LU-factorization and probability

Vincent Vigon

LU-factorization and probability

Vincent Vigon, IRMA, Université de Strasbourg

September 23, 2011

Abstract

Our initial motivation was to understand links beetween WH-factorizations for random walks and LU-factorizations for Markov chains has interpreated by Grassman [Gra87]. Actually, first ones are particular cases of second ones, up to Fourier transforms. We produce a new proof of LU-factorizations which is valid for any Markov chain with a denumerable state space equiped with a pre-order relation. Factors have nice interpretations in term of subordinated Markov chains. In particular, the LU-factorization of the potential Matrice determine the law of the global minimum of the Markov chain.

For any matrice, there are two mains LU-factorizations according you decide to entry 1 in the diagonal of the first or of the second factor. When we factorize the generator of a general Markov chain, one factorization is always valid while the other require some hypothesis on the graph of the transition matrix. This disymetry come from the fact that the class of sub-stochastic matrices is not stable under transposition. We generalize our work to the class of matrices with spectral radius less that one; this allow us to play with transposition and so with time reversal.

We study some particular cases as: skip-free Markov chains, random walks (with gives the WH-factorization), reversible Markov chains (wich gives the Cholesky factorization). We use the LU-factorization to compute invariant measures. We exhibit some pathologies: non-associativity, non-unicity which can be cured by smooth assumptions (as irreductibility).

1 Introduction

The generator of a Markov chain admits a LU-factorization which can be proved and interpreted in virtues of probabilities. This was shown by Grassman [Gra87], extended by [Hey95], and Zhao, Li, Braun [ZLB97]. In many special cases, this factorization leads to interesting method to compute invariant measures and more generally to study structured Markov chains as the one appearing in queuing systems cf. Cao , Li , Zhao [LC04] [LZ02], [LZ04]. Recently, a book by Li [Li10] was completely devoted to this subject.

In an other part of the mathematical world, the LU-factorization was extensively studied for M-matrices (which includes generators of Markov chains) see Fiedler, Ptátk [FP62], Kuo [Kuo77] , Funderlic, Plemmons [FP81], Varga, Cai

[VC82], McDonald, Schneider [MS98]. But these studies were concentrated on finite matrices while probabilistic methods allow to work with infinite matrices (e.g. matrices indexed by $\mathbb{Z}$ as we will see). Of course, probabilists are not alone to do LU-factorization with infinite matrices c.f. Andrews, Smith , Ward [AW86], [ASW86].

But before all these works was known the Wiener-Hopf factorization for random walk see e.g. Feller [Fel66]. We will explain that, up to a Fourier Transform, the LU-factorization is the natural generalization of the WH-factorization. But be carefull, the WH-factorization was already generalized in an other direction (less natural we think) by Barlow, Rogers & Willams [BRW80], [Wil84], [Wil91], [Wil08].

We now produce a "mathematical" summary: Consider $(P(x,y))_{x,y\in E}$ a sub-stochastic matrix on a denumerable state space E equipped with a pre-order relation $\preceq$ (e.g. $E=\mathbb{Z}$ and $\preceq$ is $\leq$). We think P as the transition Matrix of an eventually dying Markov chain which can goes up or goes down in E. Let I be the identity matrix indexed by E. The matrix $I-P$ is called the generator, its "inverse" $U=I+P+P^2+...$ is called the potential Matrix.

We can always give a sense to $I-P=(I-L)(I-K)$ for some matrices $L\geq 0, K\geq 0$ with "triangular shape" i.e. $L(x,y)>0 \Leftrightarrow y\succeq x$ and $K(x,y)>0 \Leftrightarrow y\prec x$. Remark that our situation is very general: we can chose $\preceq$ to be either a pre-order or an order relation, so L,K are either block-triangular or triangular matrices. We didn't need that $\preceq$ is a well-order: contrary to classical methods, we will never make any recurrence on the states.

But more interesting : $I-K$ is itself the generator of a decreasing Markov chain $n\mapsto X_{\neg_n}$ while $I-L$ is, up to a Doob-transform, the generator of an increasing Markov chain $n\mapsto X_{\check{\neg}_n}$. Both $X_{\neg}$ and $X_{\check{\neg}}$ are some time-changes of the initial Markov chain X driven by P. On the random walk case, the excessive function involved in the Doob-transform of $I-L$ is constant, so we find out the classical WH-factorization.

To arrive to the factorization of $I-P$, we start to establish a general method to factorize the potential matrix $U=\sum_{t\in\mathbb{N}}P^t$. To go from $U=VW$ to $(I-P)=(I-L)(I-K)$ simply use the fact that generators are the inverse of potential matrices.

This simple program will be accomplish in few pages during sections 3 and 5. The remaining part of this article will be devoted to specifications and generalizations:

- Section 4: The factorization $U=VW$ will be disintegrated by the formula

$$U(x,z)\mathbf{P}_{x\triangleright z}\{X_{\neg_f}=y\}=V(x,y)W(y,z)$$

  where $\mathbf{P}_{x\triangleright z}$ is the "homogeneous bridge" while $X_{\neg_f}$ is the final value of $X_{\neg}$ which is also the global minima of the trajectory. This formula shows how the LU-factors give an expression for the law of the minimum.

- Section 6: In the factorization $(I-P)=(I-L)(I-K)$ we have imposed that the second factor has 1 on its diagonal. Surprisingly, the existence of

$I-P=(I-L')(I-K')$ with 1 on the diagonal of the first factor, requires an additional condition on P: it does not exist a recurrent state which is reachable from above and not leavable to bellow. These limit was already point out for M-matrices (see Varga, Cai [VC82]).

- Section 7: We derive other factorizations: the classical three terms LDU-factorizations, and factorizations called mixed factorizations, which can also have interesting trajectorial interpretations, as the "equation amicale inversée" of ??.
- Section 8: We generalize our factorizations from the class of sub-stochastic matrices to the class of non-negative matrices with spectral radius less that one. The advantage of this larger class, is that it is stable by transposition.
- Section 9: We see how the factorization changes when we time reverse the initial Markov chain.
- Section 10: We look at some special cases: In the skip-free Markov chain we give a formula which gives the determinant of $I-P$. In the random walk case we explain why the LU-factorization is the natural generalization of the Wiener-Hopf-factorization. In the case where E is a part of $\mathbb{Z}$ we give a special formula. In the case where P is reversible, we make the Cholesky factorization.
- Section 11: We come back to the work of Grassmann and Heyman who used the LU-factorization to compute the invariant measure of a positive recurrent Markov chain.
- Section 12: On our very general case, the LU-factorization has not only good properties. We give example of non-associativity and non-unicity.
- Section 13: We give an alternative proof of $(I-P)=(I-L')(I-K')$. This new proof just use trajectorial considerations and the Markov property (no algebraic inversion as in the previous one), but this new proof is quite tricky.

2 Notations and setting

We consider: E a denumerable state space, $\mathfrak{a}: E \mapsto \mathbb{R}$ a function which gives the "altitude" of states. Letters x, y, z are always element of E. We write $x \preceq y$ when $\mathfrak{a}(x) \leq \mathfrak{a}(y)$, $x \sim y$ when $\mathfrak{a}(x) = \mathfrak{a}(y)$. The relation $\preceq$ (also written $\preceq_{\mathfrak{a}}$ when necessary) is a pre-order relation on E (conversely, any pre-order relation can be constructed with an altitudinal function). We write shortly $\{\succeq y\} = \{x \in E : x \succeq y\}$.

We consider $\big(P(x,y)\big)_{x,y\in E}$ a sub-stochastic matrix on E. We add a cemetery point $\dagger$ to E and prolong P to $E_\dagger = E \cup \{\dagger\}$ by $\overline{P}(x,\dagger) = 1 - \sum_{y\in E} P(x,y)$, $\overline{P}(\dagger,\dagger) = 1$.

We write U or $U_{[P]}$ the potential matrix $\sum_{t\in\mathbb{N}} P^t$. We write $x \rightsquigarrow y$ to indicates that x goes to y in the oriented graph of P. This is also equivalent to $U(x,y) > 0$.

We consider $\mathbb{N}$ as the set of times. Letters s, t, n are always elements of $\mathbb{N}$. Intervals $[s,t],]s,t] = [s+1, t]$ are always discrete intervals.

Summations $\sum_x$ mean $\sum_{x\in E}$, summations $\sum_t$ mean $\sum_{t\in\mathbb{N}}$.

We denote by Ω the set of trajectories from $\mathbb{N}$ to $E_\dagger$. We write X the canonical process (the identity on Ω). We write $\mathbf{P}_x$ or $\mathbf{P}_x^P$ the probability on Ω which makes X a Markov chain starting at x and with transition matrix $\overline{P}$. In particular we have $\forall x, y \in E : \mathbf{P}_x\{X_1 = y\} = P(x,y)$. We write $\mathbf{E}_x$ or $\mathbf{E}_x^P$ the expectation under $\mathbf{P}_x$.

We consider α a σ-finite measure on E and write $\mathbf{P}_\alpha = \sum_x \alpha(x)\mathbf{P}_x$ and $\mathbf{E}_\alpha = \sum_x \alpha(x)\mathbf{E}_x$.

We write $\zeta = \min\{t : X_t = \dagger\} - 1$ (the last time before the death) and $T_x = \min\{t : X_t = x\}$. If $S < T$ are random times, we write $X_{[S,T]}$ the trajectory $X_S, X_{S+1}, ..., X_T, \dagger, \dagger, ...$ we write $X_{\overleftarrow{[S,T]}}$ the reversed trajectory $X_T, X_{T-1}, ..., X_S, \dagger, \dagger, ...$

By convention $\mathfrak{a}(\dagger) = +\infty$. We write shortly $X_{[S,T]} \succeq x$ to indicate that $X_S \succeq x, X_{S+1} \succeq x, ..., X_T \succeq x$.

When we have a function $h : E \mapsto \mathbb{R}_+$, we write $D_h P$ the matrix defined by $D_h P(x,y) = \frac{h(y)}{h(x)} P(x,y) 1_{\{h(x)>0\}}$ which is the Doob transformation. We write $P^\top$ the transposition of P. To avoid multiple parenthesis we use the following priority rule:

$$D_h P\, Q = (D_h P)\, Q \qquad \text{and} \qquad D_h P^\top = D_h(P^\top)$$

We write $I(x,y) = 1_{\{x=y\}}$ the identity matrix on E. Let $F \subset E$ we write $I_F = 1_{\{x=y\in F\}}$, $P_F(x,y) = 1_{\{x\in F\}} P(x,y) 1_{\{y\in F\}}$ (using Matrix multiplication, we can write $P_F = I_F P I_F$).

Quite all quantities we will use in this article depend on the main data which is P. This dependance is not always explicitly written. E.g. $U(x,y) = U_{[P]}(x,y) = I + P + P^2 + ...$ We will sometime change our data e.g. $U_{[P_F]} = I + P_F + P_F^2 + ...$, or also $U_{[qP]} = I + qP + q^2P^2 + ...$

All matrices $A_{[P]}$ we will introduce (called $K_{[P]}, V_{[P]}, L_{[P]}, W_{[P]}, ..$) will have the shape $A_{[P]}(x,y) = \mathbf{E}_x^P\big[\sum_t \mathfrak{f}(X_{[0,t]}) 1_{\{X_t=y\}}\big]$ for some positive functional $\mathfrak{f}$. For $q \in]0,1]$ we have:

$$A_{[qP]}(x,y) = \mathbf{E}_x^P\Big[\sum_{t\le\tau_q} \mathfrak{f}(X_{[0,t]}) 1_{\{X_t=y\}}\Big] = \sum_s q(1-q)^s \mathbf{E}_x^P\Big[\sum_{t\le\tau_s} \mathfrak{f}(X_{[0,t]}) 1_{\{X_t=y\}}\Big]$$

where τ_q is an independent geometric time. Thus $]0,1] \ni q \mapsto A_{[qP]}(x,y)$ is increasing and continuous.

3 General factorization of the potential matrix

3.1 A time changed process

Let S be any stopping time taking values in $[1,\zeta]\cup\{+\infty\}$. We define a time-change by "iterating" S as follows

$$\urcorner_0 = 0,\ \urcorner_1 = S,\ \urcorner_2 = \urcorner_1 + S\circ X_{[\urcorner_1,\zeta]},\ ...\urcorner_{n+1} = \urcorner_n + S\circ X_{[\urcorner_n,\zeta]} \qquad (1)$$

All $\urcorner_n$ are stopping times and $n\mapsto\urcorner_n$ is strictly increasing until it eventually jumps to $+\infty$. We write $\urcorner_f$ the last finite $\urcorner_n$ or, in other words, $f=\min\{n:\urcorner_n<\infty\}-1$.

Example 3.1 :

- If we chose $S=\min\{t\in[1,\zeta]: X_t\in F\}$, for $F\subset E$, then $(\urcorner_n)$ are all the passage times in F.
- But the example which will interest us after is $S=\min\{t: X_t\prec X_0\}$. In this case $X_\urcorner$ is a decreasing process (which means $X_{\urcorner_n}\prec X_{\urcorner_{n+1}}$ whenever $\urcorner_n<\infty$).

Proposition 3.2 *Under $\mathbf{P}_\alpha$, the process $n\mapsto X_{\urcorner_n}$ is a Markov chain.*

Proof: Applying the strong Markov property at the stopping time $\urcorner_n$ we get

$$\mathbf{E}_\alpha[\mathfrak{f}(X_{[0,\urcorner_n]})\,1_{\{X_{\urcorner_n}=x\}}\,\mathfrak{g}(X_{[\urcorner_n,\zeta]})] = \mathbf{E}_\alpha[\mathfrak{f}(X_{[0,\urcorner_n]})\,1_{\{X_{\urcorner_n}=x\}}]\ \mathbf{E}_x[\mathfrak{g}(X_{[0,\zeta]})]$$

for all positive or bounded functionals $\mathfrak{f},\mathfrak{g}$ on the space of trajectories. Take any $f: E_\dagger^{n+1}\mapsto\mathbb{R}$ and $g: E_\dagger\mapsto\mathbb{R}$. Put $\mathfrak{f}=f(X_{\urcorner_0},X_{\urcorner_1},...,X_{\urcorner_n})$, which satisfies $\mathfrak{f}(X_{[0,\urcorner_n]})=\mathfrak{f}$. Put $\mathfrak{g}(X_{[0,\zeta]})=g(X_{\urcorner_1})$ which satisfies $\mathfrak{g}(X_{[\urcorner_n,\zeta]})=g(X_{\urcorner_{n+1}})$. Applying the previous equality to these $\mathfrak{f},\mathfrak{g}$, we get :

$$\mathbf{E}_\alpha[f(X_{\urcorner_0},X_{\urcorner_1},...,X_{\urcorner_n})\,1_{\{X_{\urcorner_n}=x\}}\,g(X_{\urcorner_{n+1}})] = \\ \mathbf{E}_\alpha[f(X_{\urcorner_0},X_{\urcorner_1},...,X_{\urcorner_n})\,1_{\{X_{\urcorner_n}=x\}}]\ \mathbf{E}_x[g(X_{\urcorner_1})]$$

which clearly indicates the Markov property. □

Remark 3.3 the reader can verify that $n\mapsto(X_{\urcorner_n},\urcorner_n)$ is also a Markov chain.

3.2 factorization

Let us write

$$V(x,y)=\mathbf{E}_x[\sum_n 1_{\{X_{\urcorner_n}=y\}}]$$
$$W(x,y)=\mathbf{E}_x[\sum_{t<S} 1_{\{X_t=y\}}]$$

The first matrix is the potential matrix of the Markov chain $X_\urcorner$.

Proposition 3.4 *We have*

$$U(x,z) = \sum_{y \in E} V(x,y) W(y,z)$$

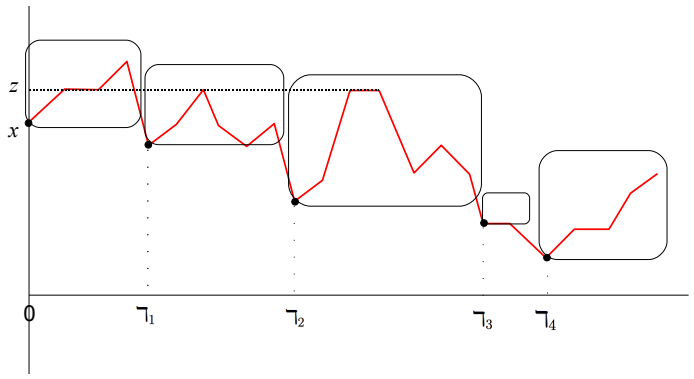
Proof: We can split $[0,\infty[$ into $\cup_n [\urcorner_n, \urcorner_{n+1}[$ to obtain:

$$\begin{aligned} U(x,z) &= \mathbf{E}_x \sum_t 1_{\{X_t = z\}} \\ &= \sum_n \sum_y \mathbf{E}_x \Big[1_{\{X_{\urcorner_n} = y\}} \Big(\sum_{t \in [0, \urcorner_1[} 1_{\{X_t = z\}} \Big) \circ X_{[\urcorner_n, \zeta]} \Big] \end{aligned} \tag{2}$$

Applying the Markov property at each stopping time $\urcorner_n$ we get:

$$U(x,z) = \sum_y V(x,y) W(y,z) \tag{3}$$

□



Here is an illustration of the equation (2) on the special case where $S = \min\{t : X_t \prec X_0\}$. This special case will occupy us during the sequel.

4 Disintegration of the general factorization

This section is independent to next ones.

In this section we suppose $U < \infty$. We fix three states x, y, z and suppose that $U(x,z) > 0$ (which means that $x \rightsquigarrow z$ in the oriented graph of P).

4.1 Recall about bridges

Let us denote by $\mathbf{P}_{x \triangleright z}$ the probability on Ω which makes X a Markov chain starting from x and with transition matrix $D_{U(\cdot,z)} P$. Under $\mathbf{P}_{x \triangleright z}$ the canonical process dies at z with probability one. We also have the following interpretation:

$$\mathbf{P}_{x \triangleright z} = \mathbf{P}_x [X_{[0,\tau_z]} \in \bullet \ / \ \tau_z > -\infty] \tag{4}$$

where $\tau_z = \sup\{t : X_t = z\}$. To a complete study of this bridge we send to Vigon [Vig11]. On this article we will see the proof of the following proposition:

Proposition 4.1 *For all functionals $\mathfrak{f}, \mathfrak{g} : \Omega \mapsto \mathbb{R}_+$, we have the "past-future extraction" under $\mathbf{P}_x$ i.e.:*

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

and we have the "past-future extraction" under $\mathbf{P}_{x\triangleright z}$ i.e.:

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

4.2 Let us disintegrate the factorization

Recall that we write $\daleth_f$ the last finite $\daleth_n$ or, in other words, $f = \min\{n : \daleth_n < \infty\} - 1$. On the special case where $S = \min\{t : X_t \prec X_0\}$, then $\daleth_f$ is the first global minimum of the trajectory.

Proposition 4.2 *We have:*

$$\mathbf{P}_{x\triangleright z}\{X_{\daleth_f} = y\}\, U(x,z) \;=\; V(x,y)\, W(y,z)$$

Remark 4.3 Summing over all y in the previous equation, we get $U(x,z) = VW(x,z)$ which is proposition 3.4.

Proof: Firstly, by the past-future extraction (proposition 4.1) applied to $\mathbf{P}_{x\triangleright z}$:

$$\begin{aligned}\mathbf{P}_{x\triangleright z}\{X_{\daleth_f} = y\} &= \sum_n \mathbf{P}_{x\triangleright z}\{X_{\daleth_n} = y, \daleth_{n+1} = \infty\} \\ &= \sum_t \sum_n \mathbf{P}_{x\triangleright z}\{\daleth_n = t, X_t = y, t + S(X_{[t,\zeta]}) = \infty\} \\ &= \sum_t \sum_n \mathbf{P}_{x\triangleright z}\{\daleth_n(X_{[0,t]}) = t, X_t = y, S(X_{[t,\zeta]}) = \infty\} \\ &= \sum_n \mathbf{P}_{x\triangleright y}\{X_{\daleth_n} = y\}\ \frac{U(x,y)U(y,z)}{U(x,z)}\ \mathbf{P}_{y\triangleright z}\{S = \infty\} \qquad (5)\end{aligned}$$

Secondly, by the past extraction applied to $\mathbf{P}_x$:

$$\begin{aligned}V(x,y) = \sum_n \mathbf{P}_x\{X_{\daleth_n} = y\} &= \sum_t \sum_n \mathbf{P}_x\{\daleth_n(X_{[0,t]}) = t, X_t = y\} \\ &= \sum_n \mathbf{P}_{x\triangleright y}\{X_{\daleth_n} = y\}\ U(x,y) \qquad (6)\end{aligned}$$

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

$$W(y,z) = \mathbf{E}_y[\sum_{t<S} 1_{\{X_t=z\}}] = \mathbf{E}_y \sum_t (1_{\{S=\infty\}} \circ X_{[0,t]})\, 1_{\{X_t=z\}}$$
$$= \mathbf{P}_{y \triangleright z}\{S=\infty\}\, U(y,z) \qquad (7)$$

To gather formulae (5), (6), (7) gives the result. □

Proposition 4.4 *We have*

$$\mathbf{P}_x\{X_{\neg_f} = z\} U(z,z) = V(x,z) W(z,z)$$

Proof: Let τ_z be the last passage at z. Using (4), we have:

$$\begin{aligned}
\mathbf{P}_x\{X_{\neg_f} = z\} &= \mathbf{P}_x\{X_{\neg_f} \circ X_{[0,\tau_z]} = z, \tau_z > -\infty\} \\
&= \mathbf{P}_x\{X_{\neg_f} \circ X_{[0,\tau_z]} = z / \tau_z > -\infty\}\ \mathbf{P}_x\{T_z < \infty\} \\
&= \mathbf{P}_{x \triangleright z}\{X_{\neg_f} = z\}\ \frac{U(x,z)}{U(z,z)} && \text{from (4)} \\
&= \frac{V(x,z) W(z,z)}{U(z,z)} && \text{from prop. 4.2}
\end{aligned}$$

In the particular case where $S = \min\{t : X_t \prec X_0\}$, the previous proposition gives us a formula for the law of the global minimum under $\mathbf{P}_x$.

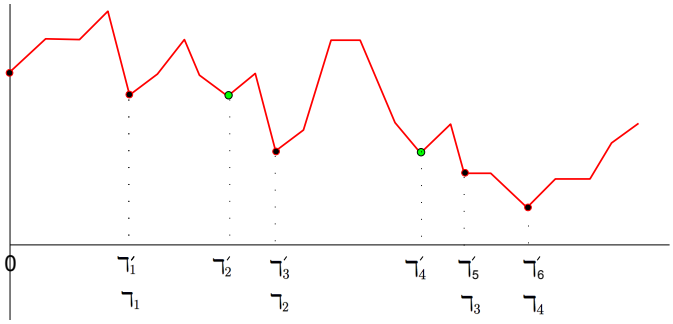
5 LU-factorizations

5.1 Factorizations of U

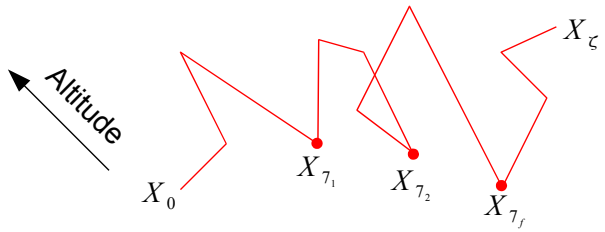
From now on, we chose two stopping times :

$$S = \min\{t : X_t \prec X_0\}$$
$$S' = \min\{t \geq 1 : X_t \preceq X_0\}$$

We keep all the notations from the previous section, the prime objects will be relative to S'.



Here we compare $\neg$ and $\neg'$. The state space is $\mathbb{Z}$ and the altitude is given by $\mathfrak{a}(x) = x$ so $\preceq$ is $\leq$.



Here the state space is $\mathbb{Z}^2 = \{(x_1, x_2)\}$ and the altitude is given by $\mathfrak{a}(x) = -x_1 + x_2$.

In this situation

- The time γ_f (resp. γ'_f) is the first (resp. the last) time where the process X reaches its global minimum. We write them shortly ρ (resp. ρ').
- The process X_γ (resp. $X_{\gamma'}$) is strictly (resp. largely) decreasing.
- $V(x, y) = 0$ for $y \succ x$ and moreover $V(x, y) = 1$ for $y \sim x$.
- $V'(x, y) = 0$ for $y \succ x$.
- $W(x, y) = 0$ for $y \prec x$.
- $W'(x, y) = 0$ for $y \prec x$ and moreover $W'(x, y) = 1$ for $y \sim x$.

The factorization given in proposition 3.4 admits two versions

$$U = VW$$
$$U = V'W'$$

which are the two classical LU-factorizations of the matrix U.

5.2 New functions

We write K and K' the transition matrices of Markov chains X_γ and $X_{\gamma'}$.

We write

$$k(x) = K(x, \dagger) = \mathbf{P}_x[X_{[0,\zeta]} \succeq x]$$
$$k'(x) = K'(x, \dagger) = \mathbf{P}_x[X_{]0,\zeta]} \succ x]$$

We recall the convention $\dagger \succ x$ for all $x \in E$. In particular: $X_1 = \dagger$ implies $X_{]0,\zeta]} \succ x\}$, $X_{\gamma_1} = \dagger$ and $X_{\gamma'_1} = \dagger$. We deduce the following lemma which will help us later:

Lemma 5.1 *$P(x, \dagger) > 0$ implies $k(x) > 0$ and $k'(x) > 0$.*

The next lemma links k and k' to laws of X_ρ and $X_{\rho'}$. We recall that $P_{\{\succ y\}}(a, b) = P(a, b)1_{\{a \succ y\}}1_{\{b \succ y\}}$ and $U_{[P_{\{\succ y\}}]} = I + P_{\{\succ y\}} + (P_{\{\succ y\}})^2 + ...$

Lemma 5.2 *We have:*

$$\mathbf{P}_x\{X_\rho = y\} = U_{[P_{\{\succ y\}}]}(x, y)k(y)$$
$$\mathbf{P}_x\{X_{\rho'} = y\} = U_{[P_{\{\succeq y\}}]}(x, y)k'(y)$$

Proof: First line:

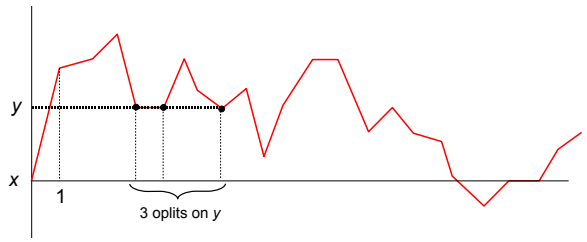
$$\begin{aligned}\mathbf{P}_x\{X_\rho = y\} &= \mathbf{E}_x\Big[\sum_t 1_{\{X_{[0,t[}\succ y\}} 1_{\{X_t=y\}} 1_{\{X_{[t,\zeta]}\succeq y\}}\Big] \\ &= \mathbf{E}_x\Big[\sum_t 1_{\{X_{[0,t[}\succ y\}} 1_{\{X_t=y\}}\Big]\mathbf{P}_y\{X_{[0,\zeta]}\succeq y\} \\ &= U_{[P_{\{\succ y\}}]}(x,y)k(y)\end{aligned}$$

The prime version is obtained replacing $\{X_{[0,t[}\succ y\}$ by $\{X_{[0,t]}\succeq y\}$ and $\{X_{[t,\zeta]}\succeq y\}$ by $\{X_{]t,\zeta]}\succ y\}$. □

5.3 New matrices

In the new definition we define the "oplits" (a made up word) which will help our intuition several times during the sequel.

Definition 5.3 A time t such that $X_t = y \succ X_0$ and $X_{]0,t]}\succeq y$ is called an oplit on y. We write oplit_y the set of oplits on y.



Then we denote by :

$$L(x,y) = \mathbf{E}_x\Big[\sum_t 1_{\{X_{]0,t[}\succ y\succeq X_0\}} 1_{\{X_t=y\}}\Big] \tag{8}$$

$$= 1_{\{x\preceq y\}}\mathbf{P}_x\{T_y = 1\} \;+\; 1_{\{x\preceq y\}}\mathbf{P}_x\{T_y\in[2,\infty[\ ,\ X_{]0,T_y-1]}\succ y\} \tag{9}$$

$$L'(x,y) = \mathbf{E}_x\Big[\sum_t 1_{\{X_{]0,t]}\succeq y\succ X_0\}} 1_{\{X_t=y\}}\Big] \tag{10}$$

$$= \mathbf{E}_x[\sharp\text{oplit}_y] \tag{11}$$

We remark immediately that L is always finite, but L' can be eventually infinite.

We define:

$$\begin{aligned}\rho_* &= \rho\circ X_{]0,\zeta]} \qquad \text{the first minimum strictly after 0} \\ \rho'_* &= \rho'\circ X_{]0,\zeta]} \qquad \text{the last minimum strictly after 0}\end{aligned}$$

Proposition 5.4 *We have*

$$\begin{aligned}\mathbf{P}_x\{X_{\rho_*} = y, X_{[0,\zeta]}\succeq x\} &= L(x,y)k(y) \\ \mathbf{P}_x\{X_{\rho'_*} = y, X_{]0,\zeta]}\succ x\} &= L'(x,y)k'(y)\end{aligned}$$

Remark 5.5 As a consequence we have $D_k L(x,y) = \mathbf{P}_x\{X_{\rho_*} = y / X_{[0,\zeta]} \succeq x\}$ taking the natural convention that conditioning by a null event give 0.

Proof: First line: For $x \succ y$ the equation is $0 = 0$. Let us assume $x \preceq y$. We have :

$$\begin{aligned}
&\mathbf{P}_x\{X_{\rho_*} = y, X_{[0,\zeta]} \succeq x\} \\
&= \mathbf{P}_x\{T_y = 1, X_{[1,\zeta]} \succeq y\} + \mathbf{P}_x\{T_y \in [2,\infty[\ ,\ X_{]0,T_y[} \succ y, X_{[T_y,\zeta]} \succeq y\} \\
&= \mathbf{P}_x\{T_y = 1\}\, \mathbf{P}_y\{X_{[0,\zeta]} \succeq y] + \mathbf{P}_x\{T_y \in [2,\infty[\ ,\ X_{]0,T_y[} \succ y\}\, \mathbf{P}_y\{X_{[0,\zeta]} \succeq y\} \\
&= L(x,y)k(y)
\end{aligned}$$

Second line: Having a look at the definition of an oplit we see that

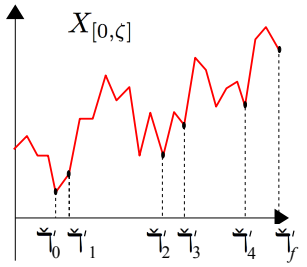
$$\mathbf{P}_x[X_{\rho'_*=y}, X_{]0,\zeta]} \succ x] = \mathbf{E}_x\Big[\sum_t 1_{\{t \in \mathrm{oplit}_y\}} 1_{\{X_t = y\}}\, 1_{\{X_{]t,\zeta]} \succ y\}}\Big] = L'(x,y)k'(y)$$

□

5.4 New processes

We define:

$$\begin{array}{llll}
\check{\neg}_0 = \rho, & \check{\neg}_1 = \rho_* \circ \check{\neg}_0, & \ldots & \check{\neg}_{n+1} = \check{\neg}_n + \rho_* \circ X_{[\check{\neg}_n,\zeta]} \\
\check{\neg}'_0 = \rho', & \check{\neg}'_1 = \rho'_* \circ \check{\neg}'_0, & \ldots & \check{\neg}'_{n+1} = \check{\neg}'_n + \rho'_* \circ X_{[\check{\neg}'_n,\zeta]}
\end{array}$$



Proposition 5.6 *Under $\mathbf{P}_\alpha$, processes $X_{\check{\neg}}$ and $X_{\check{\neg}'}$ are Markov chains whose transitions matrices are $D_k L$ and $D_{k'} L'$ and whose potential matrices are $D_k W$ and $D_{k'} W'$.*

Proof: We write as usual $X_{\check{\neg}_{[0,n]}} = (X_{\neg_1}, ..., X_{\neg_n}, \dagger, \dagger ...)$ and $X_{\check{\neg}_{[0,f]}} = (X_{\neg_1}, ..., X_{\neg_f}, \dagger, \dagger ...)$. We remark that

$$X_{\check{\neg}_{[0,n]}} = X_{\check{\neg}_{[0,n]}} \circ X_{[0,\check{\neg}_n]} = X_{\check{\neg}_{[0,f]}} \circ X_{[0,\check{\neg}_n]}$$

and that

$$1_{\{\check{\neg}_n = t\}} = 1_{\{\check{\neg}_n = \zeta\}} \circ X_{[0,t]} 1_{\{X_{[t,\zeta]} \succeq X_t\}}$$

Take $\mathfrak{f}$ a positive functional on Ω and g a positive function on $E_\dagger$. We have

$$\begin{aligned}
&\mathbf{E}_\alpha\Big[\mathfrak{f}(X_{\check{\gamma}_{[0,n]}})\,1_{\{X_{\check{\gamma}_n}=x\}}\,g(X_{\check{\gamma}_{n+1}})\Big]\\
&=\mathbf{E}_\alpha\Big[\sum_t \mathfrak{f}(X_{\check{\gamma}_{[0,f]}})\circ X_{[0,t]}\;1_{\{\check{\gamma}_n=t\}}1_{\{X_t=x\}}\;g(X_{\rho_*})\circ X_{[t,\zeta]}\Big]\\
&=\mathbf{E}_\alpha\Big[\sum_t \mathfrak{f}(X_{\check{\gamma}_{[0,f]}})\circ X_{[0,t]}\;1_{\{\check{\gamma}_n=\zeta\}}\circ X_{[0,t]}\;\;1_{\{X_t=x\}}\;1_{\{X_{[0,\zeta]}\succeq x\}}\circ X_{[t,\zeta]}\;g(X_{\rho_*})\circ X_{[t,\zeta]}\Big]\\
&=\mathbf{E}_\alpha\Big[\sum_t \mathfrak{f}(X_{\check{\gamma}_{[0,f]}})\circ X_{[0,t]}\;1_{\{\check{\gamma}_n=\zeta\}}\circ X_{[0,t]}\;\;1_{\{X_t=x\}}\Big]\;\mathbf{E}_x\Big[1_{\{X_{[0,\zeta]}\succeq x\}}\;g(X_{\rho_*})\Big]
\end{aligned}$$

With $g=1_{E_\dagger}$, that gives:

$$\begin{aligned}
&\mathbf{E}_\alpha\Big[\mathfrak{f}(X_{\check{\gamma}_{[0,n]}})\,1_{\{X_{\check{\gamma}_n}=x\}}\Big]\\
&=\mathbf{E}_\alpha\Big[\sum_t \mathfrak{f}(X_{\check{\gamma}_{[0,f]}})\circ X_{[0,t]}\;1_{\{\check{\gamma}_n=\zeta\}}\circ X_{[0,t]}\;\;1_{\{X_t=x\}}\Big]\;\mathbf{E}_x\Big[1_{\{X_{[0,\zeta]}\succeq x\}}\Big]
\end{aligned}$$

and so

$$\begin{aligned}
&\mathbf{E}_\alpha\Big[\mathfrak{f}(X_{\check{\gamma}_{[0,n]}})\,1_{\{X_{\check{\gamma}_n}=x\}}\,g(X_{\check{\gamma}_{n+1}})\Big]\\
&=\mathbf{E}_\alpha\Big[\mathfrak{f}(X_{\check{\gamma}_{[0,n]}})\,1_{\{X_{\check{\gamma}_n}=x\}}\Big]\;\mathbf{E}_x\Big[g(X_{\rho_*})/X_{[0,\zeta]}\succeq x\Big]\\
&=\mathbf{E}_\alpha\Big[\mathfrak{f}(X_{\check{\gamma}_{[0,n]}})\,1_{\{X_{\check{\gamma}_n}=x\}}\Big]\;D_kLg(x)
\end{aligned}$$

which indicate that $X_{\check{\gamma}}$ is a Markov chain with transition matrix D_kL.

Let us compute the potential matrix of $X_{\check{\gamma}}$.

$$\begin{aligned}
&\mathbf{E}_\alpha\Big[1_{\{X_{\check{\gamma}_0}=x\}}\;\sum_n 1_{\{X_{\check{\gamma}_n}=y\}}\Big]\\
&=\mathbf{E}_\alpha\Big[1_{\{X_\rho=x\}}\;\Big(\sum_n 1_{\{X_{\check{\gamma}_n}=y\}}\Big)\circ X_{[\rho,\zeta]}\Big]\\
&=\sum_t \mathbf{E}_\alpha\Big[1_{\{X_{[0,t[}\succ x\}}1_{\{X_t=x\}}1_{\{X_{[t,\zeta]}\succeq x\}}\;\Big(\sum_n 1_{\{X_{\check{\gamma}_n}=y\}}\Big)\circ X_{[t,\zeta]}\Big]\\
&=\sum_t \mathbf{E}_\alpha\Big[1_{\{X_{[0,t[}\succ x\}}1_{\{X_t=x\}}\Big]\;\mathbf{E}_x\Big[1_{\{X_{[0,\zeta]}\succeq x\}}\;\sum_n 1_{\{X_{\check{\gamma}_n}=y\}}\Big]\\
&=\sum_t \mathbf{E}_\alpha\Big[1_{\{X_{[0,t[}\succ x\}}1_{\{X_t=x\}}\Big]\;\mathbf{E}_x\Big[1_{\{X_{[0,\zeta]}\succeq x\}}\;\sum_t 1_{\{X_t=y\}}1_{\{X_{[t,\zeta]}\succeq y\}}\Big]\\
&=\sum_t \mathbf{E}_\alpha\Big[1_{\{X_{[0,t[}\succ x\}}1_{\{X_t=x\}}\Big]\;\mathbf{E}_x\Big[1_{\{X_{[0,\zeta]}\succeq x\}}\;\sum_{t<S} 1_{\{X_t=y\}}1_{\{X_{[t,\zeta]}\succeq y\}}\Big] \qquad (12)\\
&=\sum_t \mathbf{E}_\alpha\Big[1_{\{X_{[0,t[}\succ x\}}1_{\{X_t=x\}}\Big]\;\mathbf{E}_x\Big[\sum_{t<S} 1_{\{X_t=y\}}1_{\{X_{[t,\zeta]}\succ y\}}\Big] \qquad (13)\\
&=\sum_t \mathbf{E}_\alpha\Big[1_{\{X_{[0,t[}\succ x\}}1_{\{X_t=x\}}\Big]\;\mathbf{E}_x\Big[\sum_{t<S} 1_{\{X_t=y\}}\Big]\mathbf{P}_y[X_{[0,\zeta]}\succeq y]
\end{aligned}$$

so

$$\mathbf{E}_\alpha\Big[\ \sum_n 1_{\{X_{\tau_n}=y\}}/X_{\tau_0}=x\Big]=D_kW(x,y)$$

Here is some details for the previous computation:
(12): Because on $\{X_{[0,\zeta]}\succeq X_0\}$ we have $S=\infty$.
(13): Because on the complementary of $\{X_{[0,\zeta]}\succeq X_0\}$, we always have $\sum_{t<S}1_{\{X_t=y\}}1_{\{X_{[t,\zeta]}\succ y\}}=0$.

The prime version is very similar: just replace $X_{[0,t[}\succ x$ by $X_{[0,t]}\succeq x$ and $X_{[t,\zeta]}\succeq y$ by $X_{]t,\zeta]}\succ y$. □

Proposition 5.7 *We have:*

$$W=\sum_n L^n$$
$$W'=\sum_n L'^n$$

Proof: ▷ First step. Suppose that P is strictly sub-stochastic i.e. $P1_E\leq q<1$. So, from lemma 5.1 we have $k>0$. The transition matrix of X_τ is D_kL while its potential matrix is D_kW. So we have $D_kW=\sum_n(D_kL)^n=\sum_n D_k(L^n)$. Simplifying the k we deduce the proposition.

▷ Second step. Applying the first step to qP, with $q\in]0,1[$, we get $W_{[qP]}=\sum_n L^n_{[qP]}$. Then we make q tends to 1. The prime version is proven identically. □

5.5 Factorizations of the generator

When A,B are infinite matrices with signed coefficients, we say that the product AB is absolutely convergent when the product $|A|\,|B|$ is finite (where $|A|,|B|$ are matrices with coefficients $|A(x,y)|,|B(x,y)|$).

Proposition 5.8 *We have the following identity between positive matrices:*

$$P+LK=K+L \tag{14}$$
$$P+L'K'=K'+L' \tag{15}$$

We have the following factorization, with an absolutely convergent product:

$$(I-P)=(I-L)(I-K)$$

When L' is finite, we have the following factorization, with a absolutely convergent product:

$$(I-P)=(I-L')(I-K')$$

Proof: ▷ First step : Assume $P1 \leq q < 1$ (i.e. P is strictly sub-stochastic). Consequently $k > 0$ on E (lemma 5.1). Moreover $U1 \leq \frac{1}{1-q}$. From their definitions, we can see that matrices K, V, L, W are dominated by U. In this situation, all our matrices can be seen as bounded operator on $\ell^\infty(E)$. Equations $U = \sum_n P^n$, $V = \sum_n K^n$, $W = \sum_n L^n$ (proposition 5.7) show that U, V, W are the inverse operators of $(I-P), (I-K), (I-L)$. Inverting $U = VW$ gives us $(I-P) = (I-L)(I-K)$. We can develop this equation to obtain $P + LK = L + K$. The prime version is the same.

▷ Step 2: Suppose P is just a sub-stochastic matrix. We can apply the previous work to qP with $q \in]0,1[$. We obtain $qP + K_{[qP]}L_{[qP]} = K_{[qP]} + L_{[qP]}$. When $q \to 1$, we have $K_{[qP]} \uparrow K_{[P]}$ and $L_{[qP]} \uparrow L_{[P]}$. From monotone convergence $L_{[qP]}K_{[qP]} \uparrow L_{[P]}K_{[P]}$ so we get $P + KL = K + L$. Prime version is identical.

▷ Step 3: Suppose again that P is any sub-stochastic matrix. By its definition, we always have $L < \infty$. From $P + KL = K + L$ we have $LK \leq K + L$ and

$$|I - L|\,|I - K| \leq I + L + K + LK \leq I + 2L + 2K$$

From their definition K, L are finite so the product $(I-L)(I-K)$ is absolutely convergent. Finally $(I-L)(I-K) = I - K - L + KL = I - P$. The prime version is the same, except that we have to suppose first of all that $L' < \infty$. □

6 About existence of the prime factorization

The theorem 5.8 indicates that $L' < \infty$ is a sufficient condition to have the prime factorization $(I-P) = (I-L')(I-K')$ with an absolutely convergent product. This condition is also sufficient because when L' can take the value $+\infty$ the product $(I-L')(I-K')$ can not be absolutely convergent (except when we accept the convention $+\infty, 0 = 0$, in this case we have to think more).

In this section, we study the finiteness of L'. We recall that $L'(x,y)$ is the mean number of oplits on y starting from x (see definition 5.3).

6.1 Reformulation of $L' < \infty$

Recall that we write $x \rightsquigarrow y$ to indicate that x goes to y in the oriented graph of P.

Proposition 6.1 *$L' < \infty$ if and only if it does not exist a state y which is in the same time :*

1/ Recurrent	*i.e.* $U(y,y) = \infty$.
2/ Not leavable to below	*i.e.* $\forall z \prec y :\ y \not\rightsquigarrow z$.
3/ Reachable from below	*i.e.* $\exists x \prec y :\ x \rightsquigarrow y$.

Proof: Suppose that it exists a state y which satisfy 1, 2 ,3. Let x be a state such that $x \prec y$. We have

$$L'(x,y) = \mathbf{P}_x\{T_y < \infty\}\mathbf{E}_y\Big[\sum_t 1_{\{X_t=y, X_{]0,t]} \succeq y\}}\Big]$$
$$= \mathbf{P}_x\{T_y < \infty\}\mathbf{E}_y\Big[\sum_t 1_{\{X_t=y\}}\Big] = +\infty$$

Conversely. Suppose that either

- y is transient. So, for any x, we have $L(x,y) \leq U(x,y) \leq U(y,y) < \infty$.
- y is not reachable from below, so $L'(x,y) = 0$ for any x.
- y is leavable to below. So there exists $z \prec y$ such that $\mathbf{P}_y\{T_z < \infty\} > 0$. Once the process pass under y, there are not possibility of oplit. So the number of oplits at y is stochastically inferior to a geometric times with parameter $\mathbf{P}_y\{T_z < \infty\}$. So the expectation of this number is finite.

□

$F \subset E$ is called a "recurrent class" when F is irreducible and when that all states in F are recurrent. We write $x \prec F$ to indicate that $\forall y \in F : x \prec y$. We write $x \rightsquigarrow F$ to indicate that $\exists y \in F : x \rightsquigarrow F$.

Corollary 6.2 *Suppose that $E \subset \mathbb{Z}$ and that $\preceq$ is $\leq$. Then $L' < \infty$ iff it does not exist a recurrent class F and a state x such that $x \prec F$ and $x \rightsquigarrow F$.*

Proof: Suppose that L' is not finite. So there exists a recurrent point y which is reachable from a state $x \prec y$ and not leavable to bellow. Because y is not leavable to bellow, we have $y \not\rightsquigarrow x$, so that x is transient. Let us write F the recurrent class containing y. We have $x \prec F$ and $x \rightsquigarrow F$.

Reciprocally. Suppose that exists a recurrent class F and a state x such that $x \prec F$ and $x \rightsquigarrow F$. So F is bounded from bellow and it admits a smallest element y. This state y is recurrent, not leavable to bellow and reachable from bellow (from x). So L' is not finite. □

6.2 Prime factorization for a chosen altitudinal function

On the beginning of this article, we fix an altitudinal function $\mathfrak{a}$ which allows us to define our pre-order relation $\preceq$. To change $\mathfrak{a}$ to an other altitudinal function $\mathfrak{b}$ is equivalent to permute simultaneously some rows and columns of $(I - P)$. Such permutation can help to perform the LU-factorization. Matrices $K, V...$ computed with this new altitudinal function will be denoted by $K_{[\mathfrak{b}]}, V_{[\mathfrak{b}]}....$

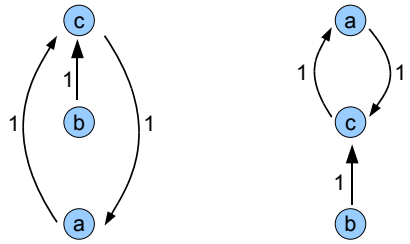
Proposition 6.3 *There always exists an altitudinal function $\mathfrak{b}$ on E such that $L'_{[\mathfrak{b}]} < \infty$. This function $\mathfrak{b}$ can even be taken injective (so the relative pre-order $\preceq_{\mathfrak{b}}$ is an order).*

Proof: Let $E_r \subset E$ be the set of recurrent states and $E_t = E \setminus E_r$ the set of transient states. Let $\mathfrak{b}$ be any function from E to $\mathbb{R}$ which maps E_r into $[0, \infty[$ and E_t into $]-\infty, 0[$ (such a function can easily been taken injective). Let us reason by contradiction: Suppose that it exists (x, y) such that y is recurrent, $x \prec_{\mathfrak{b}} y$, $x \rightsquigarrow y$ and $y \not\rightsquigarrow x$. This implies that x is transient which is absurd because, by construction of $\mathfrak{b}$, transient states are situated below recurrent states. □

6.3 Prime factorization for all altitudinal functions

Proposition 6.4 *The following points are equivalent:*

1. *For any altitudinal function $\mathfrak{b}$, we have $L'_{[\mathfrak{b}]} < \infty$.*
2. *It does not exist a couple of states (x, y) such that x is transient, y recurrent and $x \rightsquigarrow y$.*
3. *There exists a measure $\mu : E \mapsto \mathbb{R}_+^*$ such that $\mu P \leq \mu$.*



On the left we give situation where $L'_{[\mathfrak{a}]} < \infty$. But where we have a transient point which communicates to a recurrent point. On the right we changed the altitude so that $L'_{[\mathfrak{b}]}$ is not finite.

When P is irreducible, then either all states are recurrent, or all states are transient. From point 2 of the previous proposition we deduce:

Corollary 6.5 *When P is irreducible then for any altitudinal function $\mathfrak{b}$ we have $L'_{[\mathfrak{b}]} < \infty$.*

Proof of the proposition 6.4 is cut in lemmas.

Lemma 6.6 *First and second points of the proposition 6.4 are equivalent.*

Proof: Suppose that the first point is false. So it exists an altitudinal function $\mathfrak{b}$ such that there exist states x, y checking: $x \prec_{\mathfrak{b}} y$, $x \rightsquigarrow y$, y is recurrent and not leavable to below. In particular, that implies that x is transient which contradicts the second point.

Suppose the second point is false: there exist states x, y such that $x \rightsquigarrow y$, x is transient and y is recurrent. Take the altitudinal function $\mathfrak{b}$ which maps x into 0, y into 1, and $E \setminus \{x, y\}$ into $]1, \infty[$. Clearly, for this altitudinal function we have $L'_{[\mathfrak{b}]}(x, y) = +\infty$. □

Lemma 6.7 *If all points are transient or if all points are recurrent then it exists a measure $\mu : E \mapsto \mathbb{R}_+^*$ such that $\mu P \leq \mu$.*

Proof: If all points are transient then pick any probability α whose support is the whole E and define $\mu := \alpha U$. If all points are recurrent, then theorem 6-9 p.135 of [KSK66] insures us the existence of a measure $\mu : E \mapsto \mathbb{R}_+^*$ such that $\mu P = \mu$. □

Lemma 6.8 *Point two and point three of the proposition 6.4 are equivalent.*

Proof: Suppose that point two is true: Denote by E_t the set of transient points and by E_r the set of recurrent points. The previous lemma, applied to P_{E_t} and P_{E_s}, gives us two measures $\mu_t : E_t \mapsto \mathbb{R}_+^*$ and $\mu_r : E_r \mapsto \mathbb{R}_+^*$ such that $\mu_t P_{E_t} \leq \mu_t 1_{E_t}$ and $\mu_r P_{E_r} \leq \mu_r 1_{E_r}$. Now, because E_t and E_r are disconnected, we have $P = P_{E_t} + P_{E_r}$ and thus the measure $\mu = \mu_t + \mu_r$ checks $\mu P \leq \mu$.

Conversely, suppose point two is false and point three is true: So there exists $\mu : E \mapsto \mathbb{R}_+^*$ such that $\mu P = \mu$ and there exists a state x transient which communicates to a state y recurrent. We consider the measure $U(x, \cdot)$ which is finite at x and infinite at y. Put $k = U(x, x)$. We construct an other measure: $\nu(y) = \frac{k}{\mu(1)}\mu(y)$. The measure ν also checks $\nu P \leq \nu$ and moreover ν is greater that $U(x, \cdot)$ at x (both has the same value at this point). Because of the maximum principle, ν is greater that $U(x, \cdot)$ everywhere, in particular at y. This implies $\nu(y) = \mu(y) = \infty$ which contradicts our hypothesis on μ. □

6.4 Link with the existing criterion for M-matrix

A M-matrix is a finite matrix $\mathtt{A}$ of the shape $\mathtt{A} = c_{st}(I - Q)$ where Q is a matrix with non-negative entries such that the spectral radius of Q is less or equal that 1.

When E is finite, Markov chain generators $I - P$ are peculiar cases of M-matrix.

For a M-matrix $\mathtt{A}$, the $\mathtt{LU}$-factorization is possible for any ordering of indices (i.e. for any injective altitudinal function)

- if $\mathtt{A}$ is invertible (Fiedler and Ptak, 1962, [FP62])
- if $\mathtt{A}$ is irreducible (Kuo, 1977, [Kuo77])
- if there exists $\mu : E \mapsto R_+^*$ such that $\mu\mathtt{A} \geq 0$ (i.e. $\mu Q \leq \mu$) (Funderlic and Plemmons, 1981, [FP81]).

- If and only if it exists $\mu : E \mapsto \mathbb{R}_+^*$ such that $\mu \mathtt{A} \geq 0$ (Varga and Cai, 1981, [VC82]).

Finaly, Varga and Cai [VC82] gave necessary and sufficient conditions for a M-matrix to admit a LU-factorization for a fixed ordering of indices. Applying theorem 1 of [VC82] to the matrix $\mathtt{A} = I - P$, and we find our corollary 6.2. Applying theorem 5 of [VC82] to the matrix $\mathtt{A} = I - P$, and using the fact that recurrent states are absorbing, we see that $I - P$ always admits a LU-factorization; which we saw by $I - P = (I - L)(I - K)$. To help the translation between the present article and article of Varga and Cai, remark that: Firstly up and down are inverted. Secondly: to say that F is a recurrent class for P is equivalent to say that $(I_F - P_F)$ is a singular irreducible matrix.

7 More factorizations

7.1 The three terms LDU-factorization

We denote by $\tilde{S} = \min\{t \geq 1 : X_t \sim X_0, X_{]0,t[} \succ X_0\}$. We write $\tilde{\neg}$ the corresponding time-change. Proposition 3.2 shows that $n \mapsto X_{\tilde{\neg}_n}$ is a Markov chain. We write $\tilde{K}$ its transition matrix, $\tilde{V}$ its potential matrix.

If A, B, C are infinite matrices with signed coefficients, we say that the three terms product ABC is absolutely convergent when we have $|A||B||C| < \infty$. From Fubini theorem, such a product is associative.

Proposition 7.1 *We have the following identity between positive matrices*

$$V' = V\tilde{V}$$
$$U = V\tilde{V}W'$$

We have the following factorization with an absolutely convergent product:

$$(I - K') = (I - \tilde{K})(I - K)$$

Moreover, when $L' < \infty$ we have the following factorization with an absolutely convergent three terms product:

$$(I - P) = (I - L')(I - \tilde{K})(I - K)$$

Proof: All the proof is based on the application of the two factorizations $U = VW$ and $(I - P) = (I - L)(I - K)$ but starting from the process $X_{\neg'}$ instead of X. We note the following trajectorial fact:

$$X_{\neg} \circ X_{\neg'} = X_{\neg}$$
$$\Big(\sum_{t<S} 1_{\{X_t = y\}}\Big) \circ X_{\neg'} = \sum_n 1_{\{X_{\tilde{\neg}_n} = y\}}$$

So $V_{[K']} = V_{[P]}$ and $W_{[K']} = \tilde{V}_{[P]}$ and we get

$$V'_{[P]} = U_{[K']} = V_{[K']}W_{[K']} = V_{[P]}\tilde{V}_{[P]}$$

Now we have $U = V'W' = (V\tilde{V})W' = V\tilde{V}W'$ (the associativity is immediate because all terms are positive).

Similarly

$$\begin{aligned}
&L_{[K]}(x,y)\\
&= 1_{\{x\preceq y\}}\mathbf{P}_x[X_{\neg'_1} = y] + 1_{\{x\preceq y\}}\mathbf{P}_x\{T_y \circ X_{\neg'} \in [2,\infty[, X_{]0,T_y[} \circ X_{\neg'} \succ y\}\\
&= 1_{\{x\sim y\}}\mathbf{P}_x[X_{\neg_1} = y] + 0 = \tilde{K}(x,y)
\end{aligned}$$

So we have $(I - K') = (I - \tilde{K})(I - K)$.

Now we assume that $L' < \infty$. We have:

$$|I - L'|\,|I - \tilde{K}|\,|I - K| \leq I + K + \tilde{K} + \tilde{K}K + L' + L'K + L'\tilde{K} + L'\tilde{K}K$$

the only terms that are not obviously finite are $L'K, L'\tilde{K}$ and $L'\tilde{K}K$. We have:

$$\begin{aligned}
L'K(x,y) &= \sum_{a\succ y}\mathbf{E}_x\Big[\sum_t 1_{\{X_{]0,t]}\succeq a\}}1_{\{X_t = a\}}1_{\{X_{\neg_1} = y\}}\Big]\\
&= \sum_{a\succ y}\mathbf{P}_x\{X_{\rho'} \circ X_{]0,T_y]} = a\} \leq 1
\end{aligned}$$

We also have $L'\tilde{K} \leq L'$, so that $L'\tilde{K}K \leq L'K < \infty$.

Consequently, the three term product $(I - L')(I - \tilde{K})(I - K)$ is absolutely convergent, and so associative and we get $(I - P) = (I - L')[I - K'] = (I - L')[(I - \tilde{K})(I - K)] = (I - L')(I - \tilde{K})(I - K)$. □

7.2 Mixed factorization

Proposition 7.2 *We have*

$$\begin{aligned}
V + L &= I + PV\\
W + K &= I + WP\\
U &= V + UL\\
U &= W + KU
\end{aligned}$$

These equalities stay true adding prime on K, V, L, W.

Proof: Suppose firstly that P is strictly sub-stochastic, so all the involved matrices can be seen as bounded operators of $\ell^\infty(E)$.

Let us considerate $(I - P) = (I - L)(I - K)$. Multiplying it on the right by V we get the first equality. Multiplying it on the left by W we get the second equality.

Let us considerate $U = VW$. Multiplying it on the right by $(I - L)$ we get the third equality. Multiplying it on the left by $(I - K)$ we get the fourth equality.

To generalize this to any sub-stochastic P, we apply the proven equality to qP with $q \in]0, 1[$ and make q tends to 1. □

Here is a probabilistic proof and interpretation of the second factorization in the proposition 7.2.
▷ Fix $x \succ y$. So $W(x, y) = 0$. We have

$$\forall a \in E \qquad \mathbf{P}_x\{X_{S-1} = a, X_S = y\} = \mathbf{E}_x\Big[\sum_{t<S} 1_{\{X_t=a, X_{t+1}=y\}}\Big] = W(x, a)P(a, y) \tag{16}$$

Summing over all a we get $K(x, y) = WP(x, y)$.
▷ Fix $x \preceq y$. So $K(x, y) = 0$. Let us write $T_1, T_2, ...$ the successive passages at y. So $\{X_{T_i} = y\}$ means that the process visits at least i times the state y. Using the Markov property, we have

$$\begin{aligned}\forall a \in E \qquad \mathbf{E}_x\Big[\sum_i 1_{\{X_{T_i-1}=a\}} 1_{\{X_{T_i}=y\}} 1_{\{T_i<S\}}\Big] &= \mathbf{E}_x\Big[\sum_{t<S} 1_{\{X_t=a\}} 1_{\{X_{t+1}=y\}}\Big] \\ &= W(x, a)P(a, y)\end{aligned}$$

With the convention that $X_{-1} = \dagger$. Summing over all a, with a particular care to the case $x = y$, we get $W(x, y) - 1_{\{x=y\}} = WP(x, y)$.

Remark 7.3 The quantity $X_S - X_{S-1}$ appearing in (16) is the equivalent of the famous "under-shoot" in the fluctuation theory of Levy processes. The equation $K(x, y) = WP(x, y)$ is the equivalent of which we called «amicale équation inversée» in [Vig02], see also [Don01], chap. 5.

8 Matrices with spectral radius less that one

Nota Bene: Exceptionally during this section, our main data P is any matrix with non-negative entries.

In this section, we generalize our factorizations from the class of sub-stochastic matrices to the class of non-negative matrices with spectral radius less that one. The advantage of this larger class, is that it is stable by transposition, which will help us to understand some time-reversal symmetries (see the next section).

The generalization will be straightforward, but before we need to redefine $K_{[P]}, V_{[P]}, L_{[P]}, W_{[P]}$ (and prime versions) for any matrix P with non-negative entries. We also give a quick recall about the definition of the spectral radius for infinite matrices.

8.1 Unormalized bridge and occupation matrices

We define the σ-finite measure $\mathbf{F}^P_{x\triangleright z}$ on $\{\zeta < \infty\} \subset \Omega$ as follows: $\forall \omega = (\omega_0, \omega_1, ..., \omega_n, \dagger, \dagger, ...)$:

$$\mathbf{F}^P_{x\triangleright z}\{\omega\} = 1_{\{x_0=x\}} P(\omega_0, \omega_1) P(\omega_1, \omega_2) ... P(\omega_{n-1}, \omega_n) 1_{\{\omega_n = z\}}$$

For $\mathfrak{f} : \Omega \mapsto \mathbb{R}_+$ we write $\mathbf{F}^P_{x\triangleright z}[\mathfrak{f}] = \sum_\omega \mathfrak{f}(\omega) \mathbf{F}^P_{x\triangleright z}\{\omega\} 1_{\{\zeta(\omega)<\infty\}}$.

In the particular case where P is sub-stochastic we have

$$\mathbf{F}^P_{x\triangleright z}[\mathfrak{f}] = \mathbf{E}^P_x \Big[\sum_t \mathfrak{f}(X_{[0,t]}) 1_{\{X_t = y\}} \Big] = \mathbf{E}^P_{x\triangleright z}[\mathfrak{f}]\, U(x, z)$$

The last equality justifies the fact that $\mathbf{F}^P_{x\triangleright z}$ is called the "unormalized bridge" (from x to z). For more details about normalized and unormalized bridges, see [Vig11].

We will write shortly $h > 0$ to indicate that h is a function from E to $\mathbb{R}^*_+$. Here is some easily checkable properties of unormalized bridge:

$$\forall h > 0 : \mathbf{F}^{(D_h P)}_{x\triangleright z}[\mathfrak{f}] = \frac{h(z)}{h(x)} \mathbf{F}^P_{x\triangleright z}[\mathfrak{f}] \tag{17}$$

$$\mathbf{F}^{P^\top}_{x\triangleright z}[\mathfrak{f}] = \mathbf{F}^P_{z\triangleright x}[\mathfrak{f}(X_{\overleftarrow{[0,\zeta]}})] \tag{18}$$

$$\tag{19}$$

Let us call "occupation matrix" (relative to P) any matrix A such that its exists $\mathfrak{f} : \Omega \mapsto \mathbb{R}_+$ satisfying $A(x, y) = \mathbf{F}^P_{x\triangleright y}[\mathfrak{f}]$. The reader can play to check that the product of two occupation matrices is still an occupation matrix.

Our triangular matrices can be redefined as occupation matrices relative to P:

Definition 8.1 We write:

$$\begin{aligned}
K_{[P]}(x, y) &= \mathbf{F}^P_{x\triangleright y}[1_{\{S=\zeta\}}] \\
V_{[P]}(x, y) &= \mathbf{F}^P_{x\triangleright y}[1_{\{\zeta \in \mathrm{Rg}\urcorner\}}] \qquad \text{where } \mathrm{Rg}\urcorner = \{t : \exists n : \urcorner_t = n\} \\
L_{[P]}(x, y) &= \mathbf{F}^P_{x\triangleright y}[1_{\{X_{]0,\zeta[} \succ X_\zeta \succeq X_0\}}] \\
W_{[P]}(x, y) &= \mathbf{F}^P_{x\triangleright y}[1_{\{\zeta<S\}}]
\end{aligned}$$

$$\begin{aligned}
K'_{[P]}(x, y) &= \mathbf{F}^P_{x\triangleright y}[1_{\{S'=\zeta\}}] \\
V'_{[P]}(x, y) &= \mathbf{F}^P_{x\triangleright y}[1_{\{\zeta \in \mathrm{Rg}\urcorner'\}}] \qquad \text{where } \mathrm{Rg}\urcorner' = \{t : \exists n : \urcorner'_t = n\} \\
L'_{[P]}(x, y) &= \mathbf{F}^P_{x\triangleright y}[1_{\{X_{]0,\zeta]} \succeq X_\zeta \succ X_0\}}] \\
W'_{[P]}(x, y) &= \mathbf{F}^P_{x\triangleright y}[1_{\{\zeta<S'\}}]
\end{aligned}$$

Of course, when P is sub-stochastic, this new definition coincide with the old one, e.g.

$$K_{[P]}(x,y) = \mathbf{E}_x^P[1_{\{X_S=y\}}] = \mathbf{E}_x^P\Big[\sum_t 1_{\{S=\zeta\}} \circ X_{[0,t]}\ 1_{\{X_t=y\}}\Big] \tag{20}$$

$$= \mathbf{F}_{x\triangleright y}[1_{\{S=\zeta\}}] \tag{21}$$

$$L'_{[P]}(x,y) = \mathbf{E}_x^P\Big[\sum_t 1_{\{X_{]0,\zeta]}\succeq X_\zeta \succ X_0\}}) \circ X_{[0,t]}\ 1_{\{X_t=y\}}\Big] \tag{22}$$

$$= \mathbf{F}_{x\triangleright y}[1_{\{X_{]0,\zeta]}\succeq X_\zeta \succ X_0\}}] \tag{23}$$

8.2 Definition of the spectral radius

We recall some facts about the spectral radius. For details see [Vig11]. See also Woess [Woe00] chap.2 or Seneta [Sen73] chap. 6. Let us write:

$$\Lambda_{[P]} = \{\lambda > 0 : U_{[\frac{P}{\lambda}]} < \infty\}$$

it is clearly an interval of type $]\mathtt{R}, \infty[$ or $[\mathtt{R}, \infty[$ for some $\mathtt{R} \in \mathbb{R}_+ \cup \{+\infty\}$. This $\mathtt{R}$, also noted $\mathtt{R}_{[P]}$, is called the spectral radius of P. We are tempted to compare $\Lambda_{[P]}$ with:

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

We have $\Lambda'_{[P]} = \Lambda_{[P]}$ or $\Lambda'_{[P]} = \Lambda_{[P]} \cup \{\mathtt{R}_{[P]}\}$.

We always have $\Lambda_{[P]} = \Lambda_{[P^\top]}$, but $\Lambda'_{[P]}$ and $\Lambda'_{[P^\top]}$ can differ at their extremity, as shown by the example

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

On the special case where P is irreducible and $\mathtt{R}_{[P]} < \infty$, we have

$$\Lambda'_{[P]} = \Lambda'_{[P^\top]} = [\mathtt{R}_{[P]}, \infty[.$$

When E is finite and P irreducible, from Perron-Frobenius theorem, the definition of $\mathtt{R}_{[P]}$ coincide with the classical definition of spectral radius (the largest eigen values).

Finite matrices of type $c_{st}(I-P)$ with $\mathtt{R}_{[P]} \leq 1$ are usually called M-matrices (but some authors require $\mathtt{R}_{[P]} < 1$).

The three important facts to memorize about the spectral radius are:

- $\mathtt{R}_{[P]} = \mathtt{R}_{[P^T]}$.
- If $\mathtt{R}_{[P]} < 1$ then there exists $h > 0$ such that $D_h P$ is sub-stochastic.
- $\mathtt{R}_{[qP]} = q\mathtt{R}_{[P]}$

8.3 Generalized factorizations

Proposition 8.2 *Let P be such that $\mathtt{R}_{[P]} \leq 1$ then we have:*

$$\begin{aligned} U &= VW \\ U &= V'W' \\ P + LK &= K + L \\ P + L'K' &= K' + L' \end{aligned}$$

When $L < \infty$ and $K < \infty$ then we have the following factorization with an absolute convergent product:

$$(I - P) = (I - L)(I - K)$$

When $L' < \infty$ and $K' < \infty$ then we have the following factorization with an absolute convergent product:

$$(I - P) = (I - L')(I - K')$$

Proof: Factorization between non-negative matrices: First, suppose that $\mathtt{R}_{[P]} < 1$. So there exists $h > 0$ such that $D_h P$ is sub-stochastic. Then we can apply our known identities to $D_h P$ e.g. $U_{[D_h P]} = V_{[D_h P]} W_{[D_h P]}$, but all the Doob transforms can be simplified and we get the 4 first equality. Now suppose $\mathtt{R}_{[P]} = 1$, we can apply the previous work to qP with $q \in]0, 1[$ and make tends q to 1.

Factorization between signed matrices: Using the same technique, we firstly establish that $P + LK = K + L$, which implies $LK \leq K + L$. Then, assuming $L < \infty$ and $K < \infty$, we have $|I-L||I-K| \leq I+K+L+LK \leq I+2K+2L < \infty$. Then we finish the proof easily as in proposition 5.8. □

Remark 8.3 From their new definition, L, K, V, W and prime versions are occupations matrices of type $\mathbf{F}_{\bullet \triangleright \bullet}[\mathfrak{f}]$ with $\mathfrak{f}$ bounded by 1_Ω. As a consequence, they are dominated by $U = \mathbf{F}_{\bullet \triangleright \bullet}[1_\Omega]$. So the transience of P (i.e. $U_{[P]} < \infty$) implies that $\mathtt{R}_{[P]} \leq 1$ (from the definition of the spectral radius) but implies also the finiteness of all our matrices.

8.4 Other expressions for our matrices

Actually all our block-triangular matrices can be express by multiplying matrices $U_{[P_F]}$, P, I_F with $F = \{\succeq x\}, \{\succeq y\}, \{\succ x\}, \{\succ y\}$. We give here only two examples which are useful for the sequel.

Proposition 8.4 *Suppose that $\mathtt{R}_{[P]} \leq 1$. We have:*

$$\begin{aligned} K(x,y) &= 1_{\{x \succ y\}} (U_{[P_{\{\succeq x\}}]} P)(x,y) = (I_{\{\succ y\}} U_{[P_{\{\succeq x\}}]} P)(x,y) \\ L'(x,y) &= 1_{\{x \prec y\}} (P U_{[P_{\{\succeq y\}}]})(x,y) = (I_{\{\prec y\}} P U_{[P_{\{\succeq y\}}]})(x,y) \end{aligned}$$

Proof: ▷ Step 1. Suppose that P is sub-stochastic, then

$$\begin{aligned} K(x,y) &= \mathbf{P}_x^P\{X_{\neg_1} = y\} = \mathbf{E}_x^P\Big[\sum_t 1_{\{X_{[0,t]} \succeq x\}} 1_{\{X_t = y\}}\Big] 1_{\{y \prec x\}} \\ &= 1_{\{y \prec x\}} U_{[P_{\{\succeq x\}}]} P(x,y) \\ L'(x,y) &= \mathbf{E}_x^P\Big[\sum_t 1_{\{X_{]0,t]} \succeq y\}} 1_{\{X_t = y\}}\Big] 1_{\{y \succ x\}} \\ &= 1_{\{y \succ x\}} P U_{[P_{\{\succeq y\}}]}(x,y) \end{aligned}$$

▷ Step 2. Suppose that $\mathtt{R}_{[P]} < 1$. Then we apply the step 1 to $D_h P$ with $h > 0$ chosen such that $D_h P$ is sub-stochastic.

▷ Step 3. Suppose that $\mathtt{R}_{[P]} \leq 1$. Then we apply step 2 to qP with $q \in]0,1[$ and make q tends to 1. □

9 Time reversal

9.1 The proposition and its short proof

Proposition 9.1 *Suppose that* $\mathtt{R}_{[P]} \leq 1$. *We have:*

$$\begin{aligned} L'^{\top}_{[P]} &= K_{[P^\top]} \\ L^{\top}_{[P]} &= K'_{[P^\top]} \end{aligned}$$

Proof: From proposition 8.4:

$$\begin{aligned} L'_{[P]}(y,x) &= 1_{\{y \prec x\}} (P U_{[P_{\{\succeq x\}}]})(y,x) \\ &= 1_{\{y \prec x\}} (U^{\top}_{[P_{\{\succeq x\}}]} P^\top)(x,y) \\ &= 1_{\{y \prec x\}} (U_{[P^{\top}_{\{\succeq x\}}]} P^\top)(x,y) \\ &= K_{[P^\top]}(x,y) \end{aligned}$$

The second line is similar. □

Remark 9.2 Even if we start with P sub-stochastic, the matrix $P^\top$ is not necessary sub-stochastic. A purely sub-stochastic statement of the previous proposition 9.1 can be : $D_\pi L'^{\top}_{[P]} = K_{[D_\pi P^\top]}$ where $\pi > 0$ is an excessive measure (when its exists).

9.2 Probabilistic proof and interpretation of proposition 9.1

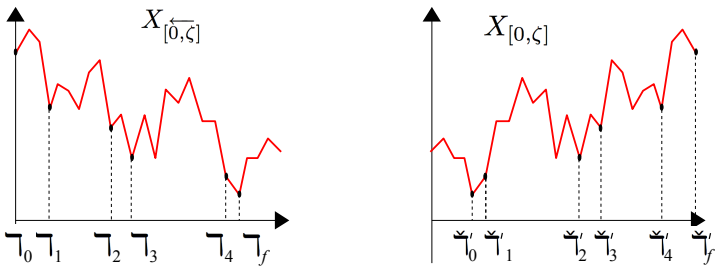
Nota bene: Our main data P is now again a sub-stochastic matrix.

During the previous sub-section, we have not really seen the link between proposition 9.1 and the title of this section. Of course, the two transpositions appearing in this proposition might be interpreted as two time reversal. Let us see this.

Proposition 9.3 *On* $\{\zeta < \infty\}$ *we have*

$$(X_{\check{\tau}'_f}, X_{\check{\tau}'_{f-1}}, ..., X_{\check{\tau}'_1}, X_{\check{\tau}'_0}) = (X_{\tau_0}, X_{\tau_1}, ..., X_{\tau_{f-1}}, X_{\tau_f}) \circ X_{\overleftarrow{[0,\zeta]}}$$

This trajectorial fact comes readily from the definition of $\check{\tau}'$ in sub-section 5.4. See also the following drawing.



The next general proposition is the discrete equivalent to the Tanaka time-reversal description. See [Vig11] for a proof.

Proposition 9.4 *Let* $Z = Z_0, Z_1, ..., Z_f, \dagger, \dagger..$ *be a dying Markov chain under a probability* $\mathbf{Q}$ *, with transition kernel* Q *and occupation measure* $\eta(y) = \sum_n \mathbf{Q}\{Z_n = y\}$*. Then the reversed Markov chain* $Z_f, Z_{f-1}, ..., Z_1, Z_0, \dagger, \dagger, ...$ *is also a Markov chain with transition kernel* $D_\eta(Q^\top)$ *i.e.*

$$\mathbf{Q}[Z_{f-1} = y / Z_f = x] = \frac{\eta(y)}{\eta(x)} Q(y,x) 1_{\{\eta(x)>0\}} \tag{24}$$

New proof of proposition 9.1: Firstly assume that P is strictly sub-stochastic. In particular, the Markov chain driven by P is dying. We write $\alpha U(y) = \mathbf{E}_\alpha[\sum_t 1_{\{X_t=y\}}]$ the occupation measure of X under $\mathbf{P}_\alpha$, so the matrix $D_{\alpha U}(P^\top) := \widehat{P}$ is the transition matrix of $X_{\overleftarrow{[0,\zeta]}}$. So we have $\mathbf{E}^P_\alpha[\mathfrak{f}(X_{\overleftarrow{[0,\zeta]}})] = \mathbf{E}^{\widehat{P}}_\beta[\mathfrak{f}(X_{[0,\zeta]})]$ where $\beta(z) = \mathbf{P}_\alpha\{X_\zeta = z\}$. From proposition 9.3, we have

$$\begin{aligned}\mathbf{E}^P_\alpha[\mathfrak{f}(X_{\check{\tau}'_f}, ..., X_{\check{\tau}'_0})] &= \mathbf{E}^P_\alpha[\mathfrak{f}(X_{\tau_0}, ..., X_{\tau_f}) \circ X_{\overleftarrow{[0,\zeta]}}] \\ &= \mathbf{E}^{\widehat{P}}_\beta[\mathfrak{f}(X_{\tau_0}, ..., X_{\tau_f})]\end{aligned}$$

So

$$\mathbf{P}^P_\alpha[X_{\check{\tau}'_{f-1}} = y \,/\, X_{\check{\tau}'_f} = x] = \mathbf{P}^{\widehat{P}}_\beta[X_{\tau_1=y} \,/\, X_{\tau_0=x}] = K_{[\widehat{P}]}(x,y) \tag{25}$$

Otherwise, the occupation measure of the Markov chain $X_{\check{\tau}'}$ under $\mathbf{E}^P_\alpha$ is given by

$$\eta(y) = \mathbf{E}^P_\alpha\Big[\sum_n 1_{\{X_{\check{\tau}_n}=y\}}\Big] = \mathbf{E}^P_\alpha\Big[\sum_n 1_{\{X_t=y\}} 1_{\{X_{]t,\zeta]} \succ y\}}\Big] = \alpha U(y) k'(y) \tag{26}$$

And if we apply the formula (24) we get

$$\begin{aligned}\mathbf{P}_\alpha^P[X_{\reflectbox{$\urcorner$}'_{f-1}} = y \,/\, X_{\reflectbox{$\urcorner$}'_f} = x] &= \frac{\alpha U(y)k'(y)}{\alpha U(x)k'(x)}(D_{k'}L')^\top(x,y)\\ &= \frac{\alpha U(y)k'(y)}{\alpha U(x)k'(x)}\frac{k'(x)}{k'(y)}L'(y,x)\\ &= \frac{\alpha U(y)}{\alpha U(x)}L'(y,x) \qquad (27)\end{aligned}$$

Gathering (26) and (27) we get $K_{[D_{\alpha U}(P^\top)]} = D_{[\alpha U]}L^\top_{[P]}$, then we simplify Doob transformations.

Secondly, if P is just sub-stochastic, we apply the previous result to qP and make q tends to 1. The proof of the second line of the proposition is similar. □

Remark 9.5 With the two previous propositions 9.3 and 9.4, we can obtain a short second proof of the Markovianity or $n \mapsto X_{\reflectbox{$\urcorner$}_n}$ (proposition 5.6) in the particular case where $\mathbf{P}_\alpha\{\zeta < \infty\} = 1$. Here are arguments: Under $\mathbf{P}_\alpha$, $X_{[0,\zeta]}$ is a Markov chain, so $X_{\overleftarrow{[0,\zeta]}}$ is a Markov chain, so $(X_{\reflectbox{$\urcorner$}_0}, ..., X_{\reflectbox{$\urcorner$}_f}) \circ X_{\overleftarrow{[0,\zeta]}}$ is a Markov chain, so $(X_{\reflectbox{$\urcorner$}'_f}, ..., X_{\reflectbox{$\urcorner$}'_0})$ is a Markov chain, so $(X_{\reflectbox{$\urcorner$}'_0}, ..., X_{\reflectbox{$\urcorner$}'_f})$ is a Markov chain.

10 Particular cases

10.1 Skip free

We say that we work on the "skip free situation" when $E = [a, b]$ is an finite interval of $\mathbb{Z}$, when $\preceq$ is $\leq$ and when $P(x, y) > 0 \;\Leftrightarrow\; y \geq x - 1$. So in this situation, our Markov chain can not make negative jumps strictly greater that 1.

We say that we work on the "conservative skip free situation" when we work on the skip free situation and when moreover $P(x, E) = 1$ for all $x \in E$, except eventually at $x = a$.

On the "conservative skip free situation", we can give without computation an expression for K, K' for all $x, y \in E$:

$$\begin{aligned}K(x,y) &= 1_{\{y=x-1\}}\\ K'(x,y) &= 1_{\{y=x\}}P(x,\{\succeq x\}) + 1_{\{y=x-1\}}P(x,x-1) \qquad (28)\end{aligned}$$

where we recall that $\{\succeq x\} = \{y : y \succeq x\} = \{y : y \geq x\}$ so $P(x, \{\succeq x\}) = \sum_{y \succeq x} P(x, y)$. Here is a consequence involving the determinant.

Proposition 10.1 *Suppose that we work on the conservative skip free situation. Then we have*

$$\det(I - P) = \prod_{x \in E}[1 - P(x, \{\succeq x\})]$$

Proof: We have $(I-P)=(I-L')(I-K')$. L' is triangular with entries 1 in its diagonal, so $\det(I-P)$ is equal to the product of the diagonal terms of K' which is given by (28). □

We can slightly generalize the previous proposition as follows:

Proposition 10.2 *Suppose that we work on the skip free situation. Suppose moreover that P is irreducible and transient. Recall that $a=\min E$. Then we have*

$$\det(I-P)=\prod_{x\in E}[1-D_{U(\cdot,a)}P(x,\{\succeq x\})]$$

Proof: From our hypothesis $U(\cdot,a)\in\mathbb{R}_+^*$. The Markov chain driven by $D_{U(\cdot,a)}P$ can only die at a, so this Doob transform carries us to the conservative skip free situation. Applying the previous proposition we get:

$$\det(I-D_{U(\cdot,a)}P)=\prod_{x\in E}[1-D_{U(\cdot,a)}P(x,\{\succeq x\})]$$

But $\det(I-D_{U(\cdot,a)}P)=\det(I-P)$. □

10.2 Random walk case.

In this sub-section we place ourselves in the "random walk situation" i.e.

- The state space is $E=\mathbb{Z}^d$.
- Our transition matrix is translation invariant: $P(x+a,y+a)=P(x,y)$. So, writing $\mathcal{P}(a)=P(0,a)$, we have $P(x,y)=\mathcal{P}(y-x)$.
- Our pre-order relation is translation invariant: $x+a\preceq y+a\ \Leftrightarrow\ x\preceq y$ (which is the case when the altitudinal function $\mathfrak{a}$ is a linear form).

From their definition it is not hard to see that matrices U,K,V,L,W and their prime versions are also translation invariant. So $K(x,y)=\mathcal{K}(y-x)$, $V(x,y)=\mathcal{V}(y-x)$, $L(x,y)=\mathcal{L}(y-x)$, $W(x,y)=\mathcal{W}(y-x)$... and our factorizations become

$$\mathcal{U}=\mathcal{V}*\mathcal{W}=\mathcal{V}'*\mathcal{W}'$$
$$(\delta-\mathcal{P})=(\delta-\mathcal{L})*(\delta-\mathcal{K})=(\delta-\mathcal{L}')*(\delta-\mathcal{K}')$$

where $*$ is the convolution product (which is commutative) and $\delta(x)=1_{\{x=0\}}$.

Functions k,k' are constant so we have $D_kL=L$, $D_{k'}L'$ so L,L' are sub-stochastic, which is not obvious from their definition. The sequel will allow us to interpret L,L' as the K',K matrices but relative to the opposite of our altitudinal function. So now we will write $K_{[P,\mathfrak{a}]},L_{[P,\mathfrak{a}]},...$ to precise our two data.

Proposition 10.3 *In the random walk situation we have:*

$$L'_{[P,\mathfrak{a}]} = K_{[P,-\mathfrak{a}]}$$
$$L_{[P,\mathfrak{a}]} = K'_{[P,-\mathfrak{a}]}$$

Proof: Remark first of all that $X\neg \circ (-X, -\mathfrak{a}) = -X\neg \circ (X, \mathfrak{a})$. Using this and the fact that $P^\top, K^\top$ are the transition matrices of $-X, -X\neg$ we get $K_{[P^\top,-\mathfrak{a}]} = K^\top_{[P,\mathfrak{a}]}$. Applying this with $P := P^\top$ we get $K_{[P,-\mathfrak{a}]} = K^\top_{[P^\top,\mathfrak{a}]}$. Finally, we use the proposition 9.1 to see that this last matrix is also $L'_{[P,\mathfrak{a}]}$.

The second line is similar. □

10.3 LU becomes WH (Wiener-Hopf).

We stay on the hypothesis on the previous sub-section but, to simplify the presentation, we add the hypothesis that $E = \mathbb{Z}^1$ and $\mathfrak{a}$ is the identity function (so $\preceq$ is $\leq$).

Let $\mathcal{G}$ be any function from E to $\mathbb{R}$. We define its Fourier/Laplace transform with the following convention:

$$\mathfrak{F}\mathcal{G}(z) = \sum_x \mathcal{G}(x)e^{-zx} \qquad \text{for all } z \in \mathbb{C} \text{ such that } \sum_x |\mathcal{G}|(x)e^{-(\Re z)x} < \infty$$

In particular, because functions $\mathcal{K}, \mathcal{L}$ are summable, their Fourier/Laplace transforms are well defined at least for z such that $\Re z = 0$. But it is usual that random walks have bounded jumps, to that $\mathcal{K}, \mathcal{L}$ have bounded support and their Fourier/Laplace transform are defined in the whole $\mathbb{C}$. The Fourier/Laplace transform turns convolution into product, so the LU-factorization becomes:

$$\big(1 - \mathfrak{F}\mathcal{P}\big)(z) = \big(1 - \mathfrak{F}\mathcal{L}\big)(z)\big(1 - \mathfrak{F}\mathcal{K}\big)(z) \qquad \text{at least for } \Re z = 0 \qquad (29)$$

This factorization is called the Wiener-Hopf factorization (see remark below). When P is strictly sub-stochastic (i.e. $\sum_a \mathcal{P}(a) < 1$) then $\mathcal{U}$ is also summable, and the LU-decomposition of U can be transform into:

$$\mathfrak{F}\mathcal{U}(z) = \mathfrak{F}\mathcal{V}(z)\ \mathfrak{F}\mathcal{W}(z) \qquad \text{at least for } \Re z = 0 \qquad (30)$$

But this equation is exactly the preceding one (29), where you inverse each terms (which cannot cancel when P is strictly sub-stochastic).

Remark 10.4 $\mathcal{K}$ is supported by $\mathbb{Z}_-$ so $\mathfrak{F}\mathcal{K}(z)$ is defined and analytic at least for $\Re z \leq 0$. While $\mathcal{L}$ is supported by $\mathbb{Z}_+$ so $\mathfrak{F}\mathcal{L}(z)$ is defined and analytic at least for $\Re z \geq 0$. In mathematical analysis, the decomposition of a function defined on band of type $\Re z \in [a, b]$ into the product or into the sum of a first analytic function defined on $\Re z \in [a, \infty[$ and a second one defined on $\Re z \in]-\infty, a]$ is called Wiener-Hopf decomposition (see [Wid97]). For this reason, the LU-factorization of random walk is called the Wiener-Hopf factorization.

10.4 Cholesky decomposition for the reversible case

Theorem 10.5 *Suppose that the altitudinal function $\mathfrak{a}$ is injective. Suppose that there exists a measure $\pi > 0$ such that $D_\pi(P^\top) = P$. Then there exists a sub-stochastic matrix Q such that $Q(x,y) > 0 \Leftrightarrow x \succeq y$ and such that we have the following factorization with an absolutely convergent product:*

$$I - P = D_\pi (I-Q)^\top \ (I-Q) \ .$$

Proof: Because $\mathfrak{a}$ is injective, the matrix $\tilde{K}$ is diagonal (i.e. $\tilde{K}(x,y) > 0 \Leftrightarrow x = y$) and its entries belong to $[0,1]$. The matrix $I - \tilde{K}$ has same properties. Let us define Z the diagonal matrix whose entries are square roots of those of $I - \tilde{K}_{[P]}$.

The existence of an excessive measure $\pi > 0$ ensure us that $L'_{[P]} < \infty$, so proposition 7.1 gives us:

$$(I-P) = (I - L'_{[P]})(I - \tilde{K}_{[P]})(I - K'_{[P]})$$

By proposition 9.1 we also have

$$L'_{[P]} = (K_{[P^\top]})^\top = (D_{\frac{1}{\pi}} K_{[D_\pi P^\top]})^\top = (D_{\frac{1}{\pi}} K_{[P]})^\top$$

So that:

$$\begin{aligned} I - P &= (I - D_{\frac{1}{\pi}} K)^\top \ (I - \tilde{K}) \ (I - K) \\ &= (I - D_{\frac{1}{\pi}} K)^\top \ Z^\top Z \ (I - K) \\ &= [Z(I - D_{\frac{1}{\pi}} K)]^\top \ [Z(I-K)] \\ &= [I - (ZD_{\frac{1}{\pi}} K + I - Z)]^\top \ [I - (ZK + I - Z)] \\ &= D_\pi [I - (ZK + I - Z)]^\top \ [I - (ZK + I - Z)] \end{aligned}$$

Now it remains to verify that $Q := (ZK + I - Z)$ checks the announced property. Clearly Q has positive entries and $Q(x,y) > 0 \Leftrightarrow x \succeq y$. Moreover we have:

$$Q1(x) = \sum_y Z(x,x)K(x,y) + 1 - Z(x,x) \leq Z(x,x) + 1 - Z(x,x) \leq 1$$

So Q is sub-stochastic. □

Remark 10.6 :

- After this proposition a natural question arises: What is the probabilistic interpretation of Q? Actually Q is the transition Matrix of the Markov chain $X_\neg$ that we speed down. Let us explain this fact in its whole generality: Consider any sub-stochastic matrix P and a diagonal matrix Z with entries in $[0,1]$. Let us consider X the Markov chain driven by P. We construct the slowed-down Markov chain X' by: Each time X'_t is in a state x, we play to dice, with probability $1 - Z(x)$, X'_{t+1} stays at x, with probability $Z(x)$, X'_{t+1} jumps as the initial process X to a state y with probability $P(x,y)$. Clearly, the transition matrix of X' is $ZP + (I - Z)$.

- Even in the non-reversible case, the technic we used in the previous proposition can be used to define other LU-factorizations, where we do not impose to the first nor the to second factors to have 1 on their diagonal.

10.5 A special formula when $E \subset \mathbb{Z}$.

In this subsection, we suppose that $E \subset \mathbb{Z}$ and $\preceq$ is $\leq$. We suppose also that $U_{[P]} < \infty$.

From proposition 4.1 we deduce that

$$\begin{aligned} U_{[P_F]}(x,z) &= \mathbf{E}_x\Big[\sum_t 1_{\{X_t \in F\}} 1_{\{X_t = z\}}\Big] \\ &= \mathbf{P}_{x \triangleright z}\{X_{[0,\zeta]} \subset F\}\, U(x,z) \end{aligned}$$

In particular, with $F = \{\succeq y\}$ we get

$$U_{[P_{\{\succeq y\}}]}(x,z) = \mathbf{P}_{x \triangleright z}\{X_{[0,\zeta]} \succeq y\} U(x,z) = \mathbf{P}_{x \triangleright z}\{X_\rho \succeq y\} U(x,z)$$

so that

$$\mathbf{P}_{x \triangleright z}\{X_\rho = y\}\, U(x,z) = U_{[P_{\succeq y}]}(x,z) - U_{[P_{\succeq y-1}]}(x,z)$$

Coupling this with proposition 4.2 we get a nice formula:

$$V(x,y)W(y,z) = U_{[P_{\succeq y}]}(x,z) - U_{[P_{\succeq y-1}]}(x,z)$$

10.6 Factorization with the reduce operator

Let us take $F \subset E$ and write $D = F^c$. So E is partitioned into D (the domain) and F (the frontier). We write $T_F = \min\{t \geq 0 : X_t \in F\}$ and

$$R_F(x,y) = \mathbf{P}_x\{X_{T_F} = y\}$$

Let us define the altitudinal function $\mathfrak{a}$ by $\mathfrak{a}(x) = 1$ on D and $\mathfrak{a}(x) = 0$ on F, then we see that $I_D K = I_D R_F$ and $I_D L = P_D$. So the LU-factorization of $(I - P)$, restricted to D gives:

$$I_D(I - P) = (I_D - P_D)(I_D - R_F)$$

But this identity can also be proven directly using the expression:

$$\begin{aligned} R_F(x,y) &= 1_{\{x=y \in F\}} + 1_{\{x \in D\}} \mathbf{E}_x\Big[\sum_t 1_{\{X_{[0,t]} \subset D\}} 1_{\{X_{t+1} = y \in F\}}\Big] \\ &= I_F(x,y) + I_D U_{[P_D]} P I_F(x,y) \end{aligned}$$

11 To compute the invariant measure

The LU-factorization is used in numerical analysis to solve linear equations. Using this idea, Grassmann [Gra87], working with P finite irreducible and recurrent give a method to compute the invariant measure. Heyman [Hey95] generalize this to $E = \mathbb{N}$ and P irreducible and positive recurrent. Both authors work with the prime factorization: $(I - P) = (I - L')(I - K')$ (always valid in the irreducible case).

In this section we explain Grassmann and Heyman technique and also transpose this technique with the factorization $(I - P) = (I - L)(I - K)$.

We suppose that $E = \{1, 2, ..., n\}$ or $E = \{1, 2, ..., n, ...\}$, that $\preceq$ is $\leq$, and that P is irreducible. In this situation $L' < \infty$ and the prime factorization is absolutely convergent (corollary 6.5).

Recall that a three terms product of infinite matrices $A.B.C$ is "absolutely convergent" when $|A||B||C| < \infty$. In this case, from Fubini theorem, this product is associative.

Lemma 11.1 *Let π be any finite measure. Three term products $\pi(I-L)(I-K)$ and $\pi(I - L')(I - K')$ are absolutely convergent.*

Proof: From (14) we have $LK \leq K + L$ so

$$\pi \left|I - L\right| \left|I - K\right| \leq \pi + \pi L + \pi K + \pi LK \leq \pi + 2\pi L + 2\pi K.$$

Because K is sub-stochastic we have $\pi K < \infty$. Then we have $\pi L(y) = \sum_x \pi(x)L(x, y)1_{\{x \leq y\}}$ which is finite because we sum is taken on a finite number of indices (because $E = \mathbb{N}^*$). The prime version is identical. □

Proposition 11.2 *Suppose that P is positive recurrent (it is the case when E is finite and P is stochastic). Then the invariant measures can be determined recursively by one or the other following expressions:*

$$\pi_1 \text{ is arbitrary and for } n \geq 2\text{: } \pi(n) = \sum_{i=1}^{n-1} \pi(i)L'(i, n)$$

$$\pi_1 \text{ is arbitrary and for } n \geq 2\text{: } \pi(n) = \frac{1}{1 - L(n, n)} \sum_{i=1}^{n-1} \pi(i)L(i, n)$$

Proof: From the associativity (lemma 11.1), to compute all measures π satisfying $\pi(I - P) = \pi(I - L')(I - K') = 0$ is equivalent to compute all measures β such that $\beta(I - K') = 0$ and then all measures π such that $\pi(I - L') = \beta$.

From the recurrence hypothesis, we see that $K'(1, 1) = 1$. Looking at the triangular linear system $\beta(I - K') = 0$ we see that all solutions are given by $\pi = (\beta_1, 0, 0, ...)$ where β_1 is arbitrary. Looking at the triangular linear system $\pi(I - L') = \beta$ we find the first recursive expression.

Now let us look the equation $\pi(I-P) = \pi(I-L)(I-K) = 0$. Looking at the triangular linear system $\beta(I - K) = 0$, we see that the only solution is $\beta = 0$.

Then, from the very definition of L (see equation (9)) we have $L(1,1) = 1$ and looking at the triangular linear system $\pi(I-L) = 0$ we find the second recursive expression. □

Remark 11.3 The reader can wonder what become these calculus when P is transient.

This implies that $K'(1,1) < 1$. So the only solution of $\beta(I - K') = 0$ is $\beta = 0$. And then, the only solution of $\pi(I - L') = 0$ is $\pi = 0$.

This implies that $L(1,1) < 1$. So the only solution of $\pi(I - L) = 0$ is $\pi = 0$.

Remark 11.4 Remark that things are far much complicated with $E = \mathbb{Z}$.

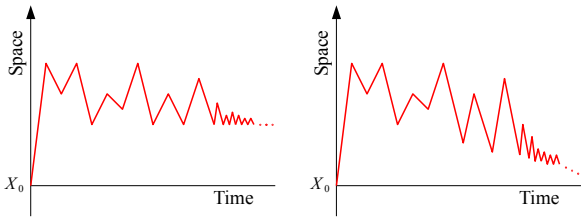
12 Pathologies

12.1 Lack of associativity

In proposition 11.2, we saw a situation where a three time product $\pi.(I-K)(I-L)$ is absolutely convergent, and thus associative. In this sub-section, we see a simple situation where the three time product $(I-K)(I-L)1_E$ is not associative, and we quantify the lack of associativity.

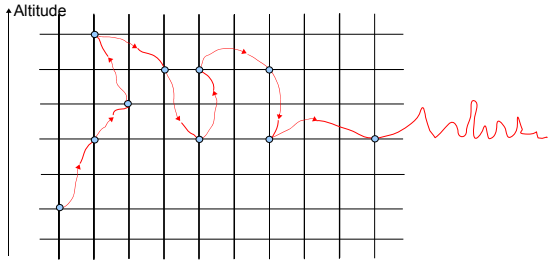
Let us consider the following events:

$$Rejection = \{X_1 \neq \dagger\,,\, X_{[0,\zeta]} \succeq X_0\,,\, \rho_* = +\infty\}$$
$$Rejection' = \{X_1 \neq \dagger\,,\, X_{]0,\zeta]} \succ X_0\,,\, \rho'_* = +\infty\}$$



Rejection' but not *Rejection* — *Rejection'* and *Rejection*

The previous drawing on the left suggests a situation where we have $Rejection'$, $Rejection^c$, but also an infinity of *oplits* (so prime factorization does not exist). In the next drawing we suggest a situation where we have $Rejection'$, $Rejection^c$, and where the process is transient. The state space is $\mathbb{Z}^2$ and we do not write the time axis.



Proposition 12.1 *We have:*

$$k(x) = Lk(x) + \mathbf{P}_x\{X_1 = \dagger\} + \mathbf{P}_x\{Rejection\}$$
$$k'(x) = L'k'(x) + \mathbf{P}_x\{X_1 = \dagger\} + \mathbf{P}_x\{Rejection'\}$$

Proof: Summing over all $y \in E$ in the first equation of proposition 5.4 we find $Lk(x) = \mathbf{P}_x\{X_{[0,\zeta]} \succeq x, \rho_* < \infty\}$ So

$$\begin{aligned} k(x) &= \mathbf{P}_x\{X_{[0,\zeta]} \succeq x\} \\ &= \mathbf{P}_x\{X_1 = \dagger\} + \mathbf{P}_x\{X_1 \neq \dagger, X_{[0,\zeta]} \succeq x, \rho_* < \infty\} + \mathbf{P}_x\{X_1 \neq \dagger, X_{[0,\zeta]} \succeq x, \rho_* < \infty\} \\ &= \mathbf{P}_x\{X_1 = \dagger\} + Lk(x) + \mathbf{P}_x\{Rejection\} \end{aligned}$$

Prime version is similar. □

Finally, when the rejection occurs with positive probability, we have a lack of associativity:

$$[(I - L)(I - K)]1_E(x) = (I - P)1_E(x) = \mathbf{P}_x\{X_1 = \dagger\}$$

while

$$(I - L)[(I - K)1_E](x) = (I - L)k(x) = \mathbf{P}_x\{X_1 = \dagger\} + \mathbf{P}_x\{Rejection\}$$

The same phenomena occurs with the prime factorization when it exists.

Remark 12.2 Let Q be a non-negative matrix. Let h be a Q-excessive function (i.e. $h \geq 0$ and $Qh \leq h$). The classical Riez decomposition of h is $h = U_{[Q]}[(I - Q)h] + Q^\infty h$.

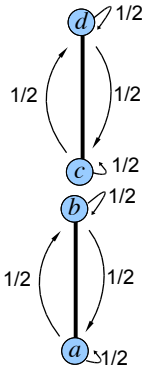
Using the previous proposition 12.1 we can compute the first term of the Riez decomposition of k as a L-excessive function. The second term can be written using the fact that $D_k L$ is the transition matrix of the Markov chain $X_{\text{ϟ}}$:

$$(I - L)k(x) = \mathbf{P}_x\{X_1 = \dagger\} + \mathbf{P}_x\{Rejection\}$$
$$L^\infty k = \mathbf{P}_x\{X_{[0,\zeta]} \succeq x, X_{\text{ϟ}} \text{ never dies}\}$$

12.2 Example of non-unicity: finite case

Our factorizations was performed by some triangular matrices $K, K^\curlyvee, V, V^\curlyvee$ etc. which are defined "automatically" from P. But, it is sometime possible to factorize $I-P$ by triangular matrices which are not our matrices.

Let $E = \{a, b, c, d\}$ with $a \prec b \prec c \prec d$. P is given by the graph or by the array below. For any parameter c_1, c_2 (in $\mathbb{R}$, $\mathbb{C}$ or anything which cancels itself when you multiply it by zero), the matrix $I-P$ can be factorized as following:



$I-P$	d	c	b	a
d	$1/2$	$-1/2$	0	0
c	$-1/2$	$1/2$	0	0
b	0	0	$1/2$	$-1/2$
a	0	0	$-1/2$	$1/2$

$=$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & c_1 & 1 & 0 \\ 0 & c_2 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1/2 & -1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & -1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The choice $c_1 = c_2 = 0$ corresponds to our factorization $(I-P) = (I-L')(I-K')$.

If we apply transposition to both sides of this equation, we obtain a second LU-factorization of $(I-P)$ with 1 on the diagonal of the second term. In this case, the choice $c_1 = c_2 = 0$ corresponds to our factorization $(I-P) = (I-L)(I-K)$.

An other interesting example is $P = \begin{pmatrix} 1 & 0 \\ p & 0 \end{pmatrix}$ with $p \in]0, 1]$. We have

$$\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$$
$$\begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ -p & 1 \end{pmatrix} \qquad \text{impossible for any choices of } *$$

But

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 0 & 0 \\ -p & b \end{pmatrix} \begin{pmatrix} 0 & 0 \\ -p & 1 \end{pmatrix} \qquad \text{possible for any } a, b \text{ such that } -ap + b = 1$$

Remark 12.3 Let us suppose that E is finite and P is irreducible. Let take $F \subsetneq E$. The Markov chain driven by the restricted matrix P_F dies, so the matrix $I_F - P_F$ is invertible (its inverse is $I_F + P_F + P_F^2 + ...$). In other word, all the main minor of the matrix P does not vanish which is the classical criterion for existence and unicity of the LU-factorization. Of course, in example given above, P is not irreducible.

12.3 Non-unicity with an infinite state space

By opposition to the finite case, when E is infinite, the non-unicity can occur even when P is irreducible. Let $E = \mathbb{N}$ and $\preceq$ is $\leq$.

Firstly imagine a situation where the Markov chain converges to $+\infty$, so the process $n \mapsto X_{\tau_n}$ does not die, so $D_{k'}L'$ is stochastic and so $(I - L')k' = 0$. Now we construct the matrix $\Gamma'(x, y) = k'(x)1_{\{y=0\}}$. For all $r \in \mathbb{R}$, we have $K'(x, y) + r\Gamma'(x, y) > 0 \;\Leftrightarrow\; x \succeq y$, so this matrix is triangular and we have

$$(I - L')(I - K' + r\Gamma') = (I - L')(I - K') = (I - P)$$

which gives us an example of non-unicity of the LU-factorization.

13 A trajectorial alternative proof

Actually, all identities we met between matrices P, U, K, V, L, W and prime versions, can been proven and interpreted by trajectorial observations, using mainly the markov property. Some of this proof as quite simple as the proof of $U = VW = V'W'$ (proposition 3.4) or the proof of the mixed factorization $W + K = I + WP$ that we made in sub-section 7.2.

But some trajectorial proofs are quite tricky as the proof of

$$P + L'K' = L' + K'$$

that we produce now.

Recall that, by definition, an *oplit* on z is a time t such that $X_t = z \succ X_0$ and such that $X_{[1,t]} \succeq z$. We write oplit_z the set of oplits on z.

▷ Step one: We fix x and z such that $x \prec z$.

Applying the Markov property at each oplit on z, we have

$$L'(x, z)K'(z, z) = \mathbf{E}_x[((\text{number of oplits on } z) - 1) \vee 0] \tag{31}$$

The -1 is corresponding to the first oplit on z which is skipped. And so

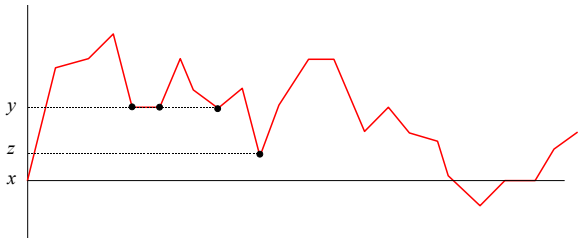
$$\begin{aligned} L'(x, z)K'(z, z) &= \mathbf{E}_x[1_{\{\mathrm{oplit}_z \neq \emptyset\}}\, (\sharp\mathrm{oplit}_z - 1)] \\ &= \mathbf{E}_x[\sharp\mathrm{oplit}_z] - \mathbf{P}_x\{\mathrm{oplit}_z \neq \emptyset\} \end{aligned} \tag{32}$$

If we apply the Markov property at each oplit on y we obtain:

$$\begin{aligned} \forall y \succ z : \qquad L'(x, y)K'(y, z) &= \mathbf{P}_x\{T_z \in [2, \infty[\ ,\ X_{\rho'} \circ X_{[1,T_z[} = y\} \\ &= \mathbf{P}_x\{\mathrm{oplit}_z \neq \emptyset\, ,\, 1 \notin \mathrm{oplit}_z\ ,\ X_{\rho'} \circ X_{[1,T_z[} = y\} \end{aligned}$$

Thus

$$\sum_{y \succ z} L'(x, y)K'(y, z) = \mathbf{P}_x\{\mathrm{oplit}_z \neq \emptyset\ ,\ 1 \notin \mathrm{oplit}_z\} \tag{33}$$



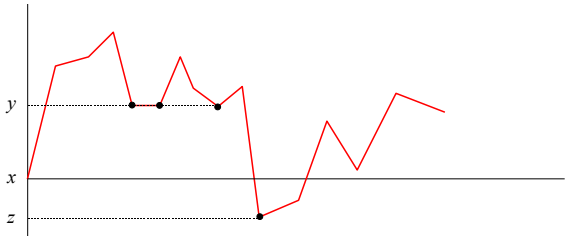
Summing (32) and (33) we get:

$$\begin{aligned} L'K'(x,z) &= \sum_{y \succeq z} L'(x,y)K'(y,z) \\ &= \mathbf{E}_x[\sharp \text{oplit}_z] - \mathbf{P}_x\{1 \in \text{oplit}_z\} \end{aligned}$$

We we remark that $\mathbf{P}_x\{1 \in \text{oplit}_z\} = P(x,z)$. Add to the fact that $K'(x,z) = 0$ we get $P(x,z) + L'K'(x,z) = L'(x,z) + K'(x,z)$.

▷ Step two. We fix x, z such that $x \succeq z$. Applying the Markov property at each oplit on $y \succ x$, we get, as previously:

$$\forall y \succ x : \qquad L'(x,y)K'(y,z) \quad = \quad \mathbf{P}_x\{T_z \in [2,\infty[\ ,\ X_{\rho'} \circ X_{[1,T_z[} = y\} \quad (34)$$



Thus

$$L'K'(x,z) = \sum_{y \succ x} L'(x,y)K'(y,z) = \mathbf{P}_x\{X_{\urcorner'_1} = z\ ,\ \urcorner'_1 \geq 1] = K'(x,z) - P(x,z)$$

Adding to the fact that $L'(x,z) = 0$ we get: $P(x,z) + L'K'(x,z) = K'(x,z) + L'(x,z)$ □

14 General comments

About calculation of LU*-factors:* To have expressions of $L, K, V, W...$ allows us to have informations about our Markov chain. Example: it allows to compute the law of the minimum (proposition 4.4), or to have expressions for the invariant measure (proposition 11.2). But how can we compute these matrices? When E is finite, the classical algorithm of Gauss-elimination (see e.g. [Cia82]) permits

to compute the factors. This algorithm can be immediately "extended" to the case where $E = -\mathbb{N}$ and $\preceq$ is $\leq$; but it cannot be extended when $E = \mathbb{N}$ and $\preceq$ is $\leq$ because it require to start to the upper level. In any cases, this algorithm is recursive, and does not furnish us closed formulae.

On the random walks case, the Wiener-Hopf technique allows to transform our problem into a problem of complex analysis: we have to find functions $\mathfrak{FL}$ and $\mathfrak{FK}$ such that (29) holds. This can be done with some contour integral technique (see [Wid97]). But once again, such expression are not always useful.

When we work with structured Markov chain, many techniques have been invented to access these factors. Most of them are developed in the Li's Book [Li10]. Some of this technique appeal the Wiener-Hopf technique.

About Generalization to the continuous time: Continuous time equivalent to Random walks are Lévy processes. For such processes, the Wiener-Hopf technique was exploited efficiently to compute law linked to min/max, but also to have information about the some infinitesimal behavior of trajectories. (see e.g. Bertoin [Ber96] , Doney [Don01], Kyprianou [Kyp06], Vigon [Vig02],[Vig03]). In continuous time, to find good expressions for factors is really difficult.

About space-time problem: An even harder problem is to obtain expressions for $q \mapsto K_{[qP]}$ and $q \mapsto L_{[qP]}$, which also give some expressions for the law of the time where the global minimum is reached. In continuous times, to compute the space/time Wiener-Hopf factorization is somewhere the Grail of the fluctuation theory for Lévy processes. The main advance in this direction was made by Fourati [Fou10] who shows how to obtain these bivariate factors as solutions of a Riemann-Hilbert problem.

Other Generalization: A Wiener-Hopf factorization have been developed by Williams for Symmetric Markov processes, but, as we said in introduction, this is an alternative generalization which does not correspond to our LU-factorization.

Does LU*-factorization is generalizable for any Markov process?* Yes we think. For right Markov process, the process $t \mapsto X_{\neg_t}$ can be defined as the past-minima process i.e $\inf\{X_s : s \leq t\}$, that we time-change to suppress the horizontal plateau. It is not so hard to see that this process is Markovian. The process $t \mapsto X_{\overline{\neg}_t}$ can be defined as the post-minima process i.e. $\inf\{X_s : s \in [t, \zeta]\}$, that we time-change to suppress horizontal plateau (see Fourati [Fou98] for the precise construction). But even on the Lévy case, this process is not always strongly Markov (see [Fou98]). In this article, Fourati explains how this process is anyway a good Markov process (i.e. Markovian on its natural past σ-field). But the difficulty is to understand how domains of generators of X, $X_{\neg}$ and $X_{\overline{\neg}}$ are linked in order to define properly a LU-factorization.

References

[ASW86] Kevin T. Andrews, Philip W. Smith, and Joseph D. Ward. *LU*-factorization of operators on l_1. *Proc. Amer. Math. Soc.*, 98(2):247–252, 1986.

[AW86] Kevin T. Andrews and Joseph D. Ward. LU-factorization of order bounded operators on Banach sequence spaces. *J. Approx. Theory*, 48(2):169–180, 1986.

[Ber96] Jean Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[BRW80] M. T. Barlow, L. C. G. Rogers, and David Williams. Wiener-Hopf factorization for matrices. In *Seminar on Probability, XIV (Paris, 1978/1979) (French)*, volume 784 of *Lecture Notes in Math.*, pages 324–331. Springer, Berlin, 1980.

[Cia82] Philippe G. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1982.

[Don01] Ronald Doney. Fluctuation theory for Lévy processes. In *Lévy processes*, pages 57–66. Birkhäuser Boston, Boston, MA, 2001.

[Fel66] William Feller. *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons Inc., New York, 1966.

[Fou98] S. Fourati. Points de croissance des processus de Lévy et théorie générale des processus. *Probab. Theory Related Fields*, 110(1):13–49, 1998.

[Fou10] Sonia Fourati. Fluctuations of Lévy processes and scattering theory. *Trans. Amer. Math. Soc.*, 362(1):441–475, 2010.

[FP62] Miroslav Fiedler and Vlastimil Pták. On matrices with non-positive off-diagonal elements and positive principal minors. *Czechoslovak Math. J.*, 12 (87):382–400, 1962.

[FP81] R. E. Funderlic and R. J. Plemmons. LU decomposition of M-matrices by elimination without pivoting. *Linear Algebra Appl.*, 41:99–110, 1981.

[Gra87] Winfried Grassmann. Means and variances of time averages in Markovian environments. *European J. Oper. Res.*, 31(1):132–139, 1987.

[Hey95] Daniel P. Heyman. A decomposition theorem for infinite stochastic matrices. *J. Appl. Probab.*, 32(4):893–901, 1995.

[KSK66] John G. Kemeny, J. Laurie Snell, and Anthony W. Knapp. *Denumerable Markov chains*. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, 1966.

[Kuo77] I Wen Kuo. A note on factorizations of singular M-matrices. *Linear Algebra and Appl.*, 16(3):217–220, 1977.

[Kyp06] Andreas E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Universitext. Springer-Verlag, Berlin, 2006.

[LC04] Quan-Lin Li and Jinhua Cao. Two types of RG-factorizations of quasi-birth-and-death processes and their applications to stochastic integral functionals. *Stoch. Models*, 20(3):299–340, 2004.

[Li10] Quan-Lin Li. *Constructive computation in stochastic models with applications*. Tsinghua University Press, Beijing, 2010. The RG-factorization.

[LZ02] Quan-Lin Li and Yiqiang Zhao. A constructive method for finding β-invariant measures for transition matrices of $M/G/1$ type. In *Matrix-analytic methods (Adelaide, 2002)*, pages 237–263. World Sci. Publ., River Edge, NJ, 2002.

[LZ04] Quan-Lin Li and Yiqiang Q. Zhao. A $MAP/G/1$ queue with negative customers. *Queueing Syst.*, 47(1-2):5–43, 2004.

[MS98] J. J. McDonald and H. Schneider. Block LU factorizations of M-matrices. *Numer. Math.*, 80(1):109–130, 1998.

[Sen73] E. Seneta. *Non-negative matrices*. Halsted Press [A division of John Wiley & Sons], New York, 1973. An introduction to theory and applications.

[VC82] Richard S. Varga and Da Yong Cai. On the LU factorization of M-matrices. *Numer. Math.*, 38(2):179–192, 1981/82.

[Vig02] 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.

[Vig03] V. Vigon. Comparaison des deux composantes d'un subordinateur bivarié, puis étude de l'enveloppe supérieure d'un processus de Lévy. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(6):993–1011, 2003.

[Vig11] V. Vigon. (homogeneous) markovian bridges. *To appears in the anale of IHP*, 2011.

[Wid97] Harold Widom. Wiener-Hopf integral equations. In *The Legacy of Norbert Wiener: A Centennial Symposium (Cambridge, MA, 1994)*, volume 60 of *Proc. Sympos. Pure Math.*, pages 391–405. Amer. Math. Soc., Providence, RI, 1997.

[Wil84] W. E. Williams. Recognition of some readily "Wiener-Hopf" factorizable matrices. *IMA J. Appl. Math.*, 32(1-3):367–378, 1984.

[Wil91] David Williams. Some aspects of Wiener-Hopf factorization. *Philos. Trans. Roy. Soc. London Ser. A*, 335(1639):593–608, 1991.

[Wil08] David Williams. A new look at 'Markovian' Wiener-Hopf theory. In *Séminaire de probabilités XLI*, volume 1934 of *Lecture Notes in Math.*, pages 349–369. Springer, Berlin, 2008.

[Woe00] Wolfgang Woess. *Random walks on infinite graphs and groups*, volume 138 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2000.

[ZLB97] Yiqiang Q. Zhao, Wei Li, and W. John Braun. On a decomposition for infinite transition matrices. *Queueing Systems Theory Appl.*, 27(1-2):127–130, 1997.