passage from $\mathbf{E}_x^P$ to $\mathbf{E}_{x\rhd\eta}^{DP}$.

We also use in this section the notations defined in Section 6.1. When ξ is an element of the Martin boundary, we say that $\{x \overset{P}{\rightsquigarrow} \xi\}$ when there exists an infinite sequence $x_0 = x \rightsquigarrow x_1 \rightsquigarrow x_2, \dots$ such that (x_i) belongs to the equivalent class ξ.

Theorem 11.12. *In all this theorem $x \in E$ and $\xi \in \mathcal{M}_{\min}$ are such that $\{x \overset{P_{\mathbb{A}}}{\rightsquigarrow} \xi\}$. We have*:

$$\mathbf{E}_x^P \xrightarrow{X_{[0,S_{\mathbb{A}}]}} \xrightarrow{\rhd\xi} \mathbf{E}_{x\rhd\xi}^{DP_{\mathbb{A}}}.$$

If $\mathbf{E}_x^P\{XX \subset \mathbb{A}\} > 0$ *then we have*:

$$\mathbf{E}_x^P \xrightarrow{/XX\subset\mathbb{A}} \xrightarrow{\rhd\xi} \mathbf{E}_{x\rhd\xi}^{DP_{\mathbb{A}}}.$$

If $\mathbf{E}_{x\rhd\xi}^{DP}\{XX \subset \mathbb{A}\} > 0$ *then we have*:

$$\mathbf{E}_x^P \xrightarrow{\rhd\xi} \xrightarrow{/XX\subset\mathbb{A}} \mathbf{E}_{x\rhd\xi}^{DP_{\mathbb{A}}}.$$

This theorem is the direct extension of Theorem 6.1. The extension of the proof is left to the reader. Actually, we write this extension in order to illustrate a way to condition a Markov chain "to stay inside a set" when it is apparently not possible. Let us treat an example.

We consider an immortal Markov chain with state space $\mathbb{Z}$. We suppose that its transition matrix P checks our hypothesis, and that the Martin boundary of $P_{\mathbb{Z}_+}$ is made by exactly one point that we denote by $+\infty$. We suppose moreover that

$$\forall x\ (\Leftrightarrow \exists x) \quad \mathbf{E}_x^P\{X_\infty = +\infty\} = 0.$$

This last hypothesis implies that the event $\{X \subset \mathbb{Z}_+\} := \{\forall t\colon\ X_t \geq 0\}$ is negligible under $\mathbf{E}_0$ (and under $\mathbf{E}_x$ for all x).

Now, here is a classical question: how is it possible to condition the Markov chain $\mathbf{E}_0$ to stay positive? Answer: this can be done by the two following operations:

$$\mathbf{E}_0^P \xrightarrow{\rhd+\infty} \xrightarrow{/X\subset\mathbb{Z}_+} := \mathbf{E}_+.$$

If intuitively (and not rigorously), you make the two operations commute, then $\mathbf{E}_+$ merits the name of "Markov chain conditioned to stay positive."

Such conditioning often appears in the literature, especially when X is a random walk (see e.g. Bertoin [1], Chapter 7 or Biggins [2]). The recipes to make such conditioning are numerous (you can also condition by events with smaller and smaller probabilities). Here we want to insist on the following fact: If your transition matrix (here P), restricted to the set where you want to lock the Markov chain (here this set is $\mathbb{Z}_+$) has more than one point in its Martin boundary, then such conditioning is ambiguous (you have to make an arbitrary choice for the final destination of your process).

12. Un-normalized bridges

Data 12.1. *Exceptionally in this section, P is any non-negative matrix.*

When P is sub-stochastic, the existence of a recurrent point (i.e. $\mathbf{E}_x^P\{\exists t \geq 1\colon\ X_t = x\} = 1$) implies non-transience and forbids the construction of $\mathbf{E}_{\bullet\rhd\bullet}^{DP}$. In this section, we define σ-finite measures $\mathbf{F}_{x\rhd z}^P$ which, in the transient case are $U_{[P]}(x,z)\mathbf{E}_{x\rhd z}^{DP}$ but which can be defined for all non-negative P with finite spectral radius.

12.1. *Spectral radius*

We recall some facts about spectral radii. These facts are quite well known, see Seneta [29], Chapter 6 or Woess [35], Chapter 2. Let us write:

$$\Lambda_{[P]} = \left\{\lambda > 0\colon\ \frac{1}{\lambda} P \text{ is transient}\right\},$$

it is clearly an interval of type $]\rho, \infty[$ or $[\rho, \infty[$ for some $\rho \in \mathbb{R}_+ \cup \{+\infty\}$. This ρ, also noted $\rho_{[P]}$, is called the spectral radius of P. Proposition 2.4 indicates that $\Lambda_{[P]} = \Lambda_{[P^\top]}$. Proposition 2.7 indicates that $\Lambda_{[P]} = \{\lambda > 0\colon \exists h > 0$ which is $\frac{1}{\lambda}P$-potential$\}$. So we are tempted to compare $\Lambda_{[P]}$ with:

$$\Lambda'_{[P]} := \{\lambda \geq 0\colon\ \exists h > 0\colon\ Ph \leq \lambda h\}.$$

Lemma 12.2. *We have $\Lambda'_{[P]} = \Lambda_{[P]}$ or $\Lambda'_{[P]} = \Lambda_{[P]} \cup \{\rho_{[P]}\}$.*

Proof. Λ' is clearly greater than Λ. So it is sufficient to show that the interior of Λ' is included in the interior of Λ. Take $\lambda \in \Lambda'$ and h a $\frac{1}{\lambda}P$-excessive function. Take $q \in]0, 1[$. The matrix $Q := \frac{q}{\lambda} D_h P$ is sub-stochastic and satisfies $U_{[Q]} 1_E \leq \frac{1}{1-q} 1_E$. So $\frac{\lambda}{q} \in \Lambda$; and this is true for all $q \in]0, 1[$. □

Next we give an example of a sub-stochastic matrix P such that $\Lambda_{[P]} \neq \Lambda'_{[P]}$ and $\Lambda'_{[P]} \neq \Lambda'_{[P^\top]}$. The matrix P and its diagonalization are given by:

$$P = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}.$$

Here is its potential matrix:

$$U_{[P]} = (I - P)^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \infty & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \infty & 0 \\ \infty & 2 \end{bmatrix}.$$

So we get $\Lambda_{[P]} = \Lambda_{[P^\top]} =]1, \infty[$ while $\Lambda'_{[P]} = [1, \infty[$ (because $P1_E \leq 1_E$). As it is impossible to solve $P^\top h \leq h$ with $h > 0$, we deduce that $\Lambda'_{[P^\top]} =]1, \infty[$.

The reader is also invited to check that the matrix P defined on $E = \mathbb{N}$ by $P(n, m) = 1_{\{m=n+1\}}$ has a spectral radius $\rho = 0$, while the matrix $\forall m, n\ P(n, m) = 1$ has a spectral radius $\rho = +\infty$.

Of course, we are more accustomed to the irreducible case where the situation is easier to handle. The argument in Woess [35], p. 81, gives:

Proposition 12.3. *Suppose that P is irreducible and $\rho_{[P]} < \infty$. Then we have $\Lambda'_{[P]} = [\rho_{[P]}, \infty[$.*

Proof. Denote by $\mathcal{H}_\lambda$ the set of $\frac{1}{\lambda}P$-excessive functions normalized by $h(o) = 1$ for some reference point $o \in E$. Take λ such that $\mathcal{H}_\lambda \neq \varnothing$. By Fatou lemma $\mathcal{H}_\lambda$ is closed. By irreducibility, for all x there exists n_x such that $P^{n_x}(o, x) > 0$. Thus, for $h \in \mathcal{H}_\lambda$:

$$P^{n_x}(o, x) h(x) \leq P^{n_x} h(o) \leq \lambda^{n_x} h(o).$$

Thus every function h of $\mathcal{H}_\lambda$ is bounded by the same function $x \mapsto \frac{\lambda^{n_x} f(o)}{P^{n_x}(o,x)}$, so $\mathcal{H}_\lambda$ is compact. Finally, we see that $\bigcap_{\lambda > \rho} \mathcal{H}_\lambda = \mathcal{H}_\rho \neq \varnothing$. □

To finish with these reminders about spectral radii, we explain the link with eigenvalues. Let us write $\tilde{\Lambda}_{[P]} = \{\lambda\colon\ \exists h > 0\colon\ Ph = \lambda h\}$. Suppose P irreducible and E finite. Then the Perron–Frobenius theorem (see Seneta [29], p. 1) indicates that $\tilde{\Lambda}_{[P]} = \{\rho_{[P]}\}$. The absolute value of other eigenvalues are strictly less than $\rho_{[P]}$ and $\rho_{[P]}$ is the only one to have an eigenvector everywhere positive. On the other hand, suppose P irreducible and E infinite, the Martin boundary (see Section 11) theory indicates that $\tilde{\Lambda}_{[P]} = [\rho_{[P]}, \infty[$. For a short direct proof of this fact, see Woess [35], Lemma 7.6, p. 83).

12.2. *Construction of un-normalized bridges*

Data 12.4. *From now on, and until the end of this section, we suppose that* $\rho_{[P]} < \infty$.

Lemma 12.5. *For* $i = 1, 2$, *pick a real* $q_i > 0$ *and a function* $h_i > 0$ *such as* $Q_i := q_i D_{h_i} P$ *is sub-stochastic. We have*:

$$\frac{h_1(x)}{h_1(z)} \mathbf{E}_x^{Q_1}\left[\sum_t q_1^{-t} \mathfrak{f}(X_{[0,t]}) 1_{\{X_t = z\}}\right] = \frac{h_2(x)}{h_2(z)} \mathbf{E}_x^{Q_2}\left[\sum_t q_2^{-t} \mathfrak{f}(X_{[0,t]}) 1_{\{X_t = z\}}\right].$$

Proof. It is sufficient to establish this formula, replacing $\mathfrak{f}$ by $1_{\{X_0 = x_0, \dots, x_n = x_n\}}$. Doing this and computing, we see that all h_i, q_i disappear. □

Definition 12.6. *Pick real* $q > 0$ *and a function* $h > 0$ *such that* $q D_h P$ *is sub-stochastic. We define the kernel*

$$\mathbf{F}^P_{x \rhd z}\big[\mathfrak{f}(X)\big] = \frac{h(x)}{h(y)} \mathbf{E}_x^{q D_h P}\left[\sum_t q^{-t} \mathfrak{f}(X_{[0,t]}) 1_{\{X_t = z\}}\right]$$

which is called the un-normalized Markov-bridge-kernel. From the previous lemma, this definition does not depend on the choice of (q, h).

We see immediately that $\mathbf{F}^P_{x \rhd z}$ is supported by $\{\zeta < \infty\}$ and that its total mass is $U_{[P]}(x, z)$ which can be infinite. We give now some direct consequences of Definition 12.6 and Proposition 12.5:

- When P is sub-stochastic:

$$\mathbf{F}^P_{x \rhd z} = \mathbf{E}^P_x\left[\sum_t 1_{\{X_{[0,t]} \in \bullet\}} 1_{\{X_t = z\}}\right]. \tag{22}$$

- When P is transient, we have:

$$\mathbf{F}^P_{x \rhd z} = \mathbf{E}^{DP}_{x \rhd z}\{X \in \bullet\} U(x, z). \tag{23}$$

- For $q \in \mathbb{R}$, we have:

$$\mathbf{F}^{qP}_{x \rhd z} = \mathbf{F}^P_{x \rhd z}\big[1_{\{X \in \bullet\}} q^{\zeta}\big]. \tag{24}$$

Remark 12.7.

- *From* (24), *taking* q *sufficiently small, we see that for any* P, $\forall s$ $\mathbf{F}^P_{x \rhd z}\{\zeta = s\} < \infty$.
- *For* P *sub-stochastic, from* (22) *we can compute* $\mathbf{F}^P_{x \rhd z}\{X \in \bullet / \zeta = s\} = \mathbf{E}^P_x\{X \in \bullet / X_s = z\}$ *which is the same formula that the one with the normalized bridge* (*conditionning make the normalization*). *We can also rewrite this as*: $\mathbf{F}^P_{x \rhd z}\{X \in \bullet, \zeta = s\} = \mathbf{E}^P_x\{X \in \bullet / X_s = z\} P^s(x, z)$ *and compare this with the formula given by Fourati* [11], *Proposition* 5.1.

12.3. *Past-future extraction and time reversal*

Theorem 12.8. *We have*:

$$\mathbf{F}^P_{x \rhd y}\big[\mathfrak{f}(X)\big] \mathbf{F}^P_{y \rhd z}\big[\mathfrak{g}(X)\big] = \mathbf{F}^P_{x \rhd z}\left[\sum_t \mathfrak{f}(X_{[0,t]}) 1_{\{X_t = y\}} \mathfrak{g}(X_{[t,\zeta]})\right] \tag{25}$$

for all states $x, y, z \in E$, *all test functions* $\mathfrak{f}, \mathfrak{g} : \Omega \mapsto \mathbb{R}_+$.

Proof. First suppose that P is transient. Formula (25) is simply the combination of (23) and of the past–future extraction (Theorem 3.5) wrote as follows:

$$\mathbf{E}^{DP}_{x\rhd y}\big[\mathfrak{f}(X)\big]U(x,y)\mathbf{E}^{DP}_{y\rhd z}\big[\mathfrak{g}(X)\big]U(y,z)=\mathbf{E}^{DP}_{x\rhd z}\Bigg[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Bigg]U(x,z).$$

Now suppose P has just a finite spectral radius. Take $q>0$ such that qP is transient. Applying (25) to qP and using (24) we get:

$$\begin{aligned}\mathbf{F}^{P}_{x\rhd y}\big[\mathfrak{f}(X)q^{\zeta}\big]\mathbf{F}_{y\rhd z}\big[\mathfrak{g}(X)q^{\zeta}\big]&=\mathbf{F}^{P}_{x\rhd z}\Bigg[\sum_t \mathfrak{f}(X_{[0,t]})q^t 1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})q^{\zeta-t}\Bigg]\\&=\mathbf{F}^{P}_{x\rhd z}\Bigg[\sum_t \big(\mathfrak{f}q^{\zeta}\big)\circ X_{[0,t]}1_{\{X_t=y\}}\big(\mathfrak{g}q^{\zeta}\big)\circ X_{[t,\zeta]}\Bigg].\end{aligned}$$

The result follows by substituating $\mathfrak{f}:=\mathfrak{f}q^{-\zeta}$ and $\mathfrak{g}:=\mathfrak{g}q^{-\zeta}$. □

Proposition 12.9. *When $x\rightsquigarrow z$ we have $\mathbf{F}^{P}_{x\rhd z}\{X\in\bullet\}=\mathbf{F}^{P^\top}_{z\rhd x}\{X_{\overleftarrow{[0,\zeta]}}\in\bullet\}$.*

Proof. Suppose P transient. From Theorem 4.3: $\mathbf{E}^{DP}_{x\rhd z}=\mathbf{E}^{DP^\top}_{z\rhd x}\{X_{\overleftarrow{[0,\zeta]}}\in\bullet\}$. Multiplying at the left by $U_{[P]}(x,z)$ and at the right by the same quantity written $U_{[P^\top]}(z,x)$, we obtain Proposition 12.9. The non-transient case follows by working with $P:=qP$, as in the previous proof. □

12.4. *Application*

Let us explain briefly how the last proposition can be used to prove the following classical result:

Proposition 12.10. *Suppose P sub-stochastic. Let $\mu>0$ a P-excessive measure. Denote by $\widehat{P}=D_\mu P^\top$. We have:*

$$\mu(a)\mathbf{E}^P_a\Bigg[\sum_{t<T^*_A}1_{\{X_t=y\}}\Bigg]1_{\{a\in A\}}=\mu(y)\mathbf{E}^{\widehat{P}}_y\{X_{T_A}=a\},$$

where $T_A=\inf\{t\ge 0\colon X_t\in A\}$ and $T^*_A=\inf\{t\ge 1\colon X_t\in A\}$.

Remark 12.11. *Suppose that $\mathbf{E}^{\widehat{P}}_y\{T_x<\infty\}=1$. The above equation, applied with $A=\{x\}$ becomes $\mu(x)\mathbf{E}^P_x[\sum_{t<T^*_x}1_{\{X_n=y\}}]=\mu(y)$, which is a very famous formula of the recurrent theory of Markov chains.*

Proof of Proposition 12.10. Pick $a\in A$, $y\in E$. Having a look at Fig. 2 we see that:

$$\begin{aligned}1_{\{X_0=a\}}1_{\{t<T^*_A\}}1_{\{X_t=y\}}&=1_{\{X_0=a\}}1_{\{T_A=t\}}\circ X_{\overleftarrow{[0,t]}}1_{\{X_t=y\}}\\&=1_{\{X_0=a\}}1_{\{T_A=\zeta\}}\circ X_{\overleftarrow{[0,\zeta]}}\circ X_{[0,t]}1_{\{X_t=y\}}.\end{aligned}$$

Summing over all t, integrating, and using (22) we get:

$$\mathbf{E}^P_a\Bigg[\sum_{t<T^*_A}1_{\{X_t=y\}}\Bigg]=\mathbf{F}^P_{a\rhd y}[1_{\{T_A=\zeta\}}\circ X_{\overleftarrow{[0,\zeta]}}].$$

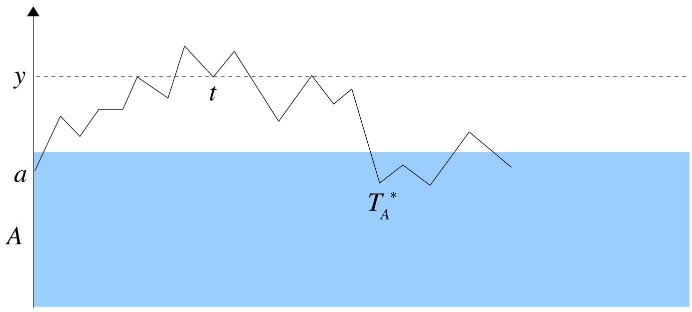

Fig. 2. This figure illustrates the trajectorial identity used in the proof of Proposition 12.10.

Multiplying by $\frac{\mu(a)}{\mu(y)}$ and using the time reversal Proposition 12.9, we get:

$$\begin{aligned}\mathbf{E}_a^P\left[\sum_{t<T_A^*} 1_{\{X_t=y\}}\right] &= \frac{\mu(a)}{\mu(y)}\mathbf{F}_{y\rhd a}^{P^\top}[1_{\{T_A=\zeta\}}] \\ &= \mathbf{F}_{y\rhd a}^{\widehat{P}}[1_{\{T_A=\zeta\}}].\end{aligned}$$

The proof is conclude by the next lemma (also interesting for itself). □

Lemma 12.12. *For any stopping time T and sub-stochastic matrix Q we have $\mathbf{F}_{x\rhd z}^Q\{T=\zeta\} = \mathbf{E}_x^Q\{X_T=z\}$.*

Proof. $\mathbf{F}_{x\rhd z}^Q\{T=\zeta\} = \mathbf{E}_x^Q[\sum_t 1_{\{T=\zeta\}} \circ X_{[0,t]} 1_{\{X_t=z\}}] = \mathbf{E}_x^Q[\sum_t 1_{\{T=t\}} 1_{\{X_t=z\}}]$. □

Comments. *Let us compare the advantages of normalized and un-normalized bridges. Normalized bridges have the advantage of depending only on the D-class and of being a probability which aids intuition. Un-normalized bridges cover the recurrent cases. Un-normalized bridges are also the simplest "path integral" we can imagine: let us suppose P sub-stochastic and compute*

$$\begin{aligned}\mathbf{F}_{x\rhd z}^P\{X=a_0a_1\cdots a_n\dagger\cdots\} &= \mathbf{E}_{x\rhd z}^P\left[\sum_t 1_{\{X_{[0,t]}=a_0a_1\cdots a_n\dagger\cdots\}} 1_{\{X_t=z\}}\right] \\ &= \mathbf{E}_{x\rhd z}^P\{X_0=a_0, X_1=a_1,\dots,X_n=a_n=z\} \\ &= I(x,a_0)P(a_0,a_1)\cdots P(a_{n-1},a_n)I(a_n,z).\end{aligned}$$

Such a formula can be easily extended to every P with $\rho_{[P]}<\infty$ (replacing P by qD_hP etc.). So an alternative definition for the un-normlalized bridge is simply the measure supported by $\{\zeta<\infty\}$ which weighs every singleton $\omega=a_0\cdots a_n\dagger\dagger\cdots$ by

$$\mathbf{F}_{x\rhd z}^P\{\omega\} = I(x,a_0)P(a_0,a_1)\cdots P(a_{n-1},a_n)I(a_n,z).$$

From this definition it is quite easy to establish the past–future extraction. The time reversal property is immediate. The only little difficulty (when E is infinite) is to prove that $\forall t\ \mathbf{F}_{x\rhd z}^P\{\zeta=t\}<\infty$ and to do this, the best way is to come back to sub-stochastic matrices by D-transformation.

13. When two bridges partially coincide

Data 13.1. *Throughout this section, we fix P, P' two transient matrices.*

Notations. *For $K \subset E$ we write $P_K(x,y)$ for $P(x,y)1_{\{x\in K\}}1_{\{y\in K\}}$ and DP_K for $D(P_K)$. Thus $DP_K = DP'_K$ if and only if there exists $h>0$ such that $\forall x,y \in K$ $P(x,y) = \frac{h(y)}{h(x)}P'(x,y)$. In this situation we say that DP, DP' coincide on K.*

For $\mathbb{K} \subset E \times E$ we write $P_{\mathbb{K}}(x,y)$ for $P(x,y)1_{\{xy\in\mathbb{K}\}}$ etc.

13.1. *Closed set*

A set $K \subset E$ is said P-closed when $K = \{K \overset{P}{\rightsquigarrow} \bullet \overset{P}{\rightsquigarrow} K\}$. This means that every chain beginning and ending in K has all its elements in K. It is easy to see that any set of type $\{A \overset{P}{\rightsquigarrow} \bullet \overset{P}{\rightsquigarrow} C\}$ is P-closed. In particular, P-absorbing and $P^\top$-absorbing sets are P-closed.

The main interest of a P-closed set K is that:

$$\forall x,z \in K \quad \mathbf{E}^{DP}_{x\rhd z}\{X \in \bullet\} = \mathbf{E}^{DP}_{x\rhd z}\{X \in \bullet / X \subset K\} = \mathbf{E}^{DP_K}_{x\rhd z}\{X \in \bullet\} \tag{26}$$

(the last equality comes from Theorem 6.1). As a consquence, when K is P-closed and P'-closed then:

$$DP_K = DP'_K \quad \Rightarrow \quad \forall x,z \in K \quad \mathbf{E}^{DP}_{x\rhd z} = \mathbf{E}^{DP'}_{x\rhd z}.$$

The purpose of the rest of this section is to establish some converse of this implication.

13.2. *Simple coincidence*

Proposition 13.2. *Fix $x,z \in E$. The following points are equivalents*:

1. $\mathbf{E}^{DP}_{x\rhd z} = \mathbf{E}^{DP'}_{x\rhd z}$.
2. *We have $\{x \overset{P}{\rightsquigarrow} \bullet \overset{P}{\rightsquigarrow} z\} = \{x \overset{P'}{\rightsquigarrow} \bullet \overset{P'}{\rightsquigarrow} z\}$ and DP, DP' coincide on this set.*

Proof. $1 \Rightarrow 2$. We have:

$$\begin{aligned}\{b : x \overset{P}{\rightsquigarrow} b \overset{P}{\rightsquigarrow} z\} &= \bigl\{b\colon \mathbf{E}^{DP}_{x\rhd z}\{\exists t\colon X_t = b\} > 0\bigr\} \\ &= \bigl\{b\colon \mathbf{E}^{DP'}_{x\rhd z}\{\exists t\colon X_t = b\} > 0\bigr\} = \{b : x \overset{P'}{\rightsquigarrow} b \overset{P'}{\rightsquigarrow} z\}\end{aligned}$$

this set will be simply denoted by $\{x \rightsquigarrow \bullet \rightsquigarrow z\}$. If this set is reduced to $\{x,z\}$, then $x \overset{P}{\not\rightsquigarrow} z$ and $x \overset{P'}{\not\rightsquigarrow} z$ and in this case the second point is obviously true. Let us now assume that $x \overset{P}{\rightsquigarrow} z$. We fix $a,b \in \{x \rightsquigarrow \bullet \rightsquigarrow z\}$. We have

$$\begin{aligned}P_{\rhd z}(a,b) &= \mathbf{E}^{DP}_{x\rhd z}\{X_{T_a+1} = b / T_a < \infty\} \\ &= \mathbf{E}^{DP'}_{x\rhd z}\{X_{T_a+1} = b / T_a < \infty\} = P'_{\rhd z}(a,b),\end{aligned}$$

this equality means: $\frac{U(b,z)}{U(a,z)}P(a,b) = \frac{U'(b,z)}{U'(a,z)}P'(a,b)$, so we have

$$P'(a,b) = \frac{h(b)}{h(a)}P(a,b) \quad \text{with } h(\bullet) = \frac{U(\bullet,z)}{U'(\bullet,z)}$$

which said that DP and DP' coincide on $\{x \rightsquigarrow \bullet \rightsquigarrow z\}$.

$2 \Rightarrow 1$. This comes from the fact that $\{x \overset{P}{\rightsquigarrow} \bullet \overset{P}{\rightsquigarrow} z\} = \{x \overset{P'}{\rightsquigarrow} \bullet \overset{P'}{\rightsquigarrow} z\}$ is a P and P' closed set (see previous subsection). □

The previous proposition cannot be generalized replacing $\{x \overset{P}{\rightsquigarrow} \bullet \overset{P}{\rightsquigarrow} z\}$ by any closed set. To understood the problem, we need to make some specific graph theory.

13.3. *Directed-spanning-tree*

An directed tree is a directed graph $(B, \xrightarrow{\Upsilon})$ such that its undirected version has no cycle, and such that there exists one element o called "center" such that for every $b \in B$ we have $o \overset{\Upsilon}{\rightsquigarrow} b$ or $b \overset{\Upsilon}{\rightsquigarrow} o$. The degenerated graph with one vertex, zero edge is a directed tree.

We say that a graph $(B, \rightarrow)$ *admits a directed-spanning-trees* when each of its connected component C admits a sub-graph which is a directed-tree linking every vertex of C.

If there exists $x \in B$ such that $B = \{x \rightsquigarrow \bullet\}$ then $(B, \rightarrow)$ admits a directed-spanning-tree (to see this, explore B from x, replacing cycles by branchings). By symmetry: if there exists $z \in B$ such that $B = \{\bullet \rightsquigarrow z\}$ then $(B, \rightarrow)$ admits a directed-spanning-tree. In particular, if P is irreducible then its directed graph admits a directed-spanning-tree.

Lemma 13.3. *Suppose that $(E, \xrightarrow{P})$ or $(E, \xrightarrow{P'})$ admits a directed-spanning-tree. We have $DP = DP'$ if and only if, for all $x, z \in E$, we have $\{x \overset{P}{\rightsquigarrow} \bullet \overset{P}{\rightsquigarrow} z\} = \{x \overset{P'}{\rightsquigarrow} \bullet \overset{P'}{\rightsquigarrow} z\}$ and DP, DP' coincide on this set.*

Proof. Direct sense is obvious. Let us prove the converse. The equality $\forall x, z$ $\{x \overset{P}{\rightsquigarrow} \bullet \overset{P}{\rightsquigarrow} z\} = \{x \overset{P'}{\rightsquigarrow} \bullet \overset{P'}{\rightsquigarrow} z\}$ implies $(E, \xrightarrow{P}) = (E, \xrightarrow{P'})$. This graph is now denoted by $(E, \rightarrow)$. Moreover, without loss of generality, we can assume that this graph has only one connected component (if not, we solve the problem on each component separately).

We denote by $\gamma(a, b) = \frac{P(a,b)}{P'(a,b)}$ with $\frac{0}{0} = 0$. The hypothesis indicates that:

$$\forall x, z \in E\ \exists h_{x,z} > 0\colon\ \forall a, b \in \{x \rightsquigarrow \bullet \rightsquigarrow z\}\colon \qquad \gamma(a, b) = \frac{h_{xz}(b)}{h_{xz}(a)} 1_{\{a \rightsquigarrow b\}}$$

(our aim is to find a function h not depending on x, z). This implies in particular that:

$$x \rightsquigarrow y \rightsquigarrow z \quad \Rightarrow \quad \gamma(x, z) = \gamma(x, y)\gamma(y, z). \tag{27}$$

Let us consider a directed spanning tree $(E, \xrightarrow{\Upsilon})$ with central point o. Let $x \in E$ and let $x = a_0, a_1, a_2, \ldots, a_{n-1}, a_n = z$ be the unique chain going from o to x or from x to o in the tree. We define a function $h > 0$ as follows:

$$\begin{aligned} & h(o) = 1, \\ \text{when } o \overset{\Upsilon}{\rightsquigarrow} x, \quad & h(x) = \gamma(o, a_1)\gamma(a_1, a_2) \cdots \gamma(a_{n-1}, x), \\ \text{when } x \overset{\Upsilon}{\rightsquigarrow} o, \quad & h(x) = \frac{1}{\gamma(o, a_1)} \frac{1}{\gamma(a_1, a_2)} \cdots \frac{1}{\gamma(a_{n-1}, x)}. \end{aligned}$$

Therefore, by construction:

$$a \overset{\Upsilon}{\rightsquigarrow} o \quad \Rightarrow \quad \gamma(a, o) = \frac{h(o)}{h(a)} \quad \text{and} \quad o \overset{\Upsilon}{\rightsquigarrow} a \quad \Rightarrow \quad \gamma(o, a) = \frac{h(a)}{h(o)}. \tag{28}$$

Let us verify that:

$$\forall a, b \in E \quad \gamma(a, b) = \frac{h(b)}{h(a)} 1_{\{a \rightsquigarrow b\}}. \tag{29}$$

If $a \not\rightsquigarrow b$ then this equation become $0 = 0$. We suppose now that $a \rightsquigarrow b$.

- If $a \overset{\Upsilon}{\rightsquigarrow} o$ and $b \overset{\Upsilon}{\rightsquigarrow} o$ then we have $a \rightsquigarrow b \rightsquigarrow o$ and (27) implies $\gamma(a, b)\gamma(b, o) = \gamma(a, o)$. Using (28) we get (29).
- If $a \overset{\Upsilon}{\rightsquigarrow} o \overset{\Upsilon}{\rightsquigarrow} b$ then (28) gives directly (29).
- If $o \overset{\Upsilon}{\rightsquigarrow} a$ and $o \overset{\Upsilon}{\rightsquigarrow} b$ then we have $o \rightsquigarrow a \rightsquigarrow b$ and (27) implies $\gamma(o, b) = \gamma(a, o)\gamma(a, b)$. Using (28) we get (29).
- If $o \overset{\Upsilon}{\rightsquigarrow} a$ and $b \overset{\Upsilon}{\rightsquigarrow} o$ then we have $a \rightsquigarrow b \rightsquigarrow o$ and (27) implies $\gamma(a, o) = \gamma(a, b)\gamma(b, o)$. Using (28) we get (29). $\square$

13.4. *Multiple coincidence and counter-example*

Theorem 13.4. *Consider K be a P-closed and P'-closed set. Suppose moreover that $(K, \xrightarrow{P})$ or $(K, \xrightarrow{P'})$ admits a directed-spanning-tree. $\mathbf{E}^{DP}_{\bullet\triangleright\bullet}$, $\mathbf{E}^{DP'}_{\bullet\triangleright\bullet}$ coincide on K if and only if DP, DP' coincide on K.*

Proof. Direct sense is the mix of Proposition 13.2 and Lemma 13.3 (applied with $E := K$). Converse is the direct consequence of (26) and does not require existence of a directed spanning tree. □

Counter-example 1. *See Fig. 3.*

13.5. *Application to inhomogeneous bridge*

Pick $x, z \in E$ and $t \in \mathbb{N}$. We denote by $\mathbb{K}_{[P]}(x \overset{t}{\triangleright} z) \subset E \times E$ the set of pairs (a, b) such that exists a path with exactly t arrows: $x = c_0 \xrightarrow{P} c_1 \xrightarrow{P} \cdots \xrightarrow{P} c_{t-1} \xrightarrow{P} c_t = z$ such that $a = c_i$, $b = c_{i+1}$ for some i. Fig. 4 helps to understand such a set.

Proposition 13.5. *Suppose that P, P' are sub-stochastic. The following points are equivalent*:

- $\mathbf{E}^{P}_{x}\{X_{[0,t]}/X_t = z\} = \mathbf{E}^{P'}_{x}\{X_{[0,t]}/X_t = z\}$.
- $\mathbb{K}_{[P]}(x \overset{t}{\triangleright} z) = \mathbb{K}_{[P']}(x \overset{t}{\triangleright} z)$ *and DP, DP' coincide on this set.*

Proof. To come back on our homogeneous situation, we use the classical technique of the passage to the space–time: On $\mathbb{N}$ we define the matrix $J(s,t) = 1_{\{t=s+1\}}$. On $E \times \mathbb{N}$, we consider the matrix $(P \otimes J)(xs, yt) = P(x,y)J(s,t)$.

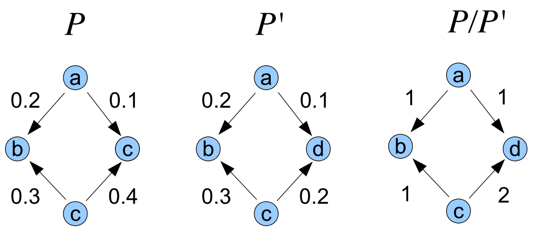

Fig. 3. For the two matrices above, it is impossible to find a function $h > 0$ such that $\forall x, y$ $P(x,y) = \frac{h(y)}{h(x)} P'(x,y)$. Meanwhile, we have $\forall x, z$: $\mathbf{E}^{DP}_{x\triangleright z} = \mathbf{E}^{DP'}_{x\triangleright z}$.

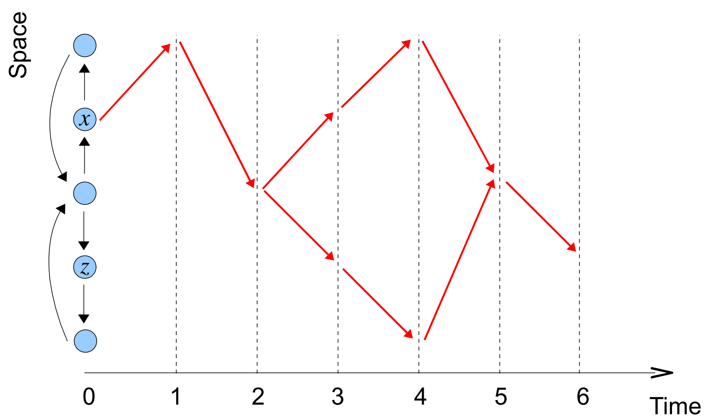

Fig. 4. On the space axis, we have represented the directed graph of a matrix P. Then, we draw all trajectories linking x to z in exactly 6 steps. This allows to find all the pairs (a, b) in $\mathbb{K}_{[P]}(x \overset{6}{\triangleright} z)$.

The canonical process taking values in $E \times \mathbb{N}$ is denoted by $(\mathcal{X}, \mathcal{T})$. The last passage of this process to zt is denoted by τ_{zt}. We have:

$$\mathbf{E}^{P\otimes J}_{x0\triangleright zt}\{\mathcal{X} \in \bullet\} = \mathbf{E}^{P\otimes J}_{x0}\{\mathcal{X}_{[0,\tau_{zt}]} \in \bullet / \tau_{zt} \in [0, \infty[\} = \mathbf{E}^P_x\{X_{[0,t]} \in \bullet / X_t = z\}.$$

Translating Proposition 13.2 on this case gives the equivalence. □

For $t \in \mathbb{N}$, we say that P is t-primitive when $\forall x, y\ P^t(x, y) > 0$.

Corollary 13.6. *Suppose that P, P' are sub-stochastic and t-primitive. The following points are equivalents*:

- $\mathbf{E}^P_x\{X_{[0,2t+1]}/X_{2t+1} = z\} = \mathbf{E}^{P'}_x\{X_{[0,2t+1]}/X_{2t+1} = z\}$.
- $DP = DP'$.

Proof. Let us look at the consequence of primitivity: Let $a \overset{P}{\to} b$. There exists a path of length t joining x to a, a path of length 1 joining a to b, a path of length t joining b to z, so that $(a, b) \in \mathbb{K}_{[P]}(x \overset{2t+1}{\triangleright} z)$. Consequently, the restriction of DP to $\mathbb{K}_{[P]}(x \overset{2t+1}{\triangleright} z)$ is DP itself. The same fact is true for P'. Now it appears clearly that this corollary is the consequence of the previous proposition applied with $t := 2t + 1$. □

Comments. *This last corollary is the discrete version of a result by Fitzsimmons [9] who explains the coincidence of inhomogeneous bridges in term of the existence of an invariant function. To see clearly the link, just remark that $DP = DP' \Leftrightarrow \exists h > 0$: $P(x, y)h(y) = P'(x, y)h(x)$. Summing on y we see that h is P-excessive. If moreover P' is stochastic, then h is P-invariant.*

14. All axiomatic bridges can be constructed

Data 14.1. *Throughout this section, we fix $\mathbf{E}_{\bullet\triangleright\bullet}$ a Markov-bridge-kernel as defined in the axiomatic Definition 1.5 (also rewrite below).*

14.1. *Purpose*

To $\mathbf{E}_{\bullet\triangleright\bullet}$, we associate a graph $(E, \overset{\triangleright}{\to})$ defined by $a \overset{\triangleright}{\to} b$ when $\mathbf{E}_{a\triangleright b}\{X_1 = b\} > 0$.

Theorem 14.2. *There exists a D-class DP such that $\mathbf{E}_{\bullet\triangleright\bullet} = \mathbf{E}^{DP}_{\bullet\triangleright\bullet}$. The directed graph of DP is given by $(E, \overset{\triangleright}{\to})$. When this directed graph admits a directed-spanning-tree, the mentioned D-class is unique.*

The proof will be perform at the end of this section.

To avoid multiple page turning, we repeat here the four axioms defining a Markov-bridge-kernel:

- "Degeneracy": If $\mathbf{E}_{x\triangleright z}\{\zeta = 0\} = 1$ then $\mathbf{E}_{x\triangleright z}$ is the Dirac measure on the trajectory $x\dagger\dagger\cdots$. In this case we say that $\mathbf{E}_{x\triangleright z}$ is degenerated.
- "Support": If $\mathbf{E}_{x\triangleright z}$ is non-degenerated then $\mathbf{E}_{x\triangleright z}\{X_0 = x, X_\zeta = z\} = 1$.
- "Cohesion": If $\mathbf{E}_{x\triangleright y}, \mathbf{E}_{y\triangleright z}$ are non-degenerated then $\mathbf{E}_{x\triangleright z}\{\exists t\colon X_t = y\} > 0$.
- "Past-future extraction": We have

$$\mathbf{E}_{x\triangleright z}\Big[\sum_t \mathfrak{f}(X_{[0,t]})1_{\{X_t=y\}}\mathfrak{g}(X_{[t,\zeta]})\Big] = \mathbf{E}_{x\triangleright y}\big[\mathfrak{f}(X)\big]\mathbf{E}_{x\triangleright z}\Big[\sum_t 1_{\{X_t=y\}}\Big]\mathbf{E}_{y\triangleright z}\big[\mathfrak{g}(X)\big].$$

Remark 14.3. *The less natural axiom is perhaps "cohesion." Let us make a construction showing that this axiom cannot be deduced from the three others: Take two parts A_1, A_2 of E such that $A_1 \cap A_2 = \{z\}$. Take two transient matrices P_1, P_2 supported by A_1, A_2 and irreducible on their support. Define $\mathbf{E}_{\bullet\triangleright\bullet}$ by: if x, z belong to the same A_i then $\mathbf{E}_{x\triangleright z} = \mathbf{E}^{P_i}_{x\triangleright z}$, if not $\mathbf{E}_{x\triangleright z}$ is taken degenerated. We see that this $\mathbf{E}_{\bullet\triangleright\bullet}$ satisfies "degeneracy," "support," "past–future extraction" but not "cohesion."*

14.2. *Lemmas*

Corollary 1.9 indicates that $\mathbf{E}_{\bullet\triangleright z}$ is a mortal Markov-chain-kernel. We denote by $\mathbf{P}_{\triangleright z}$ its transition matrix.

Lemma 14.4. *We have*

$$\mathbf{E}_{x\triangleright z}\{\exists t\colon\ X_t = y\} > 0 \quad\Rightarrow\quad \mathbf{E}_{x\triangleright z}\big\{X_{[0,\tau_y]} \in \bullet/\tau_y \in [0,\infty[\big\} = \mathbf{E}_{x\triangleright y}. \tag{30}$$

Proof. We already proved this for constructed bridges $\mathbf{E}^{P}_{\bullet\triangleright\bullet}$ (Theorem 5.1, first item). This proof only required the past–future extraction which is also true for our axiomatic bridge $\mathbf{E}_{\bullet\triangleright\bullet}$. $\square$

For a graph $(E,\to)$ we write $x \ast\!\!\rightsquigarrow y$ to indicate that there exists a chain $x \to \cdots \to y$ and that this chain is not degenerated (has at least one arrow). For example, for the directed graph of a sub-stochastic matrix Q, we have $x \overset{Q}{\ast\!\!\rightsquigarrow} y \Leftrightarrow \mathbf{E}^{Q}_{x}\{\exists t \geq 1\colon\ X_t = y\} > 0$, while $x \overset{Q}{\rightsquigarrow} y \Leftrightarrow \mathbf{E}^{Q}_{x}\{\exists t \geq 0\colon\ X_t = y\} > 0$.

Lemma 14.5. *We have $x \overset{\triangleright}{\ast\!\!\rightsquigarrow} y$ if and only if $\mathbf{E}_{x\triangleright y}$ is not degenerated.*

Proof. Suppose $x \overset{\triangleright}{\ast\!\!\rightsquigarrow} y$. Suppose that the linking chain has $n \geq 1$ arrows: $x \overset{\triangleright}{\to} a_1 \overset{\triangleright}{\to} \cdots \overset{\triangleright}{\to} a_n = y$. This force $\mathbf{E}_{x\triangleright a_1}$ and $\mathbf{E}_{a_1\triangleright a_2}$ to be non-degenerated. From the "cohesion" axiom, $\mathbf{E}_{x\triangleright a_2}$ is non-degenerated. By induction we deduce that $\mathbf{E}_{x\triangleright y}$ is non-degenerated.

Conversely, suppose $\mathbf{E}_{x\triangleright y}$ non-degenerated i.e. $\mathbf{E}_{x\triangleright y}\{\zeta \geq 1\} > 0$. From "support axiom" $\mathbf{E}_{x\triangleright y}\{\exists t \geq 1\colon\ X_t = y\} > 0$ and so there exists a chain with $n \geq 1$ arrows $x \overset{\mathbf{P}_{\triangleright y}}{\to} a_1 \overset{\mathbf{P}_{\triangleright y}}{\to} \cdots \overset{\mathbf{P}_{\triangleright y}}{\to} a_n = y$. From (30):

$$\begin{aligned}\mathbf{E}_{x\triangleright a_1}\{X_1 = a_1\} &= \mathbf{E}_{x\triangleright y}\big\{(X_{[0,\tau_{a_1}]})_1 = a_1 \in \bullet/\tau_{a_1} \in [0,\infty[\big\}\\ &= \mathbf{E}_{x\triangleright y}\{X_1 = a_1\} = \mathbf{P}_{\triangleright y}(x, a_1) > 0.\end{aligned}$$

By induction we deduce that $\mathbf{E}_{a_i\triangleright a_{i+1}}\{X_1 = a_{i+1}\} > 0$ for all i. $\square$

Lemma 14.6. *Fix $z \in E$. For all $a, b \in \{\bullet \overset{\triangleright}{\rightsquigarrow} z\}$ we have $\mathbf{E}_{a\triangleright b} = \mathbf{E}^{\mathbf{P}_{\triangleright z}}_{a\triangleright b}$.*

Proof. If $b = z$ then the statement is a tautology, so we suppose $b \neq z$.

Case 1: $a \overset{\triangleright}{\ast\!\!\rightsquigarrow} b$. So, from Lemma 14.5, $\mathbf{E}_{a\triangleright b}$ is not degenerated. Because $b \overset{\triangleright}{\rightsquigarrow} z$ and $b \neq z$ we also have $a \overset{\triangleright}{\ast\!\!\rightsquigarrow} b$ and so $\mathbf{E}_{b\triangleright z}$ is non-degenerated. From "cohesion" axiom $\mathbf{E}_{a\triangleright z}\{\exists t\colon\ X_t = b\} > 0$ and from (30) we deduce:

$$\mathbf{E}_{a\triangleright b} = \mathbf{E}_{a\triangleright z}\big\{X_{[0,\tau_b]} \in \bullet/\tau_b \in [0,\infty[\big\} = \mathbf{E}^{\mathbf{P}_{\triangleright z}}_{a\triangleright b}.$$

Case 2: $a \overset{\triangleright}{\ast\!\!\not\rightsquigarrow} b$. From Lemma 14.5 $\mathbf{E}_{a\triangleright b}$ is degenerated. If $\mathbf{E}^{\mathbf{P}_{\triangleright y}}_{a\triangleright b}$ would not be degenerated, (30) would bring a contradiction. $\square$

14.3. *Proof of Theorem* 14.2

Existence: We add to our state space a point $\ddagger$ which will play the role of a birth point. So our new state space is now $\{\ddagger\} \cup E \cup \{\dagger\}$. Let $\alpha > 0$ be a probability on E. We define $(\overline{\mathbf{E}}_{x\triangleright z})_{x,z\in\{\ddagger\}\cup E}$ by: $\forall x, z \in E$ $\overline{\mathbf{E}}_{x\triangleright z} = \mathbf{E}_{x\triangleright z}$, $\forall x \in E$ $\mathbf{E}_{x\triangleright\ddagger}$ is degenerated, $\forall z \in E$ $\mathbf{E}_{\ddagger\triangleright z}$ is the law of a trajectory going from $\ddagger$ to a in one step with probability $\alpha(a)$ and then going from a to z following the law $\mathbf{E}_{a\triangleright z}$. It is straightforward that this new kernel is still a Markov-bridge-kernel.

We consider its reversed version $\overline{\mathbf{E}}^{\top}_{\bullet\triangleright\bullet}$ (see Remark 1.6). The associated graph is denoted by $(\{\ddagger\} \cup E, \overset{\triangleleft}{\to})$. By construction we have $\forall a \in E$ $a \overset{\triangleleft}{\rightsquigarrow} \ddagger$. Let Q be the transition matrix of $\overline{\mathbf{E}}^{\top}_{\bullet\triangleright\ddagger}$; this matrix is defined on $\{\ddagger\} \cup E$. Lemma 14.6 shows that $\overline{\mathbf{E}}^{\top}_{\bullet\triangleright\bullet} = \mathbf{E}^{DQ}_{\bullet\triangleright\bullet}$ and from Theorem 4.3 we get that $\overline{\mathbf{E}}_{\bullet\triangleright\bullet} = \mathbf{E}^{DQ^{\top}}_{\bullet\triangleright\bullet}$. Taking P as the restriction of $Q^{\top}$ to E, E being $Q^{\top}$-closed, we get $\mathbf{E}_{\bullet\triangleright\bullet} = \mathbf{E}^{DP}_{\bullet\triangleright\bullet}$.

Graph: Let P be the matrix constructed previously. We have:

$$\begin{aligned} a \overset{\triangleright}{\to} b \quad &\Leftrightarrow \quad \mathbf{E}_{a\triangleright b}\{X_1 = b\} > 0 \\ &\Leftrightarrow \quad \mathbf{E}^{DP}_{a\triangleright b}\{X_1 = b\} > 0 \\ &\Leftrightarrow \quad \frac{U_{[P]}(b,b)}{U_{[P]}(a,b)} P(a,b) > 0 \quad \text{with } \frac{0}{0} = 0 \\ &\Leftrightarrow \quad P(a,b) > 0. \end{aligned}$$

We deduce that the directed graph of P is $(E, \overset{\triangleright}{\to})$.

Unicity: Suppose $\mathbf{E}_{\bullet\triangleright\bullet} = \mathbf{E}^{DP}_{\bullet\triangleright\bullet} = \mathbf{E}^{DP'}_{\bullet\triangleright\bullet}$. From the previous point, directed graphs of P, P' are $(E, \overset{\triangleright}{\to})$ whose, by hypothesis, admits a directed spanning tree. Theorem 13.4, applied with $K = E$, gives $DP = DP'$.

Acknowledgments

Thank you to Andrew Eastman who kindly change my frenglish into english. Thank you to the referee who, by his numerous suggestions, has really contributed to improve this article.

References

[1] J. Bertoin. *Lévy Processes*. *Cambridge Tracts in Mathematics* **121**. Cambridge Univ. Press, Cambridge, 1998. MR1406564
[2] J. D. Biggins. Random walk conditioned to stay positive. *J. London Math. Soc. (2)* **67** (2003) 259–272. MR1942425
[3] R. M. Blumenthal and R. K. Getoor. *Markov Processes and Potential Theory*. *Pure and Applied Mathematics* **29**. Academic Press, New York–London, 1968. MR0264757
[4] C. Dellacherie and P.-A. Meyer. *Probabilités et potentiel. Chapitres IX à XI: Théorie discrète du potentiel. Publications de l'Institut de Mathématique de l'Université de Strasbourg* **XVIII**. *Actualités Scientifiques et Industrielles* **1410**. Hermann, Paris, 1983. MR0727641
[5] J. L. Doob. Conditional Brownian motion and the boundary limits of harmonic functions. *Bull. Soc. Math. France* **85** (1957) 431–458. MR0109961
[6] E. B. Dynkin. Boundary theory of Markov processes (the discrete case). *Russian Math. Surveys* **24** (1969) 1–42. MR0245096
[7] W. Feller. *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edition. Wiley, 1966. MR0210154
[8] P. J. Fitzsimmons. On the excursions of Markov processes in classical duality. *Probab. Theory Related Fields* **75** (1987) 159–178. MR0885460
[9] P. Fitzsimmons. Markov processes with identical bridges. *Electron. J. Probab.* **3** (1998). MR1641066
[10] P. Fitzsimmons, J. Pitman and M. Yor. Markovian bridges: Construction, Palm interpretation, and splicing. In *Seminar on Stochastic Processes* 101–134. *Progress in Probability* **33**. Birkhäuser Boston, Boston, 1992. MR1278079
[11] S. Fourati. Vervaat et Lévy. *Ann. Inst. H. Poincaré Probab. Statist.* **41** (2005) 461–478. MR2139029
[12] R. K. Getoor and M. J. Sharpe. The Markov property at co-optional times. *Z. Wahrsch. Verw. Gebiete* **48** (1979) 201–211. MR0534845
[13] R. K. Getoor and M. J. Sharpe. Excursions of dual processes. *Adv. Math.* **45** (1982) 259–309. MR0673804
[14] J. Hoffmann-Jørgensen. Markov sets. *Math. Scand*. **24** (1969) 145–166. MR0256460
[15] G. A. Hunt. Markoff processes and potentials. *Illinois J. Math.* **2** (1958) 151–213. MR0107097
[16] G. A. Hunt. Markoff chains and Martin boundaries. *Illinois J. Math.* **4** (1960) 316–340. MR0123364
[17] K. Itô. Poisson point processes attached to Markov processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III: Probability Theory* 225–239. Univ. California Press, Berkeley, 1972. MR0402949
[18] M. Jacobsen. Splitting times for Markov processes and a generalised Markov property for diffusions. *Z. Wahrsch. Verw. Gebiete* **30** (1974) 27–43. MR0375477
[19] M. Jacobsen and J. M. Pitman. Birth, death and conditioning of Markov chains. *Ann. Probab.* **5** (1977) 430–450. MR0445613
[20] O. Kallenberg. *Foundations of Modern Probability*, 2nd edition. *Probability and Its Applications (New York)*. Springer, New York, 2002. MR1876169
[21] A. N. Kolmogorov. Zur Theorie der Markoffschen Ketten. *Math. Ann.* **112** (1936) 155–160. MR1513044
[22] B. Maisonneuve. *Sytèmes régénératifs. Astérique* **15**. Société mathématique de France, 1974. MR0350879
[23] P. A. Meyer, R. T. Smythe and J. B. Walsh. Birth and death of Markov processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III: Probability Theory* 295–305. Univ. California Press, Berkeley, 1972. MR0405600
[24] P. W. Millar. Zero–one laws and the minimum of a Markov process. *Trans. Amer. Math. Soc.* **226** (1977) 365–391. MR0433606
[25] M. Nagasawa. Time reversions of Markov processes. *Nagoya Math. J.* **24** (1964) 177–204. MR0169290
[26] J. Pitman and M. Yor. Itô's excursion theory and its applications. *Japan J. Math.* **2** (2007) 83–96. MR2295611
[27] A. O. Pittenger and C. T. Shih. Coterminal families and the strong Markov property. *Trans. Amer. Math. Soc.* **182** (1973) 1–42. MR0336827
[28] H. Rost. Markoff–Ketten bei sich füllenden Löchern im Zustandsraum. *Ann. Inst. Fourier (Grenoble)* **21** (1971) 253–270. MR0299755

[29] E. Seneta. *Non-Negative Matrices. An Introduction to Theory and Applications.* Allen & Unwin, London, 1973. MR0389944
[30] T. Simon. Subordination in the wide sense for Lévy processes. *Probab. Theory Related Fields* **115** (1999) 445–477. MR1728917
[31] H. Thorisson. *Coupling, Stationarity, and Regeneration. Probability and Its Applications (New York)*. Springer, New York, 2000. MR1741181
[32] W. Vervaat. A relation between Brownian bridge and Brownian excursion. *Ann. Probab.* **7** (1979) 141–149. MR0515820
[33] V. Vigon. *Simplifiez vos Lévy en titillant la factorisation de Wiener–Hopf*. Editions Universitaires Europeennes, also disposable on HAL and on my web page, 2002.
[34] D. Williams. Decomposing the Brownian path. *Bull. Amer. Math. Soc.* **76** (1970) 871–873. MR0258130
[35] W. Woess. *Random Walks on Infinite Graphs and Groups. Cambridge Tracts in Mathematics* **138**. Cambridge Univ. Press, Cambridge, 2000. MR1743100