Electronic Journal of Probability

Electron. J. Probab. **28** (2023), article no. 156, 1–44.
ISSN: 1083-6489 https://doi.org/10.1214/23-EJP1052

Almost triangular Markov chains on $\mathbb{N}$

Luis Fredes* Jean-François Marckert†

Abstract

A transition matrix $\left[\mathsf{U}_{i,j}\right]_{i,j\geq 0}$ on $\mathbb{N}$ is said to be almost upper triangular if $\mathsf{U}_{i,j} > 0 \Rightarrow j \geq i-1$, so that the increments of the corresponding Markov chains are at least -1; a transition matrix $\left[\mathsf{L}_{i,j}\right]_{i,j\geq 0}$ is said to be almost lower triangular if $\mathsf{L}_{i,j} > 0 \Rightarrow j \leq i+1$, and then, the increments of the corresponding Markov chains are at most $+1$.

In the present paper, we characterize the recurrence, positive recurrence and invariant distribution for almost triangular transition matrices class. The upper case appears to be the simplest in many ways, with existence and uniqueness of invariant measures, when in the lower case, existence, as well as uniqueness, are not guaranteed. We present the time-reversal connection between upper and lower almost triangular transition matrices, which provides classes of integrable lower triangular transition matrices.

These results encompass the case of birth and death processes (BDP) that are famous (discrete or continuous time) Markov processes taking their values in $\mathbb{N}$, which are simultaneously almost upper and almost lower triangular, and whose study has been initiated by Karlin & McGregor in the 1950's. They found invariant measures, criteria for recurrence, null recurrence, among others; their approach relies on some profound connections they discovered between the theory of BDP, the spectral properties of their transition matrices, the moment problem, and the theory of orthogonal polynomials. Our approach is mainly combinatorial and uses elementary algebraic methods; it is somehow more direct and does not use the same tools.

Keywords: Markov chains.
MSC2020 subject classifications: 60J10.
Submitted to EJP on October 28, 2022, final version accepted on October 31, 2023.

Contents

*Univ. Bordeaux, CNRS, Bordeaux INP, IMB, UMR 5251, F-33400 Talence, France. E-mail: luis.fredes@u-bordeaux.fr

†Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, F-33400 Talence, France. E-mail: marckert@labri.fr

1 Introduction

Notation and conventions

The set of non-negative integers is denoted $\mathbb{N}$. For two integers $a < b$, $[a, b]$ will be the set of integers $\{a, a + 1, \ldots, b\}$. The word “interval” will be added when standard real intervals are considered.

A transition matrix over a finite or countable state space S is a matrix $\left[\mathsf{M}_{i,j}\right]_{i,j\in S}$ indexed by S, such that the $\mathsf{M}_{i,j}$ are non-negative real numbers and sum to one on each row of the matrix.

The set of non-negative measures over S (equipped with the power set sigma field), with total mass being strictly positive, or infinite, is written $\mathcal{M}_S^+$.

A measure $\pi \in \mathcal{M}_S^+$ is said to be invariant by M if

$$\sum_{a\in S} \pi_a \mathsf{M}_{a,b} = \pi_b, \quad \text{for all } \ b \in S. \tag{1.1}$$

Often, we will see π as a row matrix $\pi := \left[\pi_x, x \in S\right]$, and (1.1) will be written $\pi\mathsf{M} = \pi$.

We denote by $\mathcal{M}_S^{+,\sim}$ the set of equivalence classes of positive measures consisting of those which are equal up to a positive factor. Since the invariance by M is a class property, we will say that a transition matrix M has a single (resp. several) invariant measure in $\mathcal{M}_S^{+,\sim}$, when there is a single (resp. several) class of invariant measures.

The adjectives “recurrent”, “irreducible”, “aperiodic”, and “positive recurrent” will qualify indifferently transition matrices and Markov chains.

For any subset F of S, M_F is the matrix obtained by keeping only the rows and columns of M indexed by elements of F (that is $\mathsf{M}_F = \left[\mathsf{M}_{i,j}\right]_{i,j\in F}$).

The identity matrix is denoted Id, and this, whatever its size is (which will be however clear from the context). For example, we will simply write $(\mathsf{Id}-\mathsf{M})_F = \mathsf{Id}-\mathsf{M}_F$, without adding the precision that in the left-hand side, Id has size the cardinality of S, and in the right-hand size, that of F.

A transition matrix $\mathsf{U} = \left[\mathsf{U}_{i,j}\right]_{0\le i,j\le+\infty}$ on $\mathbb{N}$ is said to be almost upper-triangular (◹) if

$$\mathsf{U}_{i,j} > 0 \Rightarrow j \ge i-1 \tag{1.2}$$

and a transition matrix $\mathsf{L} = \left[\mathsf{L}_{i,j}\right]_{0\le i,j\le+\infty}$ is said to be almost lower-triangular (◺) if

$$\mathsf{L}_{i,j} > 0 \Rightarrow j \le i+1. \tag{1.3}$$

Here are some “pictures” explaining the iconographic notation ◹ and ◺ of the main objects:

$$\mathsf{U} := \begin{bmatrix} \mathsf{U}_{0,0} & \mathsf{U}_{0,1} & \mathsf{U}_{0,2} & \mathsf{U}_{0,3} & \mathsf{U}_{0,4} & \cdots \\ \mathsf{U}_{1,0} & \mathsf{U}_{1,1} & \mathsf{U}_{1,2} & \mathsf{U}_{1,3} & \mathsf{U}_{1,4} & \cdots \\ 0 & \mathsf{U}_{2,1} & \mathsf{U}_{2,2} & \mathsf{U}_{2,3} & \mathsf{U}_{2,4} & \cdots \\ 0 & 0 & \mathsf{U}_{3,2} & \mathsf{U}_{3,3} & \mathsf{U}_{3,4} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \cdots \end{bmatrix}, \quad \mathsf{L} := \begin{bmatrix} \mathsf{L}_{0,0} & \mathsf{L}_{0,1} & 0 & 0 & 0 & \cdots \\ \mathsf{L}_{1,0} & \mathsf{L}_{1,1} & \mathsf{L}_{1,2} & 0 & 0 & \cdots \\ \mathsf{L}_{2,0} & \mathsf{L}_{2,1} & \mathsf{L}_{2,2} & \mathsf{L}_{2,3} & 0 & \cdots \\ \mathsf{L}_{3,0} & \mathsf{L}_{3,1} & \mathsf{L}_{3,2} & \mathsf{L}_{3,3} & \mathsf{L}_{3,4} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \cdots \end{bmatrix}$$

This paper aims to provide a first systematic study of Markov chains following a ◹ or a ◺ transition matrix. Our results will appear to generalize birth and death (BD) processes results. These latter form the famous model of Markov chains on $\mathbb{N}$ having tridiagonal transition matrices, often represented as

$$\mathsf{T} = \begin{bmatrix} r_0 & p_0 & 0 & 0 & \cdots & \\ q_1 & r_1 & p_1 & 0 & 0 & \cdots \\ 0 & q_2 & r_2 & p_2 & 0 & \cdots \\ 0 & 0 & q_3 & r_3 & p_3 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \tag{1.4}$$

where $q_i = \mathsf{T}_{i,i-1}, r_i = \mathsf{T}_{i,i}, p_i = \mathsf{T}_{i,i+1}$. Tridiagonal means that $\mathsf{T}_{i,j} > 0 \Rightarrow |i-j| \le 1$, and then, this is the class of transition matrices that are simultaneously ◹ and ◺: the increments of a chain with such a transition matrix belong to $\{-1,+1,0\}$.

Two very influential papers published in 1956 by Karlin & McGregor [13, 12] showed that the main characteristics of BD models can be exactly computed, and this is a consequence of some particular features of the algebra that comes into play.

We review some of the results they obtained here. Consider T a tridiagonal transition matrix and ℓ the maximal l with the following property: $\mathsf{T}_{i,i+1}$ and $\mathsf{T}_{i+1,i}$ are positive for all $0 \le i \le l$ with $l \in \{0,1,\dots,\} \cup \{+\infty\}$. If ℓ is finite, a Markov chain with transition matrix M is irreducible when restricted to the finite state space $[0,\ell]$. This case can be studied using the finite Markov chains tools (even if BDP on compact sets are interesting from the combinatorial point of view on their own, see Flajolet & Guillemin [7]). Karlin & McGregor results concern the irreducible case over $\mathbb{N}$ (case $\ell = +\infty$). In this case, they establish the following results: The Markov chain is reversible ($\pi_{k-1}\mathsf{T}_{k-1,k} = \pi_k \mathsf{T}_{k,k-1}$) with respect to the measure

$$\pi_0 = 1, \text{ and for } a \geq 1, \quad \pi_a = \frac{p_0 \cdots p_{a-1}}{q_1 \cdots q_a} = \prod_{j=1}^{a} \frac{\mathsf{T}_{j-1,j}}{\mathsf{T}_{j,j-1}}, \tag{1.5}$$

so that this measure is invariant by T. Moreover, this measure is the unique invariant measure for T (in $\mathcal{M}_{\mathbb{N}}^{+,\sim}$). There exists an invariant probability distribution if and only if

$$\sum_{k \geq 1} \prod_{j=1}^{k} \frac{\mathsf{T}_{j-1,j}}{\mathsf{T}_{j,j-1}} < +\infty, \tag{1.6}$$

and then, this is also a necessary and sufficient condition for positive recurrence. A Markov chain with transition matrix T is recurrent if and only if

$$\sum_{k \geq 0} \prod_{j=1}^{k} \frac{\mathsf{T}_{j,j-1}}{\mathsf{T}_{j,j+1}} = +\infty. \tag{1.7}$$

There is also a connection with the theory of orthogonal polynomials (see Section A.2 in which the connection established by Karlin & McGregor is discussed; we refer to Schoutens [21] for more information on the connections between probability and orthogonal polynomials theories). In their papers, Karlin & McGregor study mainly the continuous-time version of these Markov processes, whose behaviour is similar (up to a random time change) to the discrete version preferred here. The continuous setting is important in their study since differentiation with respect to time of some quantities is considered at many steps of their study. This is not the case in our approach, and we prefer to stick with the discrete-time which makes more natural the use of combinatorial tools: we will discuss the continuous case in Section A.1 only.

1.1 Main results and contents of the paper

The nature and behaviour of almost upper or lower triangular chains are different from those of BDP. The first difference is that if M is ◹ (or ◺) but not tridiagonal, then M is not reversible with respect to any positive measure π. The reason is that for such an M, there exists a pair of indices (a, b) such that $\mathsf{M}_{a,b} > 0$ and $\mathsf{M}_{b,a} = 0$, which implies that $\pi_a \mathsf{M}_{a,b} = \pi_b \mathsf{M}_{b,a}$ is not possible (when M is irreducible).

In this paper, we will add the irreducibility hypothesis virtually everywhere: consider the strongly connected components in the graph with vertex set $V = \mathbb{N}$ and directed edge set $E = \{(i, j), \mathsf{M}_{i,j} > 0\}$ for M being either ◹ or ◺. It is easy to see that the ◹ (resp. ◺) structure imposes all strongly connected components to be intervals of $\mathbb{N}$, since the only down-steps are -1 (resp. the only up steps are $+1$). Hence, there is at most one infinite connected component (which is, in this case, equivalent, up to change of origin to $[0, +\infty)$), and the Markov chain on each finite component reduces to the study of a Markov chain over a finite state space (we refer to [18, 19] for more information on Markov chain techniques). Hence, requiring irreducibility in this setting is natural.

Convention Unless otherwise stated, all the Markov chains considered will be irreducible on $\mathbb{N}$.

Let us recall a fact concerning finite irreducible Markov chains which is important to have in mind before starting the description of our results. Let $\mathsf{M} = \left[\mathsf{M}_{a,b}\right]_{0\le a,b\le N}$ be an irreducible transition matrix over $[0,N]$ for some finite N. Such a Markov chain is positive recurrent, and, by the Perron-Frobenius theorem (see for example [16]), M possesses a unique invariant probability distribution in $\mathcal{M}^{+}_{[0,N]}$ which is

$$\rho_k = \det\left(\mathsf{Id} - \mathsf{M}^{\mathsf{dep}(k)}\right) / \, \alpha_N, \text{ for } k \in [0,N] \tag{1.8}$$

where:

- for a matrix A, the notation $A^{\mathsf{dep}(k)}$ represents the matrix A deprived of its k-th column and row,
- α_N is the only normalising constant making of ρ a probability distribution.

As a consequence of the matrix tree theorem (see e.g. [25]), or as a consequence of the properties of the Markov chain tree theorem ([24, 5, 2, 10, 8]), for any vertex r of G, $\det\left(\mathsf{Id} - \mathsf{M}^{\mathsf{dep}(r)}\right)$ is the total weight of the set $\mathsf{SpanningTrees}^{\bullet}(r)$ of rooted spanning trees (t, r) of the oriented graph $G = (V,E)$ where $V = [0,N]$, and $E = \{(a,b) : \mathsf{M}_{a,b} > 0\}$. More precisely it states that each spanning tree (t, r) is an oriented graph $(V^{\mathsf{t},r}, E^{\mathsf{t},r})$, whose edges $e = (u,v) \in E^{\mathsf{t},r}$ are oriented towards the root r and the weight of this rooted spanning tree is defined as

$$\mathsf{Weight}(\mathsf{t}, r) = \prod_{e=(u,v)\in E^{\mathsf{t},r}} \mathsf{M}_{u,v}, \tag{1.9}$$

and the matrix tree theorem in these settings writes as

$$\det\left(\mathsf{Id} - \mathsf{M}^{\mathsf{dep}(r)}\right) = \sum_{(\mathsf{t},r)\in \mathsf{SpanningTrees}^{\bullet}(r)} \mathsf{Weight}(\mathsf{t}, r). \tag{1.10}$$

1.2 Content of the paper

Section 2.1 collects the main results concerning ◹ transition matrices

In Theorem 2.1 we establish that each irreducible ◹ transition matrix U has a single invariant measure $(\pi_a, a \ge 0)$ in $\mathcal{M}^{+,\sim}_{\mathbb{N}}$ where $\pi_0 > 0$ and

$$\pi_a = \pi_0 \, \frac{\det(\mathsf{Id} - \mathsf{U}_{[0,a-1]})}{\prod_{j=1}^{a} \mathsf{U}_{j,j-1}}, \quad a \ge 1,$$

which provides a characterization for positive recurrence ($\sum_a \pi_a < +\infty$). A necessary and sufficient condition for recurrence is given in Theorem 2.8: $\lim_{b\to\infty} \mathsf{U}_{1,0} \frac{\det(\mathsf{Id}-\mathsf{U}_{[2,b-1]})}{\det(\mathsf{Id}-\mathsf{U}_{[1,b-1]})}$ $= 1$. This is done thanks to some explicit formulas for the distribution of the hitting time $\tau_S(Y)$ of a set S by a Markov chain $(Y_i, i \ge 0)$ with transition matrix U (Theorem 2.8 and Proposition 2.9), notably, it is established that $\mathbb{P}(\tau_{\{0\}}(Y) < \tau_{[b,+\infty)}(Y) \mid Y_0 = 1) = \mathsf{U}_{1,0} \frac{\det(\mathsf{Id}-\mathsf{U}_{[2,b-1]})}{\det(\mathsf{Id}-\mathsf{U}_{[1,b-1]})}$.

In Remark 2.11 we will abandon locally the irreducibility assumption to treat the case where absorption at 0 occurs with a positive probability at each passage, and we compute the probability of eventual absorption.

For each $n \ge 0$, the projection transition matrix $\mathsf{U}^{(n)}$ is defined by restricting U to $[0,n]$ (up to some boundary details, see (2.11)). Since $\mathsf{U}^{(n)}$ is a finite state space transition matrix, when irreducible, it possesses a unique invariant distribution $\rho^{(n)}$. In Theorem 2.13, the convergence of $\rho^{(n)}$ (after normalisation if necessary) to the unique invariant measure π of U in $\mathcal{M}^{+,\sim}_{\mathbb{N}}$ is established.

Last, in Proposition 2.3, it is established that some irreducible ◹ transition matrices U have several non-proportional right eigenvectors associated with the eigenvalue 1; the subspace generated by the eigenvectors with positive coordinates may have any finite dimension, and even, can be infinite-dimensional.

Section 2.2 collects the main results concerning ◺ transition matrices

First of all, the class of ◺ transition matrices appears to be more complex to study than ◹ transition matrices. Probably, the simplest explanation is the role played more or less directly by the invariant measures π in the study of a ◺ transition matrix L. In the ◺ case, an invariant measure is a solution of $\pi_a = \sum_{b:b\geq a-1} \pi_b \mathsf{L}_{b,a}$, so that it relates π_{a-1} with an infinite number of π_b with larger indices, while in the ◹ case, $\pi_a = \sum_{b:b\leq a+1} \pi_b \mathsf{U}_{b,a}$ allows expressing π_{a+1} with the π_b with smaller indices: a triangular system, easy to solve, with a unique solution.

Hence, in the ◺ case, neither uniqueness nor existence of invariant measures is guaranteed (Theorem 2.14), and the cone of invariant measures in the ◺ case can have any dimension, ranging from 0 to $+\infty$. In Theorem 2.15, we give a characterization of ◺ transition matrices L that are recurrent (the condition is $\lim_{b\to+\infty} \frac{\prod_{j=1}^{b-1} \mathsf{L}_{j,j+1}}{\det(\mathsf{Id}-\mathsf{L}_{[1,b-1]})} = 0$). This is done also by the study of the distribution of the hitting times $\tau_{\{j\}}(Y)$ of a Markov chain with transition matrix L. The case where absorption at 0 occurs with a certain probability is discussed in Remark 2.18, and the probability of absorption is computed. In Theorem 2.19, it is shown that the invariant distribution of a ◺ transition matrix L on the finite set $[0,s]$, is proportional to $(\eta_a, a \in [0,s])$ where $\eta_a = \eta_0 \det\left(\mathsf{Id} - \mathsf{L}_{[a+1,s]}\right) \prod_{i=1}^{a} \mathsf{L}_{i-1,i}$. This result allows stating Proposition 2.20 and Proposition 2.21 which provide some conditions for the convergence of the (rescaled) invariant distribution $\eta^{(m)}$ of the projected transition matrix $\mathsf{L}^{(m)}$ to an invariant measure η of L.

Connections between ◺ and ◹ transition matrices

The time-reversal of a trajectory with jumps bounded from above by 1, is a trajectory with jumps bounded from below by $-1\dots$ so it is tempting to guess that the time-reversal of a ◹ Markov chain is a ◺ Markov chain, and vice-versa (under their stationary regime). It turns out that the complete picture is more complex than that because ◹ transition matrices have a single invariant measure, when the existence and uniqueness of invariant measures is not guaranteed in the ◺ case. Hence:

- time-reversal of ◹ transition matrices are ◺ transition matrices,
- time-reversal to ◺ transition matrices, may exist or not, and in the case where L possesses several invariant measures, several time-reversals of L can be defined, all being ◹ transition matrices (see Theorems 3.1 and 3.3). Recurrence and positive recurrence of any associated time-reversal transition matrices are shown to be equivalent to those of L.

The spectral properties of ◹ and ◺ matrices are quite rich, and they are scattered in various theorems (already cited above): Theorem 3.8 states that each complex number is a simple left-eigenvalue of all infinite ◹ transition matrices, and also simple right-eigenvalues of infinite ◺ transition matrices: eigenvectors are totally explicit. For ◹ and ◺ transition matrices U and L that are time-reversal of each other, for each Λ, there is an explicit linear map which sends the right eigenspace of U associated with the eigenvalue Λ (the set $\{v : \mathsf{U}v = \Lambda v\}$ possibly reduced to $\{0\}$), to the left eigenspace of L.

Sub-stochastic almost triangular matrices have some comparable properties to stochastic almost triangular matrices in terms of spectral properties (Proposition 3.11 and Proposition 3.12).

Since ◺ transition matrices are in general more difficult to study than ◹ transition matrices, finding the time-reversal U of a transition matrix L provides at once an important tool to study the behaviour of L-Markov chains. We then provide some results allowing one to better understand the algebraic relation between pairs (L, U), time-reversal of each other with respect to some measures (Proposition 3.5, and Remark 3.6).

In Section 3.3, we will make a slight change in the representation of ◺ transition matrices L using the so-called descent kernel

$$\begin{cases} \mathsf{L}_{b,a} &= v_b\, \mathsf{D}_{b,a}, \text{ for } b \geq a, \\ \mathsf{L}_{b,b+1} &= 1 - v_b, \quad b \geq 0. \end{cases}$$

Hence, v_b is the probability to "go down" from level b, and $\mathsf{D}_{b,a}$ is the probability to descent from b to a, when the "go down" direction is chosen. This representation, of course, equivalent to the initial representation of ◺ transition matrices, provides some different formulas for the researched time-reversal transition matrix U (which involves D too, see Proposition 3.14).

In Section 3.3.2, we will change a bit of perspective – fix D but let $(v_i, i \geq 0)$ be freely chosen: this will provide a way to construct many integrable ◺ transition matrices L (Theorem 3.16), meaning that the invariant measure associated to these L can be expressed in the form of a closed formula. This point of view is reminiscent of "catastrophe transition matrices" in which the descent transition matrix is the important feature of the model. This allows us to revisit some known results of the literature (see Section 3.3.3).

In Section 4.1 we show that our results are equivalent to the results of Karlin & McGregor in the tridiagonal case (our formulas use determinants when it is not the case for those of Karlin & McGregor so that proofs are needed).

In Section 4.2, we provide a family of integrable ◺ transition matrices: in words, when the columns of L are almost proportional (see Definition 4.1), then the system which allows computing the invariant distribution is triangular (in some sense), and then can be solved.

In Section 4.3, another family of integrable models is given: these are some models of ◺ and ◹ that can be expressed in terms of birth-death processes decomposed between some stopping times.

In Section 4.4, a fourth list of integrable models, called the repair shop Markov chain, is revisited, and treated with our main theorems (criteria of recurrence and positive recurrence are found using new methods).

Section A.1 is devoted to continuous-time counterparts of our models of ◺ and ◹ Markov chains.

Finally, since many proofs we give use combinatorial facts (notably matrix tree theorem and heap of pieces techniques), Section 1.3 recalls these tools.

1.3 Tools for the proofs of the main theorems

About the determinants of almost triangular matrices

Lemma 1.1. *Let $A = \left[A_{i,j}\right]_{0 \leq i,j \leq N}$ be a <u>finite</u> ◹ matrix (transition matrix or not, with complex coefficients). Denote by S^N the set of increasing integer valued sequences $s = (s_0, \ldots, s_k)$ with $1 \leq k \leq N$, $s_0 = -1$ and $s_k = N$. For such a sequence denote by*

$\ell(s) = k$ *its final index. We have*

$$\det(A) = \sum_{j=0}^{N} A_{0,j} \left(\prod_{i=1}^{j} (-A_{i,i-1}) \right) \det(A_{[j+1,N]}). \tag{1.11}$$

As a consequence,

$$\det(A) = \sum_{s \in S^N} \left(\prod_{j=1}^{\ell(s)} A_{s_{j-1}+1, s_j} \right) \prod_{j \in [0,N-1] \setminus s} (-A_{j+1,j}) \tag{1.12}$$

where $[0, N-1] \setminus s$ *is the set obtained by removing the elements of* s *from the set* $[0, N]$*.*

Remark 1.2. *If* A *is a finite* ◺ *matrix, then it is immediate by transposition that*

$$\det(A) = \sum_{s \in S^N} \left(\prod_{j=1}^{\ell(s)} A_{s_j, s_{j-1}+1} \right) \prod_{j \in [0,N-1] \setminus s} (-A_{j,j+1}) \,.$$

Proof. Expand the determinant along the first row: $\det(A) = \sum_{j=0}^{N} (-1)^j A_{0,j} \det(A^{\mathsf{dep}(0,j)})$ where the matrix $A^{\mathsf{dep}(0,j)}$ is obtained by removing row 0 and column j of A. Now, the conclusion follows from the fact that in the matrix $H := A^{\mathsf{dep}(0,j)}$ the j first columns have non-zero entries only above the diagonal, i.e. $(H_{a,b} > 0, b < j) \Rightarrow a \geq b$. Hence when one expands the determinant, to get a non-zero result, the diagonal entries of the first j columns must be selected, and then they are multiplied by $\det(A_{[j+1,N]})$, this concludes (1.11). Formula (1.12) is proved by recursively applying (1.11) on j for $\det(A_{[j,N]})$. $\square$

1.3.1 The matrix tree theorem and related facts

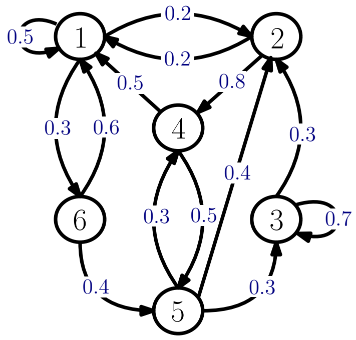
Let (G, W) be a weighted oriented graph, where $G := (V, E)$ is a graph, V is the set of nodes and $E \subset V^2$ is the set of edges. As usual, the (oriented) edge $e = (u, v)$ is oriented toward v, and its weight is $W_{u,v}$. The matrix tree theorem asserts that

$$\sum_{(\mathsf{t},r) \in \mathsf{SpanningTrees}^\bullet(r)} \prod_{e \in E^{\mathsf{t}}} W_e = \det\left(\mathsf{Laplacian}(W)^{\mathsf{dep}(r)} \right),$$

where each edge of the tree (t, r) is oriented toward the root r, and where $\mathsf{Laplacian}(W)^{\mathsf{dep}(r)}$ is the Laplacian matrix of W, in which the rth row and column have been removed. When $W_{u,v} = \mathsf{M}_{u,v}$ for a transition matrix M, $\mathsf{Laplacian}(\mathsf{M}) = (\mathsf{Id} - \mathsf{M})$ (this is equivalent to (1.10)).

Definition 1.3. *Let* $\mathsf{Roots} \subset V$ *be a set of roots, and* $\mathsf{Nodes} \subset V$ *a set of nodes ("of other nodes", we should say). We denote by* $\mathsf{Forests}(\mathsf{Nodes}, \mathsf{Roots})$ *the set of forests, a forest being a sequence of rooted trees* $\{(\mathsf{t}_1, r_1), \ldots, (\mathsf{t}_k, r_k)\}$*, satisfying the following constraints:*

- *at least one tree* $(k \geq 1)$*, and each tree is not reduced to its root,*
- *the set of nodes* $V(\mathsf{t}_i)$ *of the* t_i *are disjoint, and all of them are included in* $\mathsf{Nodes} \cup \mathsf{Roots}$*,*
- *each element of* Nodes *belongs to (exactly) a single tree,*
- *the set of roots* $\{r_1, \ldots, r_k\}$ *is a subset of* Roots *(the* r_i *are distinct),*
- *each tree* t_i *has* r_i *as the unique vertex belonging to* Roots*.*



$$\mathsf{Laplacian}(M) = \begin{pmatrix} 0.5 & -0.2 & 0 & 0 & 0 & -0.3 \\ -0.2 & 1 & 0 & -0.8 & 0 & 0 \\ 0 & -0.3 & 0.3 & 0 & 0 & 0 \\ -0.5 & 0 & 0 & 1 & -0.5 & 0 \\ 0 & -0.4 & -0.3 & -0.3 & 1 & 0 \\ -0.6 & 0 & 0 & 0 & -0.4 & 1 \end{pmatrix}$$

Figure 1: Example for the Matrix tree theorem. Here matrices rows and columns are indexed from 1 to 6 as the vertices of the graph. To the left we give the description of a Markov transition matrix as the weights of the (oriented) edges and to the right the Laplacian matrix associated to this Markov chain. In gray the column and row suppressed to compute the total weight of spanning trees rooted at vertex 1, which is equal to $\det\left(\mathsf{Laplacian}(M)^{\mathsf{dep}(1)}\right) = 0.171$.

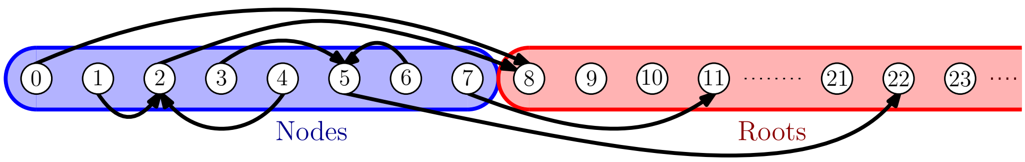


Figure 2: Representation of forests as in Definition 1.3 where $\mathsf{Nodes} = [0,7]$ and $\mathsf{Roots} = [8,+\infty)$. The figure illustrates a forest with three trees (not reduced to their root); their sets of vertices are: $\{0,1,2,4,8\}$, $\{3,5,6,22\}$ and $\{7,11\}$.

For any forest $\mathsf{f} = \{(\mathsf{t}_1, r_1), \dots, (\mathsf{t}_k, r_k)\}$, set

$$\mathsf{Weight}_W(\mathsf{f}) = \prod_{i=1}^{k} \mathsf{Weight}(\mathsf{t}_i, r_i),$$

where $\mathsf{Weight}(\mathsf{t}_i, r_i)$ is as in (1.9) with W in place of M (the edges of each tree are oriented toward their root).

Proposition 1.4. *Consider the graph $G = (\mathbb{N}, \{(i,j) : \mathsf{U}_{i,j} > 0\})$, weighted by the transition matrix U, that is $W = \mathsf{U}$. We have*

$$\sum_{F \in \mathsf{Forest}([0,x-1],[x,+\infty))} \mathsf{Weight}_{\mathsf{U}}(F) = \det((\mathsf{Id} - \mathsf{U})_{[0,x-1]}). \tag{1.13}$$

Proof. This can be viewed as a consequence of the matrix tree theorem in which the set Roots is identified with one node. $\square$

1.3.2 Heap of cycles

We recall some aspects of the theory of heaps of pieces [23, 14], and more specifically heap of cycles, which will be a useful tool to prove some of our results.

Consider M a transition matrix on a finite or infinite countable graph $G = (V, E)$, meaning that $\mathsf{M} = (\mathsf{M}_{u,v}, u, v \in V)$, $\mathsf{M}_{u,v} > 0 \Rightarrow \{u,v\} \in E$, and as usual, for all $u \in V$, $\sum_{v \in V} \mathsf{M}_{u,v} = 1$.

Attribute to each (oriented) path $w = (w_0, \dots, w_{|w|})$ on G, the weight

$$\mathsf{Weight}(w) = \prod_{j=1}^{|w|} \mathsf{M}_{w_{j-1},w_j}.$$

A path w is a cycle if $w_{|w|} = w_0$ and if moreover, for all $0 \le i < j < |w|$, $w_i \ne w_j$ then, it is called a simple cycle. We extend the map Weight to collections of paths $C := (w(1), \dots, w(|C|))$ in which case we set

$$\mathsf{Weight}(C) = \prod_{j=1}^{|C|} \mathsf{Weight}(w(i)). \tag{1.14}$$

Path decomposition

A standard result from combinatorics which has proved its importance notably in the study of loop-erased random walks (see e.g. Tyler [9, Appendix A], Lawler [15], Wilson [24], Marchal [17]), is that

Lemma 1.5. *There exists a weight-preserving bijective map that sends the set of paths on G starting at some point v onto the set of pairs $(\mathsf{saw}, \mathsf{hc})$ where saw is a self-avoiding walk on G starting at v, and hc is a heap of cycles with maximal pieces incident to saw.*

(The notion of maximal pieces is recalled below.) See e.g. Prop. 6.3. in Viennot [23] for additional details (and proof). A self-avoiding path is a path w such that $w_i = w_j \Rightarrow i = j$. A heap of cycles, is a particular instance of the concept of heap of pieces, an important combinatorial concept. We refer to Viennot [23], Krattenthaler [14], Cartier & Foata [6], Zeilberger [25], to Tyler [9, Appendix A] for details, and just recall some aspects below.

A heap of pieces is, informally, a collection of pieces, that are placed on a discrete space ($E \times \mathbb{N}$, where E is a set of elements, and $\mathbb{N}$ is the height space). The definition uses a reflexive and symmetric relation $\mathcal{R}$ on the set of pieces E. Some pieces are said to be in relation, which implies that they cannot be placed at the same height (if $p\mathcal{R}p'$, then (p, i) and (p', i) cannot belong to the same heap); moreover, a piece p at height i with $i \ge 1$ must be supported by a piece p at height $i - 1$, which is related to it (that is, if (p, i) is in a heap H and $i \ge 1$, then H must contain $(p', i - 1)$ with $p'\mathcal{R}p$).

There are several ways to define formally the notion of heap of pieces:

- as an element of a partially commutative monoid: if this point of view is adopted, a heap of pieces is a word $w_1 \dots w_m$, where the letters w_i belong to E, and in which pairs of non-related letters with respect to $\mathcal{R}$ commute (Cartier & Foata [6]),
- more geometrically (Viennot [23]), in which heaps of pieces H are viewed as finite sets of pairs $\{(p, i) : p \in E, i \in \mathbb{N}\}$, such that
  1. If $(p, i), (p', j) \in H$ and $p\mathcal{R}p'$, then $i \ne j$ (pieces in relation cannot be put at the same height).
  2. If $(p, i) \in H$ and $i > 0$, then there exists $(p', i - 1) \in H$ with $p\mathcal{R}p'$ (each piece must be supported).

These points of view are equivalent (Viennot [23], Krattenthaler [14]); each heap of pieces H can also be viewed as a poset $(H, \le)$, where:

- in the geometric point of view, $(p, i) \le (p', j)$ if $i \le j$ and $p\mathcal{R}p'$ (and $\le$ is the transitive closure of this relation),
- in the Cartier-Foata point of view, for two letters a and b in a word, $a \le b$ if $a\mathcal{R}b$ and a is at the left of b, (and $\le$ is the transitive closure of this relation).

A pair (p,i) in H is said to be maximal in H, if H does not contain any pair (p',j) such that $p'\mathcal{R}p$ and $j\geq i$ (there are no pieces in relation, above it).

Each heap, as a poset, possesses some maximal pair. We say that a piece p is maximal if the pair (p,i) is maximal for some height i.

A trivial heap of pieces is a heap in which all pieces are at level 0, which means that the pieces it contains are not in relation. If one uses the partially commutative monoid point of view, a trivial heap of pieces is a heap (a word) in which all the pieces (the letters) commute.

Proposition 1.6 (Prop. 5.3 [23])**.** *Let $\mathcal{M}$ be a subset of the pieces $\mathcal{B}$. Let W be a multiplicative weight function on heaps, such that for all heap H its weight $W(H)$ is the product of the elementary weights $W(p)$ of the pairs (p,i) it contains (the weight of a piece is independent of "its place or height" in the heap). Then, the total weight of the heap of pieces having their maximal pieces included in $\mathcal{M}$ is given by*

$$\sum_{\substack{H \text{ heaps in } (\mathcal{B},\mathcal{R})\\ \text{maximal pieces } \subset\mathcal{M}}} W(H)=\Big(\sum_{\substack{T \text{ trivial}\\ \text{heap in } (\mathcal{B},\mathcal{R})}}(-1)^{|T|}W(T)\Big)^{-1}\Big(\sum_{\substack{T \text{ trivial}\\ \text{heap in } (\mathcal{B}\setminus\mathcal{M},\mathcal{R})}}(-1)^{|T|}W(T)\Big)$$

Viennot [23, Proposition 5.3] gave this result at the level of combinatorial objects; here, we preferred a projected version, in terms of their weights (which is what we need). (See also Theorem 4.1 [14], Tyler [9, Appendix A]).

In a heap of cycles, the pieces are oriented cycles on a given graph G, and two cycles are in relation if they share a vertex. The weight of a heap of cycles, according to a transition matrix M, is identified with the weight of the collection of cycles it contains. A heap of cycles is then trivial when all the cycles it contains are non-intersecting. Denote by A_G the alternating weight of trivial heaps of cycles

$$A_G=\sum_{C=(C(1),\dots,C(|C|)\in\text{Trivial heap of cycles on } G}(-1)^{|C|}\prod_{j=1}^{|C|}\mathsf{Weight}(C(j)),$$

where $W(C(j))$ is as in (1.14). A simple expansion of the determinant using the cycles present on a permutation allows us to get

$$A_G=\det(\mathsf{Id}-\mathsf{M})=0 \tag{1.15}$$

and the reason for that is that M has 1 as an eigenvalue; hence the set of heaps of cycles on G has total weight $+\infty$. What is of greater interest is the value of $A_{G\setminus S}$, the alternating weight of trivial heaps of cycles avoiding some set of vertices S, which is

$$A_{G\setminus S}=\det\left(\mathsf{Id}-\mathsf{M}_{G\setminus S}\right), \tag{1.16}$$

as well as its inverse corresponding to the total weight of heaps of cycles on $G\setminus S$:

$$\sum_{H\in\text{Heap of Cycles on } G\setminus S}\mathsf{Weight}(H)=\det\left(\mathsf{Id}-\mathsf{M}_{G\setminus S}\right)^{-1}. \tag{1.17}$$

2 Main theorems in the almost triangular cases

2.1 Almost upper triangular cases

In the ◹ irreducible case, there exists a unique invariant measure:

Theorem 2.1. *If U is an irreducible ◹ transition matrix, then U admits a unique positive invariant measure $(\pi_a,a\geq 0)\in\mathcal{M}_{\mathbb{N}}^{+,\sim}$, which is defined (up to a constant factor $\pi_0>0$) by*

$$\pi_a:=\pi_0\,\frac{\det(\mathsf{Id}-\mathsf{U}_{[0,a-1]})}{\prod_{j=1}^{a}\mathsf{U}_{j,j-1}},\quad a\geq 1.$$

The transition matrix U *is positive recurrent if and only if*

$$\sum_{a=1}^{\infty} \frac{\det\left(\mathsf{Id} - \mathsf{U}_{[0,a-1]}\right)}{\prod_{j=1}^{a} \mathsf{U}_{j,j-1}} < \infty. \tag{2.1}$$

Remark 2.2. *(i) The measure* π *is a positive measure on* $\mathbb{N}$*, this is a consequence of Proposition 1.4 and can be seen using (1.10) too, even if the matrix* $\mathsf{U}_{[0,a-1]}$ *is not a transition matrix deprived of a row and a column (but it can be obtained as such).*

(ii) We have $\pi_a = \sum_{b\le a+1} \pi_b \mathsf{U}_{b,a}$ *so that there is a second algorithmic method to compute directly* $(\pi_a, a \ge 0)$*: fix freely a value* $\pi_0 > 0$*, and then for* $a \ge 1$ *use the following recursion:*

$$\pi_{a+1} = \Big(\pi_a - \sum_{b\le a} \pi_b \mathsf{U}_{b,a}\Big) / \mathsf{U}_{a+1,a}, \tag{2.2}$$

which provides immediately the uniqueness of the invariant measure. The equivalence of this formula with Theorem 2.1 is not obvious, and it is even not obvious that (2.2) produces a positive sequence $(\pi_a, a \ge 0)$*.*

(iii) The theorem applies in the tridiagonal case even if the formula seems different from Karlin & McGregor's formula (1.5) (see Section 4.1 for a complete explanation).

The next proposition gives a property of the right eigenspace of a ◹-transition matrix, associated with the eigenvalue 1. Remark 2.4 announces a maybe unexpected link with the recurrence of this chain.

Proposition 2.3. *There exist some ◹ irreducible transition matrices* U *with several positive, non-proportional right eigenvectors associated with the eigenvalue 1: there exist cases for which the cone generated by the right eigenvectors with positive coordinates is infinite-dimensional, and cases for which it is finite-dimensional, all dimensions* ≥ 1 *being possible.*

Of course, the vector $R = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}^t$ is a right eigenvector of U associated with the eigenvalue 1.

Remark 2.4. *We will see in Theorem 3.1, that a ◹ matrix* U *having various linearly independent positive right eigenvectors (associated with the eigenvalue 1) is transient.*

Remark 2.5. *The right eigenspace associated with the eigenvalue 1 is a vector space; here we are interested in the cone generated by the vectors with positive coordinates (the cone being the set of linear combinations with non-negative coefficients of these vectors). The cone of eigenvectors with positive coordinates is a convex set, but not a vector space (except when it is reduced to* $\{0\}$*). However, we can still define the dimension of a cone, as the dimension of the vector space it generates. A basis of this vector space (which can be taken in the cone itself) characterizes the cone of vectors with positive coordinates (these are the linear combination of these vectors with all positive coordinates).*

By the way, the existence of right eigenvectors for U has a martingale interpretation:

Remark 2.6. *We claim the following: if* $(X_k, k \ge 0)$ *is a Markov chain with irreducible transition matrix* M*, and let* $\mathcal{F} = (\mathcal{F}_n, n \ge 0) = (\sigma(X_i, i \le n), n \ge 0)$ *be the filtration generated by the* X_i*'s. Consider a function* $R : \mathbb{N} \to [0, +\infty)$ *(which we want to view also as a vector* $R = \begin{bmatrix} R_0, & R_1, & R_2, & \cdots \end{bmatrix}^t$*). The process* $(R_{X_k}, k \ge 0)$ *is* $\mathcal{F}$*-martingale iff* R *is a right eigenvector of* M *associated with the eigenvalue 1.*

The proof of the claim is simple: Since $(X_k, k \geq 0)$ *is a Markov chain,* $(R_{X_k}, k \geq 0)$ *is a martingale iff* $\mathbb{E}(R_{X_{k+1}}|\mathcal{F}_k) = \mathbb{E}(R_{X_{k+1}}|X_k) = R_{X_k}$*. Since* $(X_k, k \geq 0)$ *is time-homogeneous and discrete, observe that* $(R_{X_k}, k \geq 0)$ *is a martingale iff, for all* $i \geq 0$, $E(R_{Z_1}|Z_0 = i) = R_i$*, which is equivalent to* $\sum_{j \geq i-1} M_{i,j} R_j = R_i$ *as claimed.*

When the state space of the Markov chain is finite, since 1 is a simple eigenvalue, only $R = [1, \ldots, 1]^t$ *is a right eigenvector (up to a multiplicative constant), so that, no "interesting martingale" of the type* $(R_{X_k}, k \geq 0)$ *can be defined. Proposition 2.3 states that in the case of infinite* $\triangledown$ *transition matrices, the situation may be different, and the class of martingales of this type may be huge.*

The next lemma provides a direct algebraic argument leading to the form of the invariant measures as stated in Theorem 2.1.

Lemma 2.7. *For any finite or infinite* $\triangledown$ *matrix* U*, any* $y \in \{0, \ldots, s\}$*, where the column are indexed from 0 to* $s \leq +\infty$*, we have*

$$\det(\mathsf{Id} - \mathsf{U}_{[0,y]}) = \det(\mathsf{Id} - \mathsf{U}_{[0,y-1]})\,(1 - \mathsf{U}_{y,y}) - \sum_{x \leq y-1} \det(\mathsf{Id} - \mathsf{U}_{[0,x-1]})\,\mathsf{U}_{x,y} \prod_{j=x+1}^{y} \mathsf{U}_{j,j-1}, \tag{2.3}$$

with the convention $\det(\mathsf{Id} - \mathsf{U}_{[0,-1]}) = 1$.

Proof. This is an application of Lemma 1.1 to the $\triangledown$ matrix $A = \mathsf{Id} - U_{[0,y]}$, more exactly to the matrix obtained from $\mathsf{Id} - U_{[0,y]}$ by the symmetry with respect to the <u>second</u> diagonal. □

Proof of Theorem 2.1. In the proof, we write D_x instead of $\det(\mathsf{Id} - \mathsf{U}_{[0,x]})$. Consider the formula for π provided in Theorem 2.1; let us establish that $\sum_x \pi_x \mathsf{U}_{x,y} = \pi_y$ so that π is indeed invariant by U. Since U is $\triangledown$ (with rows and columns indexed from 0 to s), $\sum_x \pi_x \mathsf{U}_{x,y} = \pi_y$ is equivalent to

$$\sum_{x=0}^{\min\{s,y+1\}} \frac{D_{x-1}}{\prod_{j=1}^{x} \mathsf{U}_{j,j-1}} \mathsf{U}_{x,y} \quad = \quad \frac{D_{y-1}}{\prod_{j=1}^{y} \mathsf{U}_{j,j-1}}. \tag{2.4}$$

- Assume first that $y < s$ (so that $\min\{s, y+1\} = y+1$), and multiply both sides of (2.4) by $\prod_{j=1}^{y} \mathsf{U}_{j,j-1}$ allows seeing that (2.4) is equivalent to

$$\sum_{x=0}^{y-1} D_{x-1} \mathsf{U}_{x,y} \prod_{j=x+1}^{y} \mathsf{U}_{j,j-1} + D_{y-1} \mathsf{U}_{y,y} + D_y \quad = \quad D_{y-1}. \tag{2.5}$$

  which is (2.3), so that it holds, and we have indeed $\sum_x \pi_x \mathsf{U}_{x,y} = \pi_y$.
- Assume that $y = s$ (and then, s is finite). Now, $\min\{s, y+1\} = s$, and the multiplication of (2.4) by $\prod_{j=1}^{s} \mathsf{U}_{j,j-1}$ gives,

$$\sum_{x=0}^{s-1} D_{x-1} \mathsf{U}_{x,s} \prod_{j=x+1}^{s} \mathsf{U}_{j,j-1} + D_{s-1} \mathsf{U}_{s,s} \quad = \quad D_{s-1}. \tag{2.6}$$

  and this holds too by (2.3): to see this, we need to additionally observe that $D_s = 0$ (because 1 is an eigenvalue of U).

The invariant measure defines a probability measure and is therefore positive recurrent if and only if $\sum_{i=0}^{\infty} \pi_i < \infty$ which gives (2.1). □

A second proof of Theorem 2.1 will be given in Section 2.3.1.

Theorem 2.8. *For a Markov chain $Y = (Y_i, i \geq 0)$ with $\triangledown$ irreducible transition matrix* U*, denote by*

$$\tau_S(Y) = \inf\{j > 0 : Y_j \in S\}$$

the hitting time of the set S by Y. Set, for any pairs of integers (x, b) such that $0 < x < b$,

$$u_b(x) = \mathbb{P}(\tau_{\{0\}}(Y) < \tau_{[b,+\infty)}(Y) \mid Y_0 = x).$$

We have

$$u_b(x) = \frac{\det(\mathsf{Id} - \mathsf{U}_{[x+1,b-1]})}{\det(\mathsf{Id} - \mathsf{U}_{[1,b-1]})} \prod_{j=1}^{x} \mathsf{U}_{j,j-1} \tag{2.7}$$

so that U *is* ***recurrent*** *if and only if*

$$\lim_{b\to+\infty} u_b(1) = \lim_{b\to+\infty} \mathsf{U}_{1,0} \frac{\det(\mathsf{Id} - \mathsf{U}_{[2,b-1]})}{\det(\mathsf{Id} - \mathsf{U}_{[1,b-1]})} = 1. \tag{2.8}$$

In Section 4.1 we will see that in the tridiagonal case, this criterion reduces to Karlin & McGregor criterion (1.7).

Proof. The formulas of $u_b(x)$ and $u_b(1)$ are direct consequences of Lemma 1.5 and of Proposition 1.6 ((1.16) and (1.17) explaining the appearance of determinants), because, a path from x to 0 which stays in $[0, b-1]$, and which hits 0 for the first time by its last step, can be decomposed as a self-avoiding path going from x to 0 (its weight is $\prod_{j=1}^{x} \mathsf{U}_{j,j-1}$) and a heap of cycles on the vertex set $[1, b-1]$ with maximal pieces incident to $[1, x]$, whose weight is given by $\frac{\det(\mathsf{Id}-\mathsf{U}_{[x+1,b-1]})}{\det(\mathsf{Id}-\mathsf{U}_{[1,b-1]})}$; this gives (2.7) at once.

Recurrence is equivalent to $u_b(1) \to 1$ when $b \to +\infty$, since U is irreducible. $\square$

Proposition 2.9. *For $x \in (a, b)$, set $\overleftarrow{u}_{a,\geq b}(x; z) = \mathbb{E}\left[z^{\tau_{\{a\}}(Y)} \boldsymbol{1}_{\tau_{\{a\}}(Y) < \tau_{[b,+\infty)}(Y)} \mid Y_0 = x\right]$ the (defective) generating function of the hitting time of a under the event that a is reached before b. We have*

$$\overleftarrow{u}_{a,\geq b}(x; z) = \frac{\det\left(\mathsf{Id} - z\mathsf{U}_{[x+1,b-1]}\right)}{\det\left(\mathsf{Id} - z\mathsf{U}_{[a+1,b-1]}\right)} \prod_{j=a+1}^{x} (z\mathsf{U}_{j,j-1}). \tag{2.9}$$

In particular when U is recurrent $\overleftarrow{u}_{0,b}(x; z)$ converges[1] *to $\overleftarrow{u}(x; z) := \mathbb{E}\left[z^{\tau_{\{0\}}(Y)} \mid Y_0 = x\right]$ as $b \to +\infty$; otherwise it converges to*

$$\overleftarrow{u}(x; z) = \mathbb{E}\left[z^{\tau_{\{0\}}(Y)} \boldsymbol{1}_{\tau_{\{0\}}(Y) < +\infty} \mid Y_0 = x\right]. \tag{2.10}$$

Proof. The proof of the first statement is the same as that of Theorem 2.8, in which the weight $\mathsf{U}_{i,j}$ of a step is replaced by $\mathsf{U}_{i,j} z$. Second statement: in case of recurrence, $\mathbb{P}(\tau_{\{0\}}(Y) \leq \tau_{\{b\}}(Y) \,|Y_0 = x) \to 1$ when $b \to +\infty$. The statement, which is then equivalent to the convergence in distribution of $\tau_{\{0\}}(Y)$ conditioning on the event $\{\tau_{\{0\}}(Y) \leq \tau_{\{b\}}(Y)\}$ follows. $\square$

[1] It converges in the sense that, for all k, the coefficient of z^k in $\overleftarrow{u}_{0,b}(x; z)$ converges to that of $\overleftarrow{u}(x; z)$ as $b \to +\infty$. Seen as a power series in z, $\overleftarrow{u}_{0,b}(x; z)$ converges uniformly on each compact included in $[0, 1)$ to $\overleftarrow{u}(x; z)$.

We define now the transition matrix $\mathsf{U}^{(n)}$, which we will call the "projected" transition matrix U on $[0,n]$:

$$\begin{cases} \mathsf{U}^{(n)}_{i,j} & = & \mathsf{U}_{i,j}, \quad \text{for} \quad i \in [0,n], j \in [0,n-1] \\ \mathsf{U}^{(n)}_{i,n} & = & \sum_{j \geq n} \mathsf{U}_{i,j}. \end{cases} \tag{2.11}$$

It will often be used in the sequel (as well as $\mathsf{L}^{(n)}$ defined in (2.21)).

Remark 2.10. *Let* $\overrightarrow{u}_{a,\geq b}(x;z) = \mathbb{E}\left[z^{\tau_{[b,+\infty)}(Y)} \boldsymbol{1}_{\tau_{\{a\}}(Y) > \tau_{[b,+\infty)}(Y)} \mid Y_0 = x\right]$ *be the (defective) generating function of the hitting time of* $[b,+\infty)$ *under the event that* $[b,+\infty)$ *is reached before* a. *We have*

$$\overrightarrow{u}_{a,\geq b}(x;z) = \left((\mathsf{Id} - z\mathsf{U}_{[a+1,b-1]})^{-1} \left[\boldsymbol{1}_{i\in[a+1,b-1]} z \textstyle\sum_{k\geq b} \mathsf{U}_{i,k}\right]\right)_{x,1} \tag{2.12}$$

where $\mathsf{U}_{[a+1,b-1]}$ *is the matrix obtained by preserving the entries of* U *whose columns and rows are indexed by* $[a+1,b-1]$*, and replacing the others by 0, and* $\left[\boldsymbol{1}_{i\in[a+1,b-1]} z \sum_{k\geq b} \mathsf{U}_{i,k}\right]$ *the column vectors in which non-zero rows correspond to the* i *in* $[a+1,b-1]$*: the entry* $\boldsymbol{1}_{i\in[a+1,b-1]} \sum_{k\geq b} \mathsf{U}_{i,k}$ *measures the probability to jump at the right of* b *from* i. *The matrices in play are infinite, but representations with finite matrices also exist (for example, keeping the* $b+1$ *first rows and columns of the involved matrices suffice).*

To prove this formula, just observe that when M *is a transition matrix of a finite state space (with size* $s \geq b$*), the coefficient* $(\mathsf{Id} - z\mathsf{U}_{[a+1,b-1]})^{-1}_{i,j}$ *gives the total mass of paths starting at* i*, ending at* j*, and whose set of vertices is included in* $[a+1,b-1]$ *(a step* (k,ℓ) *being weighted* $zM_{k,\ell}$*).*

Remark 2.11. *Some authors consider the case where* $\sum_{k\geq 0} \mathsf{U}_{0,k} < 1$*, so that, a part of the mass disappears at each passage at 0: if one adds an additional absorbing state* $\dagger$ *to the state space, and set* $\mathsf{U}_{0,\dagger} = 1 - \sum_{k\geq 0} \mathsf{U}_{0,k}$ *and* $\mathsf{U}_{\dagger,\dagger} = 1$*, then the absorbed mass at* $\dagger$ *for a* U*-Markov chain starting from* x *is*

$$A_\dagger(x) = \mathbb{P}(\tau_{\{\dagger\}}(X) < +\infty \mid X_0 = x)$$

and the corresponding (defective) hitting time generating function is

$$a_\dagger(x;z) = \mathbb{E}(z^{\tau_{\{\dagger\}}(X)} \boldsymbol{1}_{\tau_{\{\dagger\}}(X) < +\infty} \mid X_0 = x).$$

Recall (2.10). We have $A_\dagger(x) = a_\dagger(x;1)$ *and*

$$a_\dagger(x;z) = \overleftarrow{u}(x;z) z\mathsf{U}_{0,\dagger} \sum_{k=0}^{\infty} \Big(z\mathsf{U}_{0,0} + \sum_{y=1}^{\infty} zU_{0,y} \overleftarrow{u}(y;z)\Big)^k = \frac{z\mathsf{U}_{0,\dagger} \overleftarrow{u}(x;z)}{1 - (z\mathsf{U}_{0,0} + \sum_{y=1}^{\infty} zU_{0,y} \overleftarrow{u}(y;z))}$$

where $\overleftarrow{u}$ *is defined in Proposition 2.9. To show this formula, a simple decomposition is sufficient. Each path going to* $\dagger$ *can be decomposed as:*

- *a trajectory that goes to 0 (whose weight is taken into account by* $\overleftarrow{u}(x;z)$*),*
- *a sequence of* k *cycles from 0 to 0 (each of them contributes* $(z\mathsf{U}_{0,0} + \sum_{y=1}^{\infty} zU_{0,y} \overrightarrow{u}(y,z))$*),*
- *and then a step leading to* $\dagger$ *from 0.*

Definition 2.12. *Let* $\pi, \pi^{(1)}, \pi^{(2)}, \dots,$ *be a sequence of measures on* $\mathbb{N}$. *The sequence* $(\pi^{(n)})$ *is said to converge weakly to* π *if for all* $k \geq 0$, $\lim_n \pi^{(n)}_k = \pi_k$.

When these measures are probability measures on $\mathbb{N}$, this is the classical convergence in distribution.

Theorem 2.13. *Let* U *be a* $\triangledown$ *transition matrix, irreducible on* $\mathbb{N}$*, and* $\mathsf{U}^{(n)}$ *be the projected transition matrix defined in (2.11). Denote by* $\rho^{(n)}$ *the unique invariant probability distribution of* $\mathsf{U}^{(n)}$*.*

(i) The transition matrix U *admits* $\pi \in \mathcal{M}_{\mathbb{N}}^{+}$ *as an invariant measure, if and only if there exists a sequence* $(c_n, n \geq 0)$ *such that* $c_n\rho^{(n)} \to \pi$ *weakly.*

(ii) U *is positive recurrent with invariant probability distribution* ρ *iff* $\rho^{(n)} \to \rho$ *weakly.*

Proof. Since $\mathsf{U}_{i,j} = \mathsf{U}_{i,j}^{(n)}$ for $i < n$, the equilibrium equations

$$\pi_b = \sum_{a \leq b+1} \pi_a \mathsf{U}_{a,b},$$
$$\rho_b^{(n)} = \sum_{a \leq b+1} \rho_a^{(n)} \mathsf{U}_{a,b}^{(n)} = \sum_{a \leq b+1} \rho_a^{(n)} \mathsf{U}_{a,b}$$

are the same for $b \leq n-1$. These systems can be rewritten to express π_{b+1} (respectively $\rho_{b+1}^{(n)}$) in terms of π_j with smaller indices j, as follows:

$$\pi_{b+1} = \left(\pi_b - \sum_{a \leq b} \pi_a \mathsf{U}_{a,b}\right) / U_{b+1,b}, \tag{2.13}$$

$$\rho_{b+1}^{(n)} = \left(\rho_b^{(n)} - \sum_{a \leq b} \rho_a^{(n)} \mathsf{U}_{a,b}\right) / U_{b+1,b} \tag{2.14}$$

for $b \leq n-1$. Fixing a value for π_0 allows deducing the proportionality

$$(\pi_i, 0 \leq i \leq n-1) = C_n \big(\rho_i^{(n)}, 0 \leq i \leq n-1\big) \tag{2.15}$$

for a constant $C_n > 0$ (recall that $\rho^{(n)}$ is the probability distribution which is invariant by $\mathsf{U}^{(n)}$, it is normalised to have sum 1, so that one can not take $\rho_0^{(n)} = 1$). The uncontrolled weight $\rho_n^{(n)}$ is not a detail at all, since it is directly related to C_n. If C_n goes to $+\infty$, for example, it means that the mass $\rho_i^{(n)}$ vanishes when $n \to +\infty$, but this does not prevent $C_n\rho_i^{(n)}$ to converge.

Proof of (i). Assume that π is invariant by U, by (2.15), $C_n\rho^{(n)} \to \pi$ weakly. Conversely, assume that $C_n\rho^{(n)} \to \pi$. Still by (2.15) and (2.13), π is invariant by U.

Proof of (ii). First, if $\rho^{(n)} \to \rho$ weakly, then, since ρ is assumed to be a probability distribution, by (i) ρ is invariant by U, and since ρ is summable and U irreducible, then U is positive recurrent.

Conversely, assume that ρ is invariant by U, and U positive recurrent. Assume that $X = (X_i, i \geq 0)$ is a U-Markov chain, and denote by $(\tau_i^{(n)}, i \geq 0)$ the sequence of all times t such that $X_t \leq n$ (taking in their initial order). It is easy to check that $\big(X_{\tau_i^{(n)}}, i \geq 0\big)$ is a $\mathsf{U}^{(n)}$-Markov chain (because the decreasing steps are -1). Since U is positive recurrent, the ergodic theorem applies, and $|\{i : X_i = x, i \in [0, m-1]\}|/m \to \rho_x$ as $m \to \infty$ (a.s.). The ergodic theorem applies also to $(X_{\tau_i^{(n)}})$ (which is a Markov chain on a finite state space), that $|\{i : X_i^{(n)} = x, i \in [0, m-1]\}|/m \to \rho_x^{(n)}$ for $x \leq n$, and by restriction of X, $\rho_x^{(n)}$ is the time proportion of the chain X at x divided by the time passed under level n, so that $\rho_x^{(n)} = \rho_x / (\sum_{y \leq n} \rho_y)$. Since $\sum_{y \leq n} \rho_y \to 1$ as $n \to +\infty$, we have $\lim_n \rho_x^{(n)} = \rho_x$. $\square$

2.2 Almost lower triangular cases

In the ◺ case, neither uniqueness nor existence of invariant measures is guaranteed:

Theorem 2.14. *Let* L *be an irreducible* ◺ *transition matrix.*

(i) L *has a unique right eigenvector associated with the eigenvalue 1 (up to a multiplicative constant), and this is the vector whose entries are all equal to one.*

(ii) The three following cases arise: (a) L *has no invariant measure in* $\mathcal{M}_{\mathbb{N}}^{+,\sim}$, (b) L *has a unique invariant measure in* $\mathcal{M}_{\mathbb{N}}^{+,\sim}$, (c) L *has several invariant measures in* $\mathcal{M}_{\mathbb{N}}^{+,\sim}$ *(the cone generated by these measures can have any dimension in* $\{1,2,3,\ldots\}\cup\{+\infty\}$*).*

Proof. (i) Write the system $\mathsf{L}R=R$ under the triangular form

$$R_{i+1}=(R_i-\sum_{j:j\leq i}\mathsf{L}_{i,j}R_j)/\mathsf{L}_{i+1,i},\ \text{for } i\geq 0,$$

so that the choice of $R_0=1$ fixes all the other entries to 1.

$(ii)(b)$ All tridiagonal irreducible transition matrices have a unique invariant measure in $\mathcal{M}_{\mathbb{N}}^{+,\sim}$ since they are ◹, and then Theorem 2.1 applies.

$(ii)(a)$ In some lecture notes on Markov chains, as an example of Markov chain on $\mathbb{N}$ with no invariant distribution, a transition matrix of type ◺ is often given (see e.g. [18, Example 1.7.11]). For the sake of completeness, we provide a similar example here. Take $\mathsf{L}_{0,1}=a_0=1$, and for $i\geq 1$, $\mathsf{L}_{i,i+1}=a_i$, $\mathsf{L}_{i,0}=1-a_i$. Notice that if $a_i\in(0,1)$ for all $i\geq 0$, then L is irreducible. An invariant measure π would satisfy for $i>0$, $\pi_i=\pi_{i-1}\mathsf{L}_{i-1,i}$, so that $\pi_i=\pi_0\prod_{j=0}^{i-1}a_j$ and therefore $\pi_0=\sum_{k\geq 1}\pi_k\mathsf{L}_{k,0}=\pi_0\sum_{k\geq 1}(1-a_k)\prod_{j=0}^{k-1}a_j$. From this we see that: L has an invariant measure if and only if $\sum_{k\geq 1}(1-a_k)\prod_{j=0}^{k-1}a_j=1$. Consider $a_0=1$, and for $j\geq 1$, $a_j=1-\frac{1}{(j+1)^2}=\frac{j(j+2)}{(j+1)^2}$ and then, since it is a telescopic product

$$\sum_{k\geq 1}(1-a_k)\prod_{j=1}^{k-1}a_j=\sum_{k\geq 1}(1-a_k)\prod_{j=1}^{k-1}\frac{j(j+2)}{(j+1)^2}=\sum_{k\geq 1}\frac{1}{(k+1)^2}\frac{k+1}{2k}=1/2.$$

$(ii)(c)$ We give the proof here even if it depends on a result which is stated in the sequel. This statement is a consequence of Proposition 2.3 and of Theorem 3.1 (v) (whose proofs do not depend on the present theorem). In words Theorem 3.1 (v) states: time reversals of a ◹-Markov chain U is a ◺ Markov chain with a kernel L, and one can define a linear bijection between the cone generated by the right eigenspace of U (with positive coordinates, and associated with the eigenvalue 1), to the set of invariant measures for L. The conclusion follows from Proposition 2.3, where we state that there exist matrices ◹ U with a right cone (associated with the eigenvalue 1) of any possible dimension ≥ 1. □

Theorem 2.15. *Let* L *be a* ◺ *irreducible transition matrix and* $(Y_i,i\geq 0)$ *a* L*-Markov chain. Set*

$$v_b(x)=\mathbb{P}(\tau_{\{0\}}(Y)>\tau_{\{b\}}(Y)\mid Y_0=x),\quad \textit{for } 0<x<b.$$

We have

$$v_b(x)=\frac{\det(\mathsf{Id}-\mathsf{L}_{[1,x-1]})}{\det(\mathsf{Id}-\mathsf{L}_{[1,b-1]})}\prod_{j=x}^{b-1}\mathsf{L}_{j,j+1},\tag{2.16}$$

and then, the transition matrix L *is recurrent if and only if*

$$\lim_{b\to+\infty}v_b(1)=\lim_{b\to+\infty}\frac{\prod_{j=1}^{b-1}\mathsf{L}_{j,j+1}}{\det(\mathsf{Id}-\mathsf{L}_{[1,b-1]})}=0.\tag{2.17}$$

Remark 2.16. *(i) Observe that $v_b(x)$ is not defined as $u_b(x)$ (of Theorem 2.8), but rather, corresponds to $1-u_b(x)$. Note that since Y has its jumps bounded by $+1$, it hits $[b,+\infty)$ at b (starting from x). See Remark 3.7 for a subtle point.*

(ii) By the forthcoming Proposition 3.5, Condition (2.17) in Theorem 2.15 is equivalent to

$$\frac{\det(\mathsf{Id}-\mathsf{L}_{[0,b-1]})}{\mathsf{L}_{0,1}\det(\mathsf{Id}-\mathsf{L}_{[1,b-1]})}\to 0. \tag{2.18}$$

Proof of Theorem 2.15. The proof is similar to that of Theorem 2.8. Recurrence is equivalent to $v_1(b)\to 0$ when $b\to+\infty$. The number $v_b(x)$ is the weight of the set of paths that starts at x, ends at b reached for the first time at the end of the path, and does not touch 0. Each of these paths can be decomposed as a self-avoiding path from x to b (whose weight is $\prod_{j=x}^{b-1}\mathsf{L}_{j,j+1}$), and a heap of cycles on $[1,b-1]$ whose maximal pieces are incident to $[x,b-1]$; the total weight of these heap of cycles is the quotient in (2.16). □

An analogue of Proposition 2.9:

Proposition 2.17. *For $x\in(a,b)$, set $\overrightarrow{v}_{\le a,b}(x;z)=\mathbb{E}\left[z^{\tau_{\{b\}}(Y)}\mathbf{1}_{\tau_{[0,a]}(Y)>\tau_{\{b\}}(Y)}\mid Y_0=x\right]$ and $\overleftarrow{v}_{\le a,b}(x;z)=\mathbb{E}\left[z^{\tau_{\{a\}}(Y)}\mathbf{1}_{\tau_{[0,a]}(Y)<\tau_{\{b\}}(Y)}\mid Y_0=x\right]$. We have*

$$\overrightarrow{v}_{\le a,b}(x;z)=\frac{\det\left(\mathsf{Id}-z\mathsf{L}_{[a+1,x-1]}\right)}{\det\left(\mathsf{Id}-z\mathsf{L}_{[a+1,b-1]}\right)}\prod_{j=x}^{b-1}(z\mathsf{L}_{j,j+1}),$$

$$\overleftarrow{v}_{\le a,b}(x;z)=\sum_{j=0}^{a}\left[\left(\mathsf{Id}-z\mathsf{L}_{[a+1,b-1]}\right)^{-1}\left[1_{i\in[a+1,b-1]}z\mathsf{L}_{i,j}\right]\right]_{x,1},$$

where $\left[1_{i\in[a+1,b-1]}z\mathsf{L}_{i,j}\right]$ is a column vector, whose rows are indexed by i.

The proof is a simple adaptation of that of Proposition 2.9.

Remark 2.18. *Absorption at 0: consider a ◺ transition matrix L, for which $\sum_{k\ge0}\mathsf{L}_{0,k}<1$, and add again an additional absorbing state $\dagger$ to the state space, and set $\mathsf{L}_{0,\dagger}=1-\sum_{k\ge0}\mathsf{L}_{0,k}$ and $\mathsf{L}_{\dagger,\dagger}=1$. The absorbed mass at $\dagger$ starting from x is $B_\dagger(x)=\mathbb{P}(\tau_{\{\dagger\}}(Y)<+\infty\mid Y_0=x)$ for L-Markov chain Y, and the corresponding (defective) hitting time generating function is $b_\dagger(x;z)=\mathbb{E}(z^{\tau_{\{\dagger\}}(Y)}\mathbf{1}_{\tau_{\{\dagger\}}(Y)<+\infty}\mid Y_0=x)$. Recall (2.10). We have $B_\dagger(x)=b_\dagger(x;1)$ and*

$$b_\dagger(x;z)=\frac{z\mathsf{L}_{0,\dagger}\overleftarrow{v}(x;z)}{1-\left(z\mathsf{L}_{0,0}+\sum_{y=1}^{\infty}zL_{0,y}\overleftarrow{v}(y;z)\right)}$$

where $\overleftarrow{v}(x;z)=\lim_{b\to+\infty}\overleftarrow{v}_{\le 0,b}(x;z)$.

An invariant measure η satisfies $\eta_b=\sum_{a\ge b-1}\eta_a\mathsf{L}_{a,b}$ so that

$$\eta_{b-1}=\frac{\eta_b(1-\mathsf{L}_{b,b})-\sum_{a\ge b+1}\eta_a\mathsf{L}_{a,b}}{\mathsf{L}_{b-1,b}}. \tag{2.19}$$

The fact that η_{b-1} is expressed using the η_a with larger indices a brings a very important difficulty here: formula (2.19) can be used to check that a sequence $(\eta_k,k\ge0)$ is indeed invariant, but, it seems unsuitable to compute an invariant distribution; and once again, such a solution does not exist in all generality.

Theorem 2.19. *Let* L *be an irreducible* ◺ *transition matrix with* <u>*finite size*</u> *(indexed by* $[0,s]\times[0,s]$*). For any* $\eta_0>0$*, set for* $a\in[0,s]$*,*

$$\eta_a=\eta_0\ \det\left(\mathsf{Id}-\mathsf{L}_{[a+1,s]}\right)\ \prod_{i=1}^{a}\mathsf{L}_{i-1,i}. \tag{2.20}$$

The measure $(\eta_a, a\in[0,s])$ *is invariant by* L *(and by Perron-Frobenius, there is a single class of invariant measures).*

Proof. This is a consequence of (1.10) and of Proposition 1.4. Indeed, observe the geometry of the graph with vertex set $[0,s]$ and edge set $\{(i,j):\mathsf{L}_{i,j}>0\}$. Take any spanning tree rooted at a: The vertices in $[0,a-1]$ can be connected to a only using the edges $0\mapsto 1\mapsto\cdots\mapsto a$ (so that the observed tree contains this branch), and the rest of the edges of the tree, forms a forest whose root set is contained in $[0,a]$ having set of nodes $[a+1,s]$. □

Given this theorem, it is tempting to think that when L is indexed by $\mathbb{N}$, and say, irreducible, its invariant distribution is obtained by just taking $\lim_n\det\big(\mathsf{Id}-\mathsf{L}^{(n)}_{[a+1,n]}\big)\times\prod_{i=1}^{a}\mathsf{L}_{i-1,i}$ where $\mathsf{L}^{(n)}$ is the projected transition matrix of L on $[0,n]$ defined by

$$\begin{cases}\mathsf{L}^{(n)}_{i,j} &=& \mathsf{L}_{i,j}, \quad \text{for} \quad 0\le i\le n, 0\le j\le n-1\\ \mathsf{L}^{(n)}_{i,n} &=& \sum_{j\ge n}\mathsf{L}_{i,j}.\end{cases} \tag{2.21}$$

But it is not the case, since Theorem 2.14 establishes that a transition matrix L is not assured to have an invariant distribution. The complete picture is more complex and some additional conditions are needed to get this kind of convergence result. The following Propositions 2.20 and 2.21 provide two criteria (Prop. 2.21 is simpler, but can be applied only when L is positive recurrent).

Proposition 2.20. *Let* L *be a* ◺ *irreducible transition matrix. Let* $\rho^{(n)}$ *be the invariant probability distribution of* $\mathsf{L}^{(n)}$ *(see (2.20)). Set*

$$\eta^{(n)}_a:=\rho^{(n)}_a\,/\,\rho^{(n)}_0,\quad \textit{for } a\ge 0.$$

If the two following conditions hold:

- *(a) there exists a non-negative sequence* $(S_a,a\ge 0)$ *such that, for each* b*,* $\sum_{a=b-1}^{+\infty}S_a\mathsf{L}_{a,b}<+\infty$*, and which bounds uniformly* $\eta^{(n)}$*: for all* $a,n\ge 0$*,* $|\eta^{(n)}_a|\le S_a$*,*
- *(b)* $\lim_n\eta^{(n)}_a$ *exists for each* a*; set* $\eta_a:=\lim_n\eta^{(n)}_a$ *for* $n\to+\infty$*,*

then η *is invariant by* L*; that is, for all* b*,* $\eta_b=\sum_{a=b-1}^{+\infty}\eta_a\mathsf{L}_{a,b}$*.*

Proof. We will prove that η satisfies $\eta_b=\sum_{a=b-1}^{+\infty}\eta_a\mathsf{L}_{a,b}$ for all $b\ge 0$. Fix some $b\in\mathbb{N}$, and take $n>b$. The equation $\rho^{(n)}=\rho^{(n)}\mathsf{L}^{(n)}$ is equivalent to

$$\rho^{(n)}_b=\sum_{a=b-1}^{n}\rho^{(n)}_a\mathsf{L}^{(n)}_{a,b}\text{ for } b<n \tag{2.22}$$

since the conservation of the last entry $\rho^{(n)}_n$ of $\rho^{(n)}$ is ensured by the conservation of the others. Hence, (2.22) is equivalent to

$$\eta^{(n)}_b=\sum_{a=b-1}^{n}\eta^{(n)}_a\mathsf{L}^{(n)}_{a,b}\text{ for } b<n$$

and since $\mathsf{L}^{(n)}_{a,b} = \mathsf{L}_{a,b}$ when $b < n$, this is equivalent to

$$\eta^{(n)}_b = \sum_{a=b-1}^{+\infty} \eta^{(n)}_a \mathsf{L}_{a,b} \mathbf{1}_{a\le n} \text{ for } b < n \tag{2.23}$$

By (b), for all a, $\eta^{(n)}_a \mathsf{L}_{a,b}\mathbf{1}_{a\le n} \to \eta_a \mathsf{L}_{a,b}$ when $n \to +\infty$, and (a) allows us to use Lebesgue dominated convergence theorem, from which we deduce that the right-hand side of (2.23) converges to $\sum_{a=b-1}^{+\infty} \eta_a \mathsf{L}_{a,b}$, while the left-hand side goes to η_b by (b), which is the wanted relation. □

Proposition 2.21. *Let $\rho^{(n)}$ be the invariant probability distribution of $\mathsf{L}^{(n)}$. If $\rho^{(n)}$ converges weakly to some probability measure ρ on $\mathbb{N}$, then ρ is invariant by L.*

Proof. Consider (2.22) which is equivalent to $\rho^{(n)}\mathsf{L}^{(n)} = \rho^{(n)}$. It suffices to establish that for every $b \in \mathbb{N}$,

$$\sum_{a=b-1}^{n} \rho^{(n)}_a \mathsf{L}_{a,b} \underset{n\to+\infty}{\longrightarrow} \sum_{a=b-1}^{+\infty} \rho_a \mathsf{L}_{a,b}. \tag{2.24}$$

Take a small $\varepsilon > 0$. As a measure over $\mathbb{N}$, ρ is tight: there exists K such that $\rho_0 + \cdots + \rho_K > 1-\varepsilon$. Take now n large enough, so that $\rho^{(n)}_0 + \cdots + \rho^{(n)}_K > 1 - 2\varepsilon$, so that $\sum_{j>K} \rho^{(n)}_j \le 2\varepsilon$. Since $\mathsf{L}_{a,b} \le 1$, for all b,

$$\left| \sum_{a=b-1}^{n-1} \rho^{(n)}_a \mathsf{L}_{a,b} - \sum_{a=b-1}^{+\infty} \rho_a \mathsf{L}_{a,b} \right| \le \left| \sum_{a=b-1}^{K} \rho^{(n)}_a \mathsf{L}_{a,b} - \sum_{a=b-1}^{K} \rho_a \mathsf{L}_{a,b} \right| + 2\varepsilon$$

(with the empty sum being equal to 0, when $b+1 > K$). Now, since $\rho^{(n)} \to \rho$ weakly, the r.h.s. is smaller than 3ε for n large enough. □

Remark 2.22. *Tridiagonal transition matrices are ◺; some work is needed to see that the results of this section apply to the tridiagonal case (see Section 4.1).*

2.3 Remaining proofs

2.3.1 A second proof of Theorem 2.1

Consider U an irreducible ◹ transition matrix. Since the transition matrix $\mathsf{U}^{(n)}$ is finite, we have

$$\rho^{(n)}_a = \alpha_n \det\left(\mathsf{Id} - \mathsf{U}^{(n)\,\mathsf{dep}(a)}\right). \tag{2.25}$$

Here α_n is the only constant making of $\rho^{(n)}$ a probability distribution.

Now, we claim that for some constants α'_n, α''_n, for all $a \in [0, n-1]$,

$$\det\left(\mathsf{Id} - \mathsf{U}^{(n)\,\mathsf{dep}(a)}\right) = \alpha'_n \det\left(\mathsf{Id} - \mathsf{U}_{[0,a-1]}\right) \prod_{j=a+1}^{n} \mathsf{U}_{j,j-1} \tag{2.26}$$

$$= \alpha''_n \det\left(\mathsf{Id} - \mathsf{U}_{[0,a-1]}\right) / \prod_{j=1}^{a} \mathsf{U}_{j,j-1}. \tag{2.27}$$

The invariant distribution $\pi = [\pi_a, 0 \le a \le n]$ of any irreducible transition matrix M indexed by $[0,n]$ is proportional to $\det(\mathsf{Id} - \mathsf{M}^{\mathsf{dep}(a)})$ (Section 1.3.1). If such M is a ◹ transition matrix, then each tree rooted at $a \in [0,n]$ can be decomposed into two parts: a branch $n \mapsto n-1 \mapsto \cdots \mapsto a$ "above a", and a forest with set of roots on $[a,n]$, and other vertices on $[0, a-1]$, so that, as explained in Proposition 1.4 leads to (2.26) (since $\mathsf{U}^{(n)}$ is a ◹-transition matrix on a finite state space, and since $\prod_{j=a+1}^{n} \mathsf{U}^{(n)}_{j,j-1} = \prod_{j=a+1}^{n} \mathsf{U}_{j,j-1}$). Formula (2.27) is obtained by dividing (2.26) by the constant (depending only on n) $\prod_{j=1}^{n} \mathsf{U}^{(n)}_{j,j-1} = \prod_{j=1}^{n} \mathsf{U}_{j,j-1}$.

Using (2.25), (2.26) and (2.27), one sees that

$$\frac{\rho_a^{(n)}}{\alpha_n \alpha''_n} = \det\left(\mathsf{Id} - \mathsf{U}_{[0,a-1]}\right) / \prod_{j=1}^{a} \mathsf{U}_{j,j-1},$$

so that this constant sequence converges when $n \to +\infty$. Hence, Theorem 2.13(i) applies: the measure $(\det\left(\mathsf{Id} - \mathsf{U}_{[0,a-1]}\right) / \prod_{j=1}^{a} \mathsf{U}_{j,j-1}, a \geq 0)$ is invariant by U.

2.3.2 Proof of Proposition 2.3

First $R = \begin{bmatrix} 1, & 1, & 1 \cdots \end{bmatrix}$ is a right eigenvector of any matrix U in ◹.

Infinite dimensional cone We will construct explicitly a class of matrices U in ◹, for which the right eigenspace of U, associated with the eigenvalue 1 contains the cone generated by an infinite number of linearly independent vectors with positive coordinates (so that Lemma 2.23 implies the infinite-dimensional case statement). Set

$$\mathsf{U}_{0,0} = \mathsf{U}_{0,1} = 1/2,$$

and for the next rows, only three entries are not zeros:

$$\mathsf{U}_{c,c-1} = d_c,\ \mathsf{U}_{c,c} = 1 - d_c - u_c,\ \mathsf{U}_{c,2c} = u_c, \quad \text{for } c \geq 1.$$

Set, for all $m \geq 1$,

$$\Gamma_m := \prod_{c=1}^{m-1} \left(1 + 2\frac{d_c}{u_c}\right). \tag{2.28}$$

Lemma 2.23. *If (d_c) and (u_c) are positive sequences, such that $d_c + u_c < 1$, and*

$$\Gamma_\infty := \lim_m \Gamma_m < +\infty$$

then U is irreducible, and the cone generated by the right eigenvectors of U with positive coordinates, associated with the eigenvalue 1, is infinite-dimensional.

Proof. A vector $V = (V_a, a \geq 0)$ is a right eigenvector of U with eigenvalue 1, iff

$$V_0 = V_1, \text{ and } V_{2c} = (1 + d_c/u_c)V_c - (d_c/u_c)V_{c-1}, \text{ for all } c \geq 1. \tag{2.29}$$

It implies that the entries of V with even indices $2c$ are functions of the entries with smaller indices, and, better than that, are functions of V_0 and of the entries with odd smaller indices V_{2p+1}. As a consequence, $V_0 = V_1$ and the entries $(V_{2p+1}, p > 1)$ provide a parametrization of the eigenspace we are looking for, but some additional work is needed, to prove the statement concerning the cone generated by positive vectors.

We first establish that there exists an infinite dimensional set of linearly independent vectors that are bounded, and satisfies (2.29) (we drop the positivity constraint, for a small moment).

Since (2.29) allows one to determine the even indices V_{2p} (for $p \geq 1$) using the odd ones (plus V_0), let us assume that $V_0 = V_1 = 1$ and all the free parameters $(V_{2p+1}, p \geq 0)$ satisfies:

$$|V_{2p+1}| \leq \Gamma_{2p+1} \text{ for all } p \geq 0. \tag{2.30}$$

We claim that in this case for all $i \geq 0$ (odd or even, the even ones being computed thanks to (2.29)), we have $|V_i| \leq \Gamma_i$. The claim can be proved by induction: let P_C be the property, $|V_i| \leq \Gamma_i$ for all $i \leq C$. Then P_0 holds. Assume that P_C holds for some $C \geq 0$. Then either $C+1$ is odd, and by (2.30), P_{C+1} holds; either $C+1 = 2c$ is even by (2.29),

$$\begin{aligned}|V_{2c}| &\leq (1 + d_c/u_c)|V_c| + (d_c/u_c)|V_{c-1}| \\ &\leq (1 + 2d_c/u_c)\Gamma_c = \Gamma_{c+1} \leq \Gamma_{2c}\end{aligned}$$

since $(\Gamma_c, c \geq 0)$ is increasing and since $c+1 \leq 2c$ for $c \geq 1$, so that, again, P_{C+1} holds.

This argument allows proving that the (right) eigenspace (associated with the eigenvalue 1) contains an infinite family[2] of linearly independent vectors $(V^i, i \in I)$, whose coordinates are bounded in absolute value by Γ_∞, and then, using that $R = \begin{bmatrix}1, & 1, & 1 \cdots\end{bmatrix}$ is in this space too, the vectors $(2\Gamma_\infty R + V^i, i \in I)$ form an infinite family of linearly independent vectors, with positive coordinates belonging to this eigenspace. □

Finite dimensional case Fix $m \geq 1$: we will give a ◹ matrix U, whose dimension of the cone of positive vector stable by U has dimension $m+1$, where the "+1" counts the vector R, so that we will design a matrix U having an additional set of m linearly independent vectors exactly (the case $m = 0$ is fulfilled by the tridiagonal case, in which there is a unique right eigenvector associated with the eigenvalue 1, since the system $TV = V$ is triangular). We will use the tools already discussed above in the infinite case. Consider this time the kernel K defined by $\mathsf{U}_{0,0} = \mathsf{U}_{0,1} = 1/2$, and whose subsequent rows, only three entries are not zeros:

$$\mathsf{U}_{c,c-1} = d_c, \;\; \mathsf{U}_{c,c} = 1 - u_c - d_c, \;\; \mathsf{U}_{c,c+1+m} = u_c.$$

Any (right) eigenvector V of U associated with the eigenvalue 1, satisfies $V_0 = V_1 = V_{2+m} = 1$. Again, the m-entries $V_2, \ldots, V_{1+m}$ will appear to be free parameters, since, for $c \geq 2$

$$V_{c+1+m} = (1 + d_c/u_c)V_c - (d_c/v_c)V_{c-1}.$$

Now, taking the free parameters $|V_i| \leq \Gamma_i$ (for i from 2 to $1+m$), we can again show, by iteration that all V_i satisfy $|V_i| \leq \Gamma_i$, and we conclude as in the infinite case.

This ends the proof of Proposition 2.3.

3 Connections between almost upper and lower triangular cases

According to Theorem 2.1, ◹ transition matrices always have an invariant measure, while it is not the case for ◺ transition matrices (Theorem 2.14). The next theorem says that one can associate with each ◹ transition matrix U, another matrix L of type ◺, "its time-reversal".

Theorem 3.1. *Consider an irreducible ◹ transition matrix $\mathsf{U} = \left[\mathsf{U}_{i,j}\right]_{0 \leq i,j}$, with invariant measure π, then set $\mathsf{L} = \left[\mathsf{L}_{i,j}\right]_{0 \leq i,j}$ as*

$$\mathsf{L}_{i,j} = \pi_j \mathsf{U}_{j,i}/\pi_i. \tag{3.1}$$

We have

(i) L is an irreducible ◺ transition matrix on $\mathbb{N}$, with invariant measure π too.

[2]The subspace of the eigenspace associated with the eigenvalue 1 that we described, is parametrized by the vectors V that we gave, which are functions of the values V_{2p+1} (with odd indices) that are taken in intervals: the generated space is infinite-dimensional.

(*ii*) L *is recurrent if and only if* U *is recurrent,*

(*iii*) *If* π *is a probability distribution then, if* $(Y_k, k \in \mathbb{Z})$ *is a* U*-Markov chain under its stationary regime (meaning that* $Y_k \sim \pi$ *for any* $k \in \mathbb{Z}$*), then the time-reversal of this chain,* $(Y_{-k}, k \in \mathbb{Z})$ *is a* L*-Markov chain under its stationary regime.*

(*iv*) L *is positive recurrent if and only if* U *is positive recurrent.*

(*v*) *If* U *has a right eigenvector* $R = \left[R_k, k \geq 0\right]$ *with positive coordinate (associated with the eigenvalue 1) then* L *admits* $\left[\pi_k R_k, k \geq 0\right]$ *as invariant measures: more generally, if* U *has* k *linearly independent eigenvectors (associated with the eigenvalue 1, having positive coordinates), then* L *admits* k *linearly independent invariant measures (recall that the column vector with coordinates equal to 1, is such an eigenvector).*

(*vi*) *If* U *has several linearly independent positive right eigenvectors (associated with the eigenvalue 1), then* U *is not recurrent, and neither is* L*.*

Notice that in (iii), only positive recurrent matrices have an invariant probability measure.

Remark 3.2. *As a consequence of the* (ii) *of the Theorem, and of Formulas (2.8), (2.17) and (2.18),*

$$\mathsf{U}_{1,0}\frac{\det(\mathsf{Id}-\mathsf{U}_{[2,b-1]})}{\det(\mathsf{Id}-\mathsf{U}_{[1,b-1]})} \xrightarrow[b\to\infty]{} 1 \Leftrightarrow \frac{\prod_{j=1}^{b-1}\mathsf{L}_{j,j+1}}{\det(\mathsf{Id}-\mathsf{L}_{[1,b-1]})} = \frac{\det(\mathsf{Id}-\mathsf{L}_{[0,b-1]})}{\mathsf{L}_{0,1}\det(\mathsf{Id}-\mathsf{L}_{[1,b-1]})} \xrightarrow[b\to\infty]{} 0, \quad (3.2)$$

but these formulas are in nature, a bit different. See Remark 3.7 to explore further "what is equal".

Proof of Theorem 3.1. (i): straightforward.

(ii): for an irreducible U-Markov chain $(Y_j, j \geq 0)$, recurrence is equivalent to $\mathbb{P}(\tau_{\{0\}}(Y) < +\infty \mid Y_0 = 0) = 1$. This means that the total "U-weights" of the paths in the set $\cup_{k\geq 0}\{(x_0 = 0, x_1, \ldots, x_k, x_{k+1} = 0), x_i > 0, i \in [1,k]\}$ is 1 when the "U-weight" of a given path $(x_0, \ldots, x_{k+1})$ is defined to be $\prod_{j=0}^{k}\mathsf{U}_{x_j,x_{j+1}}$. Since such paths start and end at 0, then their L-weights and U-weights coincide.

(iii): by translation invariance, it suffices to write

$$\mathbb{P}(Y_k = y_k, 0 \leq k \leq a) = \pi_{y_0}\prod_{j=0}^{a-1}\mathsf{U}_{y_j,y_{j+1}} = \pi_{y_a}\prod_{j=0}^{a-1}\mathsf{L}_{y_{j+1},y_j} = \mathbb{P}(Z_{a-j} = y_j, 0 \leq j \leq a)$$

for Z an L-Markov chain under its invariant regime.

(iv): by (i), both L and U have the same invariant probability measure (which implies the statement).

(v): Assume that $R = [R_i, i \geq 0]$ with $R_0 := 1$ is a right eigenvector of U, associated with the eigenvalue 1, then $\sum_b U_{a,b}R_b = R_a$ for $a \geq 0$, which by (3.1) gives

$$\sum_b \pi_b R_b \,\mathsf{L}_{b,a} = \pi_a R_a, \quad a \geq 0;$$

in other words, L possesses $[\pi_a R_a, a \geq 0]$ as an invariant measure. Since π has non-zero entries, the map $R \mapsto [\pi_a R_a, a \geq 0]$ is a linear bijection, and the conclusion follows.

(vi): By (v), if U has several linearly independent, positive right eigenvectors associated with the eigenvalue 1, then L has several invariant measures, so that it is not recurrent (see e.g. Norris [18, Theo. 2.2, p. 102] or Brémaud [3, Theo. 3.2.3, p. 119]), and by (ii) neither is U. □

Theorem 3.3. *The ◺ transition matrix* L *admits a time-reversal transition matrix* U *if and only if it possesses a positive invariant measure* η *in which case* $\mathsf{U}_{b,a} = \eta_a \mathsf{L}_{a,b}/\eta_b$*, and* U *and* L *are both time-reversal of each other. As a consequence, for each* L*, there is a linear bijection between the set of classes of invariant measures of* L *(in* $\mathcal{M}_{\mathbb{N}}^{+,\sim}$*) and the set of time-reversal transition matrices* U*.*

The proof of this theorem is simple since any such U is the time-reversal of L, but it exists only when the positive invariant measure η exists; as explained in Theorem 2.14, some irreducible ◺ do not admit any positive invariant measure.

Remark 3.4. *This Theorem, which is deeply connected to Theorem 3.1 allows making the following observation. A ◹ transition matrix* U *possesses a unique invariant measure* π*, and then, a unique time reversal ◺ transition matrix* L*: let us say that* (π, L) *is the time reversal of* (π, U)*. Now, consider* $\mathsf{Right}(\mathsf{U})$ *the set of* <u>*right*</u> *eigenvectors of* U *with positive coordinates associated with the eigenvalue 1. The set* $\mathsf{Left}(\mathsf{L})$ *of invariant measures of* L *is* $\{[\pi_j R_j, j \geq 0], R \in \mathsf{Right}(\mathsf{U})\}$*. It is then natural to ask, in this case, what are the time reversals of* L *and their link with "the initial* U*". The time reversal of* (π', L) *with* $\pi'_j = \pi_j R_j$*, is* (π', U') *with* $\pi'_i \mathsf{U}'_{i,j} = \pi'_j \mathsf{L}_{j,i}$ *for all* i, j*. Injecting* $\mathsf{L}_{j,i} = \pi_i \mathsf{U}_{i,j}/\pi_j$ *provides* $\pi_i R_i \mathsf{U}'_{i,j} = \pi_j R_j \pi_i \mathsf{U}_{i,j}/\pi_j$ *so that*

$$\mathsf{U}'_{i,j} = \mathsf{U}_{i,j}\frac{R_j}{R_i}, \quad \text{for all } i, j.$$

To compute the set of time reversals of (π, L) *where* π *is taken in the set of invariant measures of* L*, it suffices then to know a single element* π *of this set, to compute the time reversal* (π, U) *of* (π, L) *(by* $\mathsf{U}_{i,j} = \pi_j \mathsf{L}_{j,i}/\pi_i$*), and then to compute the right eigenvectors of* U*. This observation entails that there is a correspondence between the set of ◺ transition matrices* L *possessing an invariant measures, and the set of ◹ transition matrices quotient by the equivalence relation* $\sim$*, defined by* $\mathsf{U} \sim \mathsf{U}'$ *if* U *possesses a* <u>*right*</u>*-eigenvector* R *with positive coordinate, associated with the eigenvalue 1, such that* $U'_{i,j} = \mathsf{U}_{i,j}\frac{R_j}{R_i}$ *for all* i, j*.*

3.1 Algebraic connection between U and L

Consider (U, L) a pair of irreducible transition matrices where U is ◹, L is ◺, and assume that they are time-reversal of each other. The invariant measure π of U is unique, so that $\mathsf{L}_{i,j} = \pi_j \mathsf{U}_{j,i}/\pi_i$, $\forall i, j \geq 0$. A simple expansion of the determinant using the cycle's decomposition of permutations, gives, for every $a, b \geq 0$:

$$\det\left(\mathsf{Id} - \mathsf{U}_{[a,b]}\right) = \det\left(\mathsf{Id} - \mathsf{L}_{[a,b]}\right). \tag{3.3}$$

Apart from this formula, the main relation is

$$\mathsf{L}_{a,b} = \pi_b \mathsf{U}_{b,a}/\pi_a = \mathsf{U}_{b,a}\frac{\det\left(\mathsf{Id} - \mathsf{U}_{[0,b-1]}\right)/\prod_{j=1}^{b}\mathsf{U}_{j,j-1}}{\det\left(\mathsf{Id} - \mathsf{U}_{[0,a-1]}\right)/\prod_{j=1}^{a}\mathsf{U}_{j,j-1}}. \tag{3.4}$$

If U is known, and the corresponding L is searched, then this last formula, built using Theorem 2.1 allows computing it. On the other hand, if L is known, but not U, this is more difficult since we have no simple expression of π in terms of L (and again, the existence and uniqueness of π are not assured).

The following proposition provides some relations between the elements in the tuple $(\pi, \mathsf{U}, \mathsf{L})$.

Proposition 3.5. *For any $b \geq 0$, set*

$$Z_b := \prod_{j=1}^{b} \frac{\mathsf{L}_{j-1,j}}{\mathsf{U}_{j,j-1}}, \quad Z'_b := \prod_{j=1}^{b} \frac{\mathsf{U}_{j-1,j}}{\mathsf{L}_{j,j-1}},$$

(where $Z_0 = Z'_0 = 1$, which is compatible with the convention concerning empty products),

(i) For any $a \geq 0$, $\det(\mathsf{Id} - \mathsf{L}_{[0,a-1]}) = \prod_{j=0}^{a-1} \mathsf{L}_{j,j+1}$ (with the convention, $\det(\mathsf{Id} - \mathsf{L}_{[0,-1]}) = 1$).

(ii) For any $b \geq 0$, $Z_b = Z'_b$.

(iii) The measure $(Z_0, Z_1, Z_2, \dots)$ is invariant by both U and L.

The point (ii) of Proposition 3.5 is equivalent to

$$\mathsf{L}_{a,a-1}\mathsf{L}_{a-1,a} = \mathsf{U}_{a-1,a}\mathsf{U}_{a,a-1}. \tag{3.5}$$

Proof. Taking $b = a+1$ in (3.4), gives

$$\mathsf{L}_{a,a+1} = \det\left(\mathsf{Id} - \mathsf{U}_{[0,a]}\right) / \det\left(\mathsf{Id} - \mathsf{U}_{[0,a-1]}\right) \tag{3.6}$$

from what we infer (i) (using (3.3)).

For $k \geq 1$, using (3.4) we get

$$\mathsf{L}_{b+k,b} \prod_{j=b+1}^{b+k} \mathsf{L}_{j-1,j} = \mathsf{U}_{b,b+k} \prod_{j=b+1}^{b+k} \mathsf{U}_{j,j-1}; \tag{3.7}$$

for $k = 1$ this provides (ii). Further, this equation rewrites $\mathsf{L}_{b+k,b} = Z_b\, \mathsf{U}_{b,b+k} \,/\, Z_{b+k}$, and since for $k = -1$ this is valid too (since $\mathsf{L}_{b-1,b} = Z_b \mathsf{U}_{b,b-1}/Z_{b-1} = U_{b,b-1}U_{b-1,b}/\mathsf{L}_{b,b-1}$, and this is true by (ii)) we have for all a, b, $\mathsf{L}_{a,b} = Z_b\, \mathsf{U}_{b,a} \,/\, Z_a$, which ensures (iii). □

Remark 3.6. *By Proposition 3.5, when L is known, each of the sequences $(\pi_a, a \geq 0)$ (up to a multiplicative constant), $(\mathsf{U}_{a,a-1},\ a \geq 1)$,$(\mathsf{U}_{a,a+1},\ a \geq 0)$, $((\mathsf{U}_{i,j},\ j > i), i \geq 0)$ allows computing the others. For example, if $(\mathsf{U}_{a,a+1}, a \geq 0)$ is known as well as L, (3.5) allows one to compute $(\mathsf{U}_{a,a-1}, a \geq 1)$, and then equation (3.7) let us obtain $((\mathsf{U}_{i,j}), j > i, i \geq 0)$ and $(Z_b, b \geq 0)$ which is proportional to π.*

Remark 3.7. *By reversibility, it can be seen that $u_b(x)$ and $1 - v_b(x)$ are not equal in general. This comes from a lack of symmetry in the measured event. If instead one observes the return time to 0 by random walks starting at 0, the symmetry comes back, but the formulas are more complex. Denote by Y^{U} and Y^{L} Markov chains with respective transition matrices U and L such that $\pi_a \mathsf{L}_{a,b} = \pi_b \mathsf{U}_{b,a}$. One has*

$$\mathbb{P}(\tau_{\{0\}}(Y^{\mathsf{U}}) < \tau_{[b,+\infty)}(Y^{\mathsf{U}}) \mid Y_0^{\mathsf{U}} = 0) = \mathbb{P}(\tau_{\{0\}}(Y^{\mathsf{L}}) < \tau_{\{b\}}(Y^{\mathsf{L}}) \mid Y_0^{\mathsf{L}} = 0). \tag{3.8}$$

Formula (3.7) allows seeing that the weight of a cycle $(a, a+b, a+b-1, \dots, a+1, a)$ for Y^{U} is the same as the weight of the cycle $(a, a+1, \dots, a+b, a)$ for Y^{L} (which allows proving (3.8), using combinatorial techniques).

3.2 Some spectral properties of ◺ and ◹ transitions matrices

The spectral properties of infinite ◺ and ◹ transition matrices appear to be rather interesting since their left and right eigenvalues are in general different and have different multiplicities.

We already viewed that ◹-irreducible matrices have a unique <u>left</u> eigenvector associated with the eigenvalue 1 (up to the multiplicative constant), but could have one or more (including $+\infty$ linearly independent) <u>right</u> eigenvectors associated with the eigenvalue 1. Matrices of type ◺ may have zero or several invariant measures, when, 1, seen as <u>right</u> eigenvalue has multiplicity 1.

The next theorem shows that ◹-transition matrices have a complete spectrum on the <u>left</u>-hand side, and ◺-ones, a complete spectrum on the <u>right</u>-hand side, with, in both cases, totally explicit eigenvectors.

Theorem 3.8. *(i) If U is an irreducible ◹-transition matrix on $\mathbb{N}$, then U admits all $\Lambda \in \mathbb{C}$ as a simple left eigenvalue with corresponding eigenvector $\pi(\Lambda) := \left[\pi_a(\Lambda), a \geq 0\right]$ unique (up to a constant factor $\pi_0(\Lambda) \neq 0 > 0$), defined by*

$$\pi_a(\Lambda) := \pi_0(\Lambda)\, \frac{\det(\Lambda \mathsf{Id} - \mathsf{U}_{[0,a-1]})}{\prod_{j=1}^{a} \mathsf{U}_{j,j-1}}, \quad a \geq 1. \tag{3.9}$$

(The vector $\pi(1)$ is the unique invariant measure of U, up to a multiplicative constant.)

(ii) If L is an irreducible ◺-transition matrix on $\mathbb{N}$, then for all $\Lambda \in \mathbb{C}$, $\pi'(\Lambda) := \left[\pi'_a(\Lambda), a \geq 0\right]$ defined by

$$\pi'_a(\Lambda) := \pi'_0(\Lambda)\, \frac{\det(\Lambda \mathsf{Id} - \mathsf{L}_{[0,a-1]})}{\prod_{j=1}^{a} \mathsf{L}_{j-1,j}}, \quad a \geq 1, \tag{3.10}$$

is the unique (up to a constant factor $\pi'_0(\Lambda) \neq 0$) right eigenvector of L associated with the simple eigenvalue Λ.

(iii) If L is a ◺-transition matrix time reversal of an irreducible ◹ transition matrix U (that is, $\mathsf{L}_{i,j} = \pi_j \mathsf{U}_{j,i}/\pi_i$, for all i, j) then $\left[\pi_j(\Lambda)/\pi_j(1), j \geq 0\right]$ is collinear to $\pi'(\Lambda)$.

(iv) Assume again that L is a ◺-transition matrix time reversal of an irreducible ◹ transition matrix U. For all $\Lambda \in \mathbb{C}$, the map

$$\begin{array}{rccc} \Psi : & R^{\mathbb{N}} & \longrightarrow & \mathbb{R}^{\mathbb{N}} \\ & v & \longmapsto & \left[v_j \pi_j(1), j \geq 0\right] \end{array}$$

is a linear map which sends the <u>right</u> eigenspace of U associated with the eigenvalue Λ (that is, the set $\{v : \mathsf{U}v = \Lambda v\}$, possibly reduced to $\{0\}$), to the <u>left</u> eigenspace of L associated with the same eigenvalue.

Notice that (iv) applies only to the matrices L having a time reversal.

Remark 3.9. *For any $\Lambda \in \mathbb{C}$, there exists some ◹ transition matrix U having Λ as right eigenvalue, with eigenspace having dimension k for all $k \geq 0$ (including $+\infty$). The proof can be simply adapted from that of Lemma 2.23.*

We first state the following analogue of Lemma 2.7 (which can be proved with the same argument):

Lemma 3.10. *Let $s \in \mathbb{N} \cup \{+\infty\}$ be a size parameter. If $\mathsf{U} = \left[\mathsf{U}_{i,j}\right]_{0 \leq i,j \leq s < +\infty}$ is a finite irreducible ◹ stochastic matrix with size s and eigenvalue Λ, or if $\mathsf{U} = \left[\mathsf{U}_{i,j}\right]_{0 \leq i,j \leq s = +\infty}$ is an infinite irreducible ◹-stochastic matrix, then for all $y \leq s$*

$$\begin{aligned} \det(\Lambda \mathsf{Id} - \mathsf{U}_{[0,y]}) &= \det(\Lambda \mathsf{Id} - \mathsf{U}_{[0,y-1]})\,(\Lambda - \mathsf{U}_{y,y}) \\ &\quad - \sum_{x \leq y-1} \det(\Lambda \mathsf{Id} - \mathsf{U}_{[0,x-1]})\,\mathsf{U}_{x,y} \prod_{j=x+1}^{y} \mathsf{U}_{j,j-1}, \end{aligned} \tag{3.11}$$

with the convention $\det(\Lambda \mathsf{Id} - \mathsf{U}_{[0,-1]}) = 1$.

Proof of Theorem 3.8. (i) Take $\Lambda \in \mathbb{C}$ and let us solve the system $v\mathsf{U} = \Lambda v$ with unknown, the vector $v = (v_j, j \geq 0)$. The system $\{\Lambda v_j = \sum_{i \leq j+1} v_i \mathsf{U}_{i,j}$ for $j \geq 0\}$, is a triangular system on (v_j), and v_{j+1} can be expressed uniquely in terms of the previous v_j:

$$v_{j+1} = (\Lambda v_j - \sum_{i \leq j} v_i)/\mathsf{U}_{j+1,j}, \quad j \geq 0$$

and this leads immediately to the existence and uniqueness of a unique solution v up to a multiplicative constant. Now, let us establish that $\pi(\Lambda)$ given in (3.9), is such a solution, that is, it satisfies $\sum_x \pi_x(\Lambda)\mathsf{U}_{x,y} = \Lambda \pi_y(\Lambda)$ for $y \geq 1$; the proof is a simple adaptation of that of Theorem 2.1 (with Lemma 3.10 replacing Lemma 2.7 in the argument).

(ii) Notice that in the proof of (i), we did not use that U is a transition matrix, but just that $\mathsf{U}_{j+1,j} \neq 0$. The proof of (ii) is then very similar to the proof of (i): the system $\mathsf{L}c = \Lambda c$ (for a column vector c) is equivalent to ${}^t c^t \mathsf{L} = \Lambda^t c$ (where ${}^t A$ is the transposed matrix of A).

(iii) If v is a left eigenvector of U with eigenvalue Λ, then $v\mathsf{U} = \Lambda v$ is equivalent to $\sum_i v_i \mathsf{U}_{i,j} = \Lambda v_j, \forall j$, and since $\pi_j \mathsf{L}_{j,i}/\pi_i = \mathsf{U}_{i,j}$, $\sum_i \mathsf{L}_{j,i} v_i/\pi_i = \Lambda v_j/\pi_j$ which means that $\left[v_i/\pi_i, i \geq 0\right]$ is a right eigenvector associated with the eigenvalue Λ. By uniqueness of the right eigenvector of L up to a multiplicative constant, the conclusion follows.

(iv) First, the map Ψ is clearly a bijective linear map on $\mathbb{R}^{\mathbb{N}}$. Now $\mathsf{U}R = \Lambda R$ is equivalent to $\sum_j \mathsf{U}_{i,j} R_j = \Lambda R_i = \sum_j \mathsf{L}_{j,i} \pi_j R_j/\pi_i$ (for all $i \geq 0$), which is equivalent to the fact that $\Psi(R) = [\pi_j R_j, j \geq 0]$ is in the left eigenspace of L associated with Λ. □

3.2.1 Sub-stochastic almost triangular matrices

A finite or infinite square matrix is said to be sub-stochastic if its coefficients are non-negative, and if the sums on each row belong to $[0, 1]$. The notion of irreducibility extends to sub-stochastic matrices. By Perron-Frobenius theorem, a finite irreducible sub-stochastic matrix M possesses a simple positive eigenvalue $R \leq 1$ (called the Perron-Frobenius eigenvalue) which is strictly greater than all other eigenvalues modulus (and $R < 1$, if at least one row, of M has a sum in $(0, 1)$).

Proposition 3.11. *Theorem 3.8*(i) *and* (ii) *and Lemma 3.10 extend to the sub-stochastic case (that is, their statements remain true if* U *and* L *are sub-stochastic instead of stochastic).*

Proposition 3.12. *If* U *(resp.* L*) is a* <u>*finite*</u> *irreducible sub-stochastic* ◹ *(resp.* ◺*) transition matrix (with columns indexed from 0 to* s*), and* Λ *is an eigenvalue of* U *(resp.* L*), then* Λ *is a simple eigenvalue of* U *(resp.* L*), and* π *(resp.* π'*) as defined in Theorem 3.8*(i) *is an associated eigenvector.*

Proof. The proof is the same as in the proof of the finite case of Theorem 2.1. □

We think that this extension to the sub-stochastic case is interesting for two reasons. The first one is that, sub-stochastic almost upper triangular matrices form a large class of models for which, as already said, a close formula can be provided. The second reason is that the truncation $\mathsf{U}_{[0,n]}$ of a ◹ transition matrix U is sub-stochastic. As explained in Seneta [22, Chap. 7], larger and larger truncations of infinite matrices U can be used to study the behaviour of Markov chains with transition matrices U, under some suitable hypothesis, and it is worth saying a few words concerning this aspect (as suggested by a referee of the paper). This second point has to be tempered a bit, since Theorem 2.1 provides a close formula for “the non-truncated chain”, and these close formulas do not need the computation of any eigenvalue, which made them simpler to study than truncated versions.

Proposition 3.11 allows giving a third proof of Theorem 2.1, but only under the additional hypothesis that U is recurrent. Consider $\mathsf{U}^{<}(n) = \mathsf{U}_{[0,n]}$ as the nth truncation of U (for all $n \geq 0$). We prefer to use $\mathsf{U}^{<}(n)$ instead of $\mathsf{U}_{[0,n]}$, because our formula needs further truncations. Call $\Lambda(n)$ and $v(n)$ its Perron eigenvalue and eigenvector.

Since $\mathsf{U}^{<}(n)$ is sub-stochastic and irreducible, Seneta [22, Theo. 6.8] implies that $\Lambda(n) \nearrow 1$ (this follows from Theorem 6.6. [22] and is a consequence of U being a stochastic recurrent matrix[3]). Now, Proposition 3.11 together with Theorem 6.9 [22] (fixing $v(n)_0 = \pi_0 = 1$) gives that

$$\begin{aligned}\lim_{n\to\infty} v(n)_a &= \lim_{n\to\infty} \frac{\det(\Lambda(n)\mathsf{Id} - \mathsf{U}^{<}(n)_{[0,a-1]})}{\prod_{j=1}^{a} \mathsf{U}^{<}(n)_{j,j-1}} = \lim_{n\to\infty} \frac{\det(\Lambda(n)\mathsf{Id} - \mathsf{U}_{[0,a-1]})}{\prod_{j=1}^{a} \mathsf{U}_{j,j-1}} \\ &= \frac{\det(\mathsf{Id} - \mathsf{U}_{[0,a-1]})}{\prod_{j=1}^{a} \mathsf{U}_{j,j-1}} = \pi_a\end{aligned}$$

but the Seneta theorem requires the recurrence of U for this conclusion, when Theorem 2.1 holds more generally.

3.3 General presentation of ◺ transition matrices using descent kernels

Markov chains with ◹ transition matrices, by time-reversal, provide ◺ transition matrices, on which much can be said. Nevertheless, if one is interested in a particular L models which does not come from such a time-reversal, it may be needed to have some tools allowing one to study them, (and, for example, compute their time reversal if they have one, to get access to the toolbox of ◹ transition matrices). A slight change of point of view on ◺ transition matrices will allow us to search more efficiently the form of their time-reversal when they exist (see Proposition 3.14), to design many ◺ transition matrices L for which it is possible to find the time-reversal (Section 3.3.2), and, finally, to revisit some known results of the literature (the so-called, catastrophe transition matrices, see Section 3.3.3). The results collected in this section are of interest to the user searching some complete families of ◺ transition matrices for which the invariant distribution is computable (for the sake of teaching, statistical purpose, or simple curiosity).

Definition 3.13. *A descent kernel $\mathsf{D} = \left[\mathsf{D}_{i,j}\right]_{i,j\geq 0}$ over $\mathbb{N}$ is a lower triangular transition matrix $\mathsf{D}_{i,j} > 0 \Rightarrow j \leq i$ (with non-negative coefficients, summing to one on each row).*

Each ◺ transition matrix L can be represented uniquely as a pair (v, D) where $v = (v_a, a \geq 0)$ is a sequence of elements of the interval $[0,1]$ (in fact, $(0,1)$ in the irreducible case, except $v_0 \in [0,1)$), and D a descent transition matrix, as follows:

$$\begin{cases} \mathsf{L}_{b,a} &= v_b\, \mathsf{D}_{b,a}, \text{ for } b \geq a, \\ \mathsf{L}_{b,b+1} &= 1 - v_b, \quad b \geq 0. \end{cases} \tag{3.12}$$

We will say that (v, D) is the descent representation of L. In other words: v_b is seen as the probability of descent from b, and D as the descent kernel, conditionally on a descent. With probability $1 - v_b$, there is an increment $+1$ (ascent).

In the literature, instead of descent kernel, the word "catastrophe" is sometimes used, but with a slightly different construction, relying instead upon a standard birth-death process, mixed with a descent kernel (in our representation, the random walker has to

[3]The one appearing as the limiting value of $\Lambda(n)$ is not defined as the Perron-Frobenius eigenvalue of the matrix U in the literature, since the natural definition that we use for the finite case becomes pathological; for example, it does not make sense to define it as the maximum (modulus) eigenvalue since Theorem 3.8 says that all complex values are eigenvalues of U. In the Seneta language (see definition 6.1 [22]), this limiting value is defined as the reciprocal of a radius of convergence, called convergence parameter.

choose randomly between a $+1$ step and a descent taken according to D, see (3.12)). We think that our choice, while equivalent, is more compact, and allows us to better observe the algebra into play (see e.g. Pollett & al. [20], Brockwell & al. [4], Kapodistria & al. [11], and references therein).

3.3.1 Representation of time-reversal of ◺ transition matrices (descent form)

From Theorem 3.3 we see that irreducible ◺ transition matrices L having an invariant measure and those admitting a time-reversal are the same. The representation of ◺-transition matrices L using descent kernels will allow us to have a better point of view on the form of their possible time-reversal.

Proposition 3.14. *If the set of time-reversals of an irreducible transition matrix* L *with descent representation* (v,D) *is not empty, then each of its elements* U *can be represented as follows*

$$\begin{cases} \mathsf{U}_{a,b} &=& u_a\,\alpha_a\beta_b\,\mathsf{D}_{b,a},\ \textit{for } b\geq a,\\ \mathsf{U}_{a,a-1} &=& 1-u_a, \quad a\geq 0,\ \textit{with } u_0=1\end{cases} \tag{3.13}$$

where $[\pi,u,\alpha,\beta]$ *is a 4-tuple of sequences which satisfies*

- (i) π, u, α, β *are sequences of positive real numbers, except for* β_0 *which is* 0 *iff* $v_0=0$*; moreover* $u_0=1$ *and* $u_j\in(0,1)$ *for* $j\geq 1$*,*
- (ii) $\sum_{b:b\geq a}\alpha_a\beta_b\mathsf{D}_{b,a}=1$ *for all* $a\geq 0$*,*
- (iii) *for all* $b\geq 0$*,* $a\leq b$*,*

$$u_a\alpha_a=1/\pi_a,\quad \beta_b=\pi_b v_b. \tag{3.14}$$

- (iv) $\pi_a(1-v_a)=\pi_{a+1}(1-u_{a+1})$ *for all* $a\geq 0$*.*

Proof. First, assume that π,u,α,β satisfy the properties stated in the theorem. By (i) and (ii), U is a transition matrix. Let us check that $\pi_b\mathsf{L}_{b,a}=\pi_a\mathsf{U}_{a,b}$ which is sufficient to conclude (by Theorem 3.3).

– First, we have $\pi_a\mathsf{L}_{a,a+1}=\pi_a(1-v_a)=\pi_{a+1}(1-u_{a+1})=\pi_{a+1}\mathsf{U}_{a+1,a}$,
– and for $b\geq a$, $\pi_b\mathsf{L}_{a,b}=\pi_b v_b D_{b,a}=\beta_b D_{b,a}$, while $\pi_a\mathsf{U}_{a,b}=\pi_a u_a\,\alpha_a\beta_b\,\mathsf{D}_{b,a}=\pi_b v_b D_{b,a}=\pi_b\mathsf{L}_{b,a}$ (by (3.14)).

These points show that in all cases $\pi_a\mathsf{U}_{a,b}=\pi_b\mathsf{L}_{b,a}$.

Conversely, assume that U is a time-reversal of L (with representation (v,D)) for some positive measure π, that is, it satisfies $\pi_b\mathsf{L}_{b,a}=\pi_a\mathsf{U}_{a,b}$. Since U is ◹, it can be represented as $\mathsf{U}_{a,a-1}=1-u_a$ and $\mathsf{U}_{a,b}=u_aH_{a,b}$ for a sequence u and H such that $H_{a,b}>0\Rightarrow b\geq a$ (an ascent kernel). The sequence u must satisfy $\pi_b\mathsf{L}_{b,b+1}=\pi_b(1-v_b)=\pi_{b+1}\mathsf{U}_{b+1,b}=\pi_{b+1}(1-u_{b+1})$, so that (iv) holds.

Let us show that U can be represented as stated in the Theorem. First, we must have $\mathsf{L}_{b,a}=0\Leftrightarrow\mathsf{U}_{a,b}=0$. Since for all $b>0$, the factor $v_b>0$, $\mathsf{L}_{b,a}=0\Leftrightarrow D_{b,a}=0$ and since this must be equivalent to $\mathsf{U}_{a,b}=0$, it is easily seen that $\mathsf{U}_{a,b}=u_aH_{a,b}=u_a.\mathsf{D}_{b,a}.g(a,b)$ for some positive function $g(a,b)$. This way of thinking extends to $b=0$, when $v_0>0$. If $v_0=0$, then $\mathsf{L}_{0,0}=0$, and since $\mathsf{D}_{0,0}=1$ (because D is a descent transition matrix), to satisfy $\mathsf{U}_{0,0}=u_0H_{0,0}=u_0.\mathsf{D}_{0,0}.g(0,0)=0$ too, we will take $g(0,0)=0$ (in fact $\beta_0=0$ will be the needed specification). Now, for all $b\geq a$ write

$$\pi_b v_b\mathsf{D}_{b,a}=\pi_a\mathsf{U}_{a,b}\Leftrightarrow\mathsf{U}_{a,b}=\pi_b v_b\mathsf{D}_{b,a}/\pi_a, \tag{3.15}$$

and then if $b>0$, the variables in factor to $\mathsf{D}_{b,a}$ are functions of separated variables a or b, so that $g(a,b)=\alpha_a\beta_b$ for some sequences α and β. Set $\alpha_a=1/(u_a\pi_a)$, $\beta_b=\pi_b v_b$, and for this choice, $\mathsf{U}_{a,b}=u_a\alpha_a\beta_b\mathsf{D}_{b,a}$, so that (iii) and (i) hold. It remains to check (ii). Since U is the time-reversal of L, we get $\sum_b\pi_a\mathsf{U}_{a,b}=1$ and then $\sum_{b:b\geq a}\pi_a\mathsf{U}_{a,b}=\pi_a u_a$ which implies that $\sum_{b:b\geq a}\pi_a u_a\alpha_a\beta_b\mathsf{D}_{b,a}=\pi_a u_a$ and then $\sum_{b:b\geq a}\alpha_a\beta_b\mathsf{D}_{b,a}=1$. □

3.3.2 Catalytic inversion of ◺ transition matrices

In this section, we introduce a tool allowing one to design many ◺-transition matrices L with a computable invariant measure (and computable time-reversal transition matrices U). The weakness of this approach is that it is far more efficient when, instead of fixing a given L in terms of its descent representation (v, D) only (v_0, D) is fixed. By this method, finding a complete descent kernel (v, D) with a computable invariant measure amounts to finding a positive sequence X satisfying some inequalities:

Definition 3.15. *Consider a pair (v_0, D), where $v_0 \in [0, 1)$ and D is a descent kernel. A sequence $X = (X_a, a \geq 0)$ is said to be (v_0, D) pushable iff the following three conditions are satisfied:*

- *$X_0 = v_0$ and, for all $i \geq 1$, $X_i > 0$ (so that $X_0 = 0$ is possible).*
- *For all $a \geq 0$, $Y_a := \sum_{b \geq a} X_b \mathsf{D}_{b,a}$ is finite. For short, we will write $Y = X.\mathsf{D}$.*
- *For all $a \geq 1$,*

$$\sum_{i=1}^{a} (Y_i / Y_0 - X_i) > 0. \tag{3.16}$$

Theorem 3.16. *Let $v_0 \in [0, 1)$, D be a descent kernel, and X be a (v_0, D) pushable sequence. Set $Y = X.\mathsf{D}$, and*

$$v_{a+1} = \frac{X_{a+1}}{Y_{a+1}/Y_0 + X_a(1/v_a - 1)} \text{ for all } a \geq 0 \tag{3.17}$$

(if $v_0 = 0$, take $v_1 = \frac{X_1}{Y_0}(1 - \frac{Y_1}{Y_0+Y_1}) = \frac{X_1}{Y_0+Y_1}$ instead). We have

$$v_a \in (0, 1), \text{ for all } a > 0, \tag{3.18}$$

since this is equivalent to (3.16). Define the 4-tuple $[u, \alpha, \beta, \pi]$ by $\pi_0 = Y_0$, $u_0 = 1$ and for $a \geq 0$,

$$\begin{cases} u_{a+1} & := & \dfrac{Y_{a+1} u_a}{Y_a(1 - v_a) + Y_{a+1} u_a}, \\ \pi_{a+1} & := & \pi_a \dfrac{1 - v_a}{1 - u_{a+1}}, \\ \beta_a & := & X_a, \\ \alpha_a & := & 1/(u_a \pi_a) \end{cases} \tag{3.19}$$

then for this 4-tuple, the ◺ L with representation (v, D) has time-reversal U as defined in Proposition 3.14, and then, both transition matrices L and U have π as invariant measure.

Proof. Let us first say why (3.18) is equivalent to (3.16). Set $\gamma_a = X_a(1/v_a - 1)$. From (3.17), one gets that

$$\gamma_{a+1} = \gamma_a + Y_{a+1}/Y_0 - X_{a+1}, \quad a \geq 0 \tag{3.20}$$

and then $\gamma_{a+1} = \gamma_0 + \sum_{i=1}^{a+1} (Y_i/Y_0 - X_i)$. Since $v_a \in (0, 1) \Leftrightarrow \gamma_a > 0$, we get the result (note that $\gamma_1 = Y_0 + Y_1 - X_1$ when $v_0 = 0$, so that (3.20) holds for $\gamma_0 = (Y_0 + Y_1 - X_1) - Y_1/Y_0 + X_1 = Y_0 + Y_1 - Y_1/Y_0$ in this case).

It suffices to check that Proposition 3.14 applies to the tuple of sequences $[\pi, u, \alpha, \beta]$ as defined in (3.19). The condition (i) is immediate since we took $\beta_0 = X_0 = v_0$; the fact that $u_j \in (0, 1)$ for all $j > 0$ is clear.

For (ii) observe that the first equation of the system (3.19) is equivalent to

$$\frac{1-v_a}{1-u_{a+1}} = \frac{Y_{a+1}/u_{a+1}}{Y_a/u_a} \tag{3.21}$$

so that (Y_a/u_a) is proportional to $(\pi_a, a \geq 0)$ as defined in the third equation of the system (3.19), and since $Y_0/u_0 = \pi_0$ these sequences are equal. Write

$$\sum_{b:b\geq a} \alpha_a \beta_b \mathsf{D}_{b,a} = \frac{1}{u_a \pi_a} \sum_{b:b\geq a} X_b \mathsf{D}_{b,a} = \frac{Y_a}{u_a \pi_a} = 1.$$

Now we prove condition (iii). Since $u_a\alpha_a = 1/\pi_a$, we only need to prove that $\beta_b = X_b = \pi_b v_b$. Since $\pi_0 = 1$, $X_0 = v_0$, the formula is true for $b = 0$; let us assume that it holds for $b \leq a$, for some a, and let us establish that $X_{a+1} = \pi_{a+1}v_{a+1}$. From (3.17),

$$\frac{X_{a+1}}{v_{a+1}} \quad = \quad \frac{u_{a+1}Y_{a+1}}{u_{a+1}Y_0} + X_a\left(\frac{1}{v_a} - 1\right) = u_{a+1}\pi_{a+1} + \pi_a(1-v_a) = \pi_{a+1}$$

by system (3.19), second equation. The case where $a = 0$ and $v_0 = 0$ has to be treated separately: in this case $u_1 = Y_1/(Y_0 + Y_1)$ and since $Y_0 = \pi_0$, $v_1 = \frac{X_1}{Y_0}(1 - Y_1/(Y_0 + Y_1))$, we have $X_1/v_1 = Y_0/(1-u_1) = \pi_0(1-v_0)/(1-u_1) = \pi_1$ so that X_1/v_1 is indeed equal to π_1.

Finally (iv) is immediate by the second equation of the system (3.19). $\square$

3.3.3 "Catastrophe transition matrices": analysis of ◺ transition matrices with same descent kernel

Theorem 3.16 gives a reformulation for the problem of finding the invariant distributions of the time-reversal of a given transition matrix L, and as such it may appear a bit useless since no methods are provided to compute the pair (X, Y) which is needed to conclude. The following examples show the power of this theorem: given the descent kernel D and some parameter $v_0 \in [0, 1)$, it is quite easy to find many (v_0, D) pushable sequences X. Even if it is still difficult to target a given sequence $v = (v_i, i \geq 0)$, it is possible to construct many sequences v for which it is possible to construct the time-reversal of (v, D). This allows observing the general form of integrable systems (v, D). The following results are comparable with those of Pollett & al. [20], Brockwell & al. [4], Kapodistria & al. [11] (and references therein), in which various catastrophe transition matrices are investigated (in continuous-time). In these results as well, it can be observed that very specific forms of catastrophe transition matrices are needed to find the invariant distributions or absorption probabilities: each time it is a challenge to complete all computation details.

Geometric catastrophe

This is the family of ◺ transition matrices whose (v_0, D) representation involved, for some $p \in (0, 1)$, the descent kernel

$$\mathsf{D}_{b,a} = p(1-p)^{b-a} + \mathbf{1}_{a=0}(1-p)^{b+1}, \;\; 0 \leq a \leq b.$$

For any positive sequence $X = (X_j, j \geq 0)$ and $a \geq 0$,

$$Y_a = \sum_{b\geq a} X_b \mathsf{D}_{b,a} = p \sum_{x\geq 0} X_{a+x}(1-p)^x + \mathbf{1}_{a=0} \sum_{b\geq 0} X_b(1-p)^{b+1},$$

so that this closed-form formula can be effectively computed for many sequences X (the $X_i's$ are the coefficients of a power series). It remains to extract the pushable sequences (those that satisfy $\sum_{i=1}^{a}(Y_i/Y_0 - X_i) > 0$, for non-negative parameters and sequences (v_0, X) with $X_0 = v_0$, see Definition 3.15). From this, the complete description of the vectors v and u can be obtained as explained in Theorem 3.16.

Binomial catastrophe

The descent kernel D, in this case, is defined as follows

$$\mathsf{D}_{b,a} = \binom{b}{a} p^a (1-p)^{b-a}, \;\; 0 \le a \le b.$$

For $X = (e^{-\lambda}\lambda^b/b!, b \ge 0)$ the Poisson distribution $(P_b^{(\lambda)}, b \ge 0)$ with parameter λ, the corresponding Y is Poisson distributed with parameter λp, i.e $Y_a = P_a^{(\lambda p)}$. In this case $Y_i/Y_0 = (p\lambda)^i/i!$ and then $\sum_{i=1}^{a}(Y_i/Y_0 - X_i) = \sum_{i=1}^{a}((p\lambda)^i/i! - e^{-\lambda}\lambda^i/i!) = \sum_{i=1}^{a}\frac{\lambda^i e^{-\lambda}}{i!}(p^i e^{\lambda} - 1)$ is indeed positive for p such that $\lambda p \ge 1 - e^{\lambda}$ at least, since $\sum_{i=1}^{a}\frac{\lambda^i e^{-\lambda}}{i!}(p^i e^{\lambda}) \ge \lambda p$ (the value taken for $a = 1$) and for each $a \ge 1$, $\sum_{i=1}^{a}\frac{\lambda^i e^{-\lambda}}{i!}(-1) \ge \sum_{i=1}^{+\infty}\frac{\lambda^i e^{-\lambda}}{i!}(-1) = -(1 - e^{-\lambda})$, so that (v_0, X) is pushable when $\lambda p \ge 1 - e^{-\lambda}$.

For $v_0 = X_0$, with $v_0 \in (0,1)$, then

$$v_{a+1} = \frac{e^{-\lambda}\lambda^{a+1}/(a+1)!}{e^{-\lambda p}(p\lambda)^{a+1}/(a+1)!/e^{-\lambda p} + e^{-\lambda}\lambda^a/a!(1/v_a - 1)} = \frac{1}{e^{\lambda}p^{a+1} + (1/v_a - 1)(a+1)/\lambda}.$$

From this formula the sequence $(v_a, a \ge 0)$ is characterized. From here, set $u_0 = 1$, and compute successively, for $a \ge 0$,

$$u_{a+1} = \frac{e^{-\lambda p}(\lambda p)^{a+1}/(a+1)!u_a}{e^{-\lambda p}(p\lambda)^a/a!(1-v_a) + e^{-\lambda p}(\lambda p)^{a+1}/(a+1)!u_a} = \frac{(\lambda p)u_a}{(1-v_a)(a+1) + (\lambda p)u_a}.$$

After that, the values of π can be obtained from the third formula in the system (3.19).

Uniform catastrophe

The case of uniform catastrophe is described by the following descent kernel $\mathsf{D}_{b,a} = 1/(b+1)$ for $a \in [0,b]$. The case where $X_b = (b+1)\rho_b$ for ρ_b a probability measure with full support on $\mathbb{N}$ and a finite mean, is integrable. Denote $\overline{\rho}_b = \sum_{k \ge b}\rho_k$ (the tail distribution function). The computation of Y gives $Y_a = \overline{\rho}_a$ so that $Y_0 = 1$. The pushability condition is $\sum_{i=1}^{a}(\overline{\rho}_i - (i+1)\rho_i) \ge 0$. For $v_0 = X_0$, with $v_0 \in (0,1)$, compute successively the v_a using:

$$v_{a+1} = \frac{(a+2)\rho_{a+1}}{\overline{\rho}_{a+1} + (a+1)\rho_a}.$$

Set $u_0 = 1$, and compute the u_a using:

$$u_{a+1} = \frac{\overline{\rho}_{a+1}u_a}{\overline{\rho}_a(1-v_a) + \overline{\rho}_{a+1}u_a}.$$

After that, the values of π can be obtained from the third formula in the system (3.19).

4 Particular models

4.1 Back to the tridiagonal case

About the formulas for invariant distributions

In the tridiagonal case, by (1.5), the invariant distribution $(\pi_a, a \ge 0)$ associated with the tridiagonal transition matrix T is unique (since it is $\triangledown$), and it is proportional to

$(p_a^{(t)} := \prod_{j=1}^a \frac{\mathsf{T}_{j-1,j}}{\mathsf{T}_{j,j-1}}, a \geq 0)$, and to $(p_a^{(\mathsf{U})} := \frac{\det(\mathsf{Id}-\mathsf{T}_{[0,a-1]})}{\prod_{j=1}^a \mathsf{T}_{j,j-1}}, a \geq 0)$ in Theorem 2.1 for the ◹ case. Therefore, in the tridiagonal case, these two formulas must coincide.

The fact that $p_a^{(t)} = p_a^{(\mathsf{U})}$ is a consequence of Theorem 2.1 and Remark 2.2(ii) (and it is also a consequence of (3.3) and Proposition 3.5).

In the ◺ case, it is a bit more complex since Theorem 2.19 only deals with finite ◺ transition matrices then we need to use Proposition 2.20. We then take $\mathsf{L}^{(n)}$ as described in this Proposition 2.20. In the tridiagonal case, $\mathsf{L}^{(n)}_{i,j} = \mathsf{T}_{i,j}$ for all $i,j \leq n$, except for $\mathsf{L}^{(n)}_{n,n} = \mathsf{T}_{n,n} + \mathsf{T}_{n,n+1}$. For any $a < n$, we have, by Theorem 2.19

$$\rho_a^{(n)} = c_n \det(\mathsf{Id} - \mathsf{L}^{(n)}_{[a+1,n]}) \prod_{i=1}^{a} \mathsf{T}_{i-1,i}.$$

The determinant, by the matrix tree theorem, coincides with the weight of trees rooted at a, on the graph with vertex set $[a,n]$, and edge set $\{(i,j) \;:\; \mathsf{T}_{i,j} > 0, i,j \leq n\}$. Since the only decreasing edges are the $(j,j-1)$, there is a single tree on $[a,n]$ rooted at a, it is the tree with edges $\{(j,j-1), j \in [a+1,n]\}$. Hence

$$\rho_a^{(n)} = c_n \left(\prod_{i=a+1}^{n} \mathsf{T}_{i,i-1}\right)\left(\prod_{i=1}^{a} \mathsf{T}_{i-1,i}\right) \mathbf{1}_{a\leq n} = c'_n \prod_{i=1}^{a} \frac{\mathsf{T}_{i-1,i}}{\mathsf{T}_{i,i-1}} \mathbf{1}_{a\leq n}.$$

Set, as done in Proposition 2.20,

$$\eta_a^{(n)} = \rho_a^{(n)}/\rho_0^{(n)} = \mathbf{1}_{a\leq n} \prod_{i=1}^{a} \frac{\mathsf{T}_{i-1,i}}{\mathsf{T}_{i,i-1}}.$$

It remains to prove that the two conditions (a) and (b) of Proposition 2.20 are satisfied.

- (a) We need to take $S_a = \prod_{i=1}^a \frac{\mathsf{T}_{i-1,i}}{\mathsf{T}_{i,i-1}}$ (which is $p_a^{(t)} = p_a^{(\mathsf{U})}$ by the way). For b fixed, we have $\sum_{a:a\geq b+1} S_a \mathsf{T}_{a,b} = S_{b+1}\mathsf{T}_{b+1,b}$ and this is indeed $< +\infty$.
- (b) Let $\eta_a := S_a = p_a^{(t)} = p_a^{(\mathsf{U})}$. The convergence of $\eta_a^{(n)} \xrightarrow[n\to\infty]{} \eta_a$ is obvious.

Recurrence criterion

In the tridiagonal case, the criterion for recurrence is known to be (1.7) (that is $\sum_{k\geq 0}\prod_{j=1}^{k} \frac{\mathsf{T}_{j,j-1}}{\mathsf{T}_{j,j+1}} = +\infty$). Let us show that it is equivalent to Theorem 2.8 in this case $\left(\lim_{b\to+\infty} u_1(b) = 1 \text{ with } u_1(b) = \mathsf{M}_{1,0}\frac{\det(\mathsf{Id}-\mathsf{T}_{[2,b-1]})}{\det(\mathsf{Id}-\mathsf{T}_{[1,b-1]})}\right)$. If T is tridiagonal

$$\det(\mathsf{Id} - \mathsf{T}_{[1,b-1]}) = (1 - \mathsf{T}_{1,1})\det(\mathsf{Id} - \mathsf{T}_{[2,b-1]}) - \mathsf{T}_{1,2}\mathsf{T}_{2,1}\det(\mathsf{Id} - \mathsf{T}_{[3,b-1]})$$

and writing $D_{a,b} := \det(\mathsf{Id} - \mathsf{T}_{[a,b-1]})$, we have more generally

$$D_{i,b-1} = (1 - \mathsf{T}_{i,i})D_{i+1,b-1} - \mathsf{T}_{i,i+1}\mathsf{T}_{i+1,i}D_{i+2,b-1}, \quad \text{for } i+1 \leq b-1. \tag{4.1}$$

Set

$$Z_{i,b-1} = \frac{D_{i,b-1}}{D_{i+1,b-1}\,\mathsf{T}_{i,i-1}}$$

so that $u_1(b) = 1/Z_{1,b-1}$. Formula (4.1) rewrites

$$Z_{i,b-1} \quad = c_i + a_{i+1} \,/\, Z_{i+1,b-1} \text{ for } i \leq b-2 \tag{4.2}$$

for

$$q_i := \frac{\mathsf{T}_{i,i+1}}{\mathsf{T}_{i,i-1}}, \quad a_{i+1} = -q_i, \quad c_i = 1 + q_i. \tag{4.3}$$

Notice that "the last" term, for $i = b-1$,

$$Z_{b-1,b-1} = \frac{D_{b-1,b-1}}{D_{b,b-1}\,\mathsf{T}_{b-1,b-2}} = \frac{1-\mathsf{T}_{b-1,b-1}}{\mathsf{T}_{b-1,b-2}} = \frac{\mathsf{T}_{b-1,b-2}+\mathsf{T}_{b-1,b}}{\mathsf{T}_{b-1,b-2}} = 1 + q_{b-1} = c_{b-1}$$

so that (4.2) can be used to express $Z_{1,b-1}$ in terms of the a_i's and c_i's as follows: if one sees the relation (4.2) as a kind of continued fraction expansion, what we want to do is produce a formula for the so-called convergents:

$$Z_{1,b-1} := c_1 + \cfrac{a_2}{c_2 + \cfrac{a_3}{\ddots + \cfrac{\ddots}{c_{b-2} + \cfrac{a_{b-1}}{c_{b-1}}}}}; \tag{4.4}$$

One can compute the value of the "finite continuous fraction" expressed in (4.4): set

$$A_0 = 1, A_1 = c_1 = 1 + q_1, B_0 = 0, B_1 = 1$$

and computing successively

$$\begin{cases} A_k &= c_k A_{k-1} + a_k A_{k-2} = (1+q_k)A_{k-1} - q_{k-1}A_{k-2}, \\ B_k &= c_k B_{k-1} + a_k B_{k-2} = (1+q_k)B_{k-1} - q_{k-1}B_{k-2} \end{cases} \tag{4.5}$$

for k from 2 to $b-1$, we get

$$Z_{1,b-1} = A_{b-1} \,/\, B_{b-1}.$$

We can proceed to the computation of A_{b-1} and B_{b-1}, by observing that for any $k \le b-1$, since $A_k = (1+q_k)A_{k-1} - q_{k-1}A_{k-2}$ and $B_k = (1+q_k)B_{k-1} - q_{k-1}B_{k-2}$, we then have for $k \le b-1$:

$$\alpha_k := A_k - q_k A_{k-1} = A_{k-1} - q_{k-1}A_{k-2} = \alpha_{k-1},$$
$$\beta_k := B_k - q_k B_{k-1} = B_{k-1} - q_{k-1}B_{k-2} = \beta_{k-1}.$$

From what we see that $(A_k - q_k A_{k-1}, k < h)$ and $(B_k - q_k B_{k-1}, k < h)$ are both constant, but are subject to different initial conditions. For $C_k = A_k$ or $C_k = B_k$, with c encoding the initial condition,

$$C_k \quad = \quad q_k C_{k-1} + c \Rightarrow C_k = q_k(q_{k-1}C_{k-2} + c) + c = C_0 \prod_{j=1}^{k} q_j + c \sum_{j=1}^{k} \prod_{i=j+1}^{k} q_i. \tag{4.6}$$

One gets for $k \le b-1$, setting $Q_k := \prod_{j=1}^{k} q_j = Q_{k-1}q_k$, $F_k := \sum_{j=1}^{k} \left(\prod_{i=1}^{j} q_i\right)^{-1}$

$$A_k = Q_k(1+F_k), \quad B_k = Q_k F_k.$$

Hence,

$$Z_{1,b-1} = 1 + 1/F_{b-1}$$

and since $F_k := \sum_{j=1}^{k} \left(\prod_{i=1}^{j} q_i\right)^{-1} = \sum_{j=1}^{k} \prod_{i=1}^{j} \frac{\mathsf{T}_{i,i-1}}{\mathsf{T}_{i,i+1}}$ and then one observes that the convergence of $Z_{1,b-1}$ to 1 is indeed equivalent to Karlin & McGregor criterion (1.7).

Positive recurrence criterion

The positive recurrent criterion is $\sum_{k\ge 0} \prod_{j=1}^{k} \frac{\mathsf{T}_{j-1,j}}{\mathsf{T}_{j,j-1}} < \infty$, while it is $\sum_{a=1}^{\infty} \frac{\det\left(\mathsf{Id}-\mathsf{U}_{[0,a-1]}\right)}{\prod_{j=1}^{a} \mathsf{U}_{j,j-1}} < \infty$ in the U case. Section 4.1 explained why, in the tridiagonal case, the formulae of the invariant measures obtained in the $\triangledown$ and in the tridiagonal cases coincide: since the formulae coincide, the criteria for positive recurrence coincide.

4.2 A class of integrable almost lower-triangular transition matrices

In Section 3.3.2, we described a strategy to design some ◺ transition matrices L for which the time-reversal U can be computed (for some invariant measure π computed simultaneously). Here, we present a large family of transition matrices for which the invariant distribution can be computed directly.

Denote by $\mathsf{e}_j = (\mathbf{1}_{i=j}, i \geq 0)$ the vector with a 1 in entry j only (the first vector is e_0).

Definition 4.1. *A ◺-transition matrix is said to be Col(0)-triangular if for any $c \geq 1$,*

$$\mathsf{L}_{\bullet,c} = \alpha_c \mathsf{L}_{\bullet,0} + \sum_{\ell=0}^{c} a_{\ell,c}\, \mathsf{e}_\ell, \tag{4.7}$$

where, $\mathsf{L}_{\bullet,c}$ denotes the column c of L, the first column being $\mathsf{L}_{\bullet,0}$.

The first column $\mathsf{L}_{\bullet,0}$ is general, and for any c, $\mathsf{L}_{\bullet,c}$ is essentially proportional to $\mathsf{L}_{\bullet,0}$, except its $c+1$ first entries, indexed from 0 to c. Since $\mathsf{L}_{0,c} = \cdots = \mathsf{L}_{c-2,c} = 0$, for $c \geq 1$, only the entries $\mathsf{L}_{c,c}$ and $\mathsf{L}_{c-1,c}$ are free.

We claim that it is possible to solve the system $\eta = \eta \mathsf{L}$ when L is a slt, irreducible and Col(0)-triangular. The idea is to keep on hold the first equation until the end of the resolution:

$$\eta_0 = \eta L_{\bullet,0} = \sum_k \eta_k L_{k,0}. \tag{4.8}$$

Suppose that η is the solution of (4.8) and $\eta = \eta \mathsf{L}$. For $c \geq 1$, we have

$$\eta_c = \eta \mathsf{L}_{\bullet,c} = \alpha_c\, \eta\, \mathsf{L}_{\bullet,0} + \eta \sum_{\ell=0}^{c} a_{\ell,c}\, \mathsf{e}_\ell = \alpha_c\, \eta_0 + \sum_{\ell=0}^{c} a_{\ell,c}\, \eta_\ell \tag{4.9}$$

and it is then apparent that $(\eta_c, c \geq 1)$ is a solution of a standard triangular linear system in which η_0 is seen as a parameter. It remains to check if the obtained solution of (4.9) solves (4.8) or not (which corresponds to the case where a solution exists or none, respectively).

Example: Consider a sequence of vectors $V^{\geq k} = \begin{bmatrix} V_0 1_{0 \geq k} \\ V_1 1_{1 \geq k} \\ \vdots \end{bmatrix}$ (indexed by k) where, vertically, the entries are non increasing: $1 > V_0 > V_1 > V_2 > \cdots > 0$. The vectors $V^{\geq k}$ are essentially proportional, and $V := V^{\geq 0}$ has the sequence $(V_i, i \geq 0)$ as entries. The simplest cases of Col(0)-triangular transition matrices are those in which the columns $\mathsf{L}_{\bullet,c}$ are essentially proportional. Consider $\alpha_0 = 1, \alpha_1, \alpha_2, \ldots$ and the matrix L such that

$$\mathsf{L}_{\bullet,c} = \alpha_c V^{\geq c-1} = \alpha_c \Big[L_{\bullet,0} - \sum_{j=0}^{c-2} V_j \mathsf{e}_j \Big] \text{ for all } c \geq 0. \tag{4.10}$$

Since L is a ◺-transition matrix, the condition $\sum_c \mathsf{L}_{r,c} = 1$ for all r becomes:

$$(\alpha_0 + \cdots + \alpha_{r+1})\, V_r = 1 \quad \text{for } r \geq 0, \tag{4.11}$$

which implies $\sum_{\ell=0}^{r+1} \alpha_\ell = 1/V_r$, and since $1 + \alpha_1 = \alpha_0 + \alpha_1 = \frac{1}{V_0}$, we obtain

$$\alpha_1 = \frac{1}{V_0} - 1, \text{ for } r > 0, \alpha_{r+1} = \frac{1}{V_r} - \frac{1}{V_{r-1}}. \tag{4.12}$$

Since $(V_i, i \geq 0)$ is decreasing, it is easily seen that $r \mapsto (\sum_{i=1}^{r} \alpha_i)$ is increasing so that all the α_i are positive. We have

$$\eta_k = \eta \mathsf{L}_{\bullet,k} = \eta_0 \alpha_k (1 - V_0 - \cdots - V_{k-2}), \quad k \geq 0 \tag{4.13}$$

we then need that $\sum_{i=0}^{+\infty} V_i \leq 1$ and

$$\eta_0 = \sum_k \eta_k L_{k,0} = \sum_k \eta_k V_k < +\infty,$$

which is equivalent to

$$1 = \sum_k \alpha_k V_k (1 - V_0 - \cdots - V_{k-2}). \tag{4.14}$$

Using (4.12), it can be written as

$$V_0 + V_1 \left(\frac{1}{V_0} - 1 \right) + \sum_{k \geq 2} \left(\frac{1}{V_{k-1}} - \frac{1}{V_{k-2}} \right) V_k (1 - V_0 - \cdots - V_{k-2}) = 1. \tag{4.15}$$

If L satisfies (4.10) and (4.15), then its invariant measures η can be computed thanks to (4.13).

4.3 Some ◺ and ◹ transition matrices associated with BDP Markov chains

Let T be an irreducible tridiagonal transition matrix. As explained at the beginning of Section 1, the invariant measure of T and the criterion of recurrence and positive recurrence are available. Now define two associated transition matrices U and L as follows. Let X be a Markov process with transition matrix T. The increasing steps of X are $+1$ while decreasing steps are -1. Now, define

$$\begin{cases} \mathsf{U}_{a,b} &= \mathbb{P}\left(X_{\tau^{\downarrow}} = b \mid X_0 = a\right), \\ \mathsf{L}_{a,b} &= \mathbb{P}\left(X_{\tau^{\uparrow}} = b \mid X_0 = a\right) \end{cases} \tag{4.16}$$

where

$$\tau^{\downarrow} = \inf\{t > 0 \ : \ X_t = X_{t-1} - 1\},$$
$$\tau^{\uparrow} = \inf\{t > 0 \ : \ X_t = X_{t-1} + 1\}.$$

In words, if one observes X only at the times $(t_j, j \in \mathbb{Z})$ following a decreasing step, then the sequence of observations is an U-Markov process on $\{0, 1, 2, 3, \ldots\}$. If one observes X only at the times $(t_j, j \in \mathbb{Z})$ following an increasing step, then the sequence of observations is a L-Markov process on $\{1, 2, 3, \ldots\}$.

The transition matrices U and L can be computed using path decompositions: Set $\mathsf{Loop}_i^{\mathsf{T}} := (1 - \mathsf{T}_{i,i})^{-1}$. For all $i \geq 0$, we have $\mathsf{U}_{i,i-1} = \mathsf{Loop}_i^{\mathsf{T}}\, \mathsf{T}_{i,i-1}$, and for $j \geq i$,

$$\mathsf{U}_{i,j} = \left[\prod_{b=i}^{j} \mathsf{Loop}_b^{\mathsf{T}}\, \mathsf{T}_{b,b+1} \right] \mathsf{Loop}_{j+1}^{\mathsf{T}}\, \mathsf{T}_{j+1,j}.$$

We have, for $i \geq 1$, $\mathsf{L}_{i,i+1} = \mathsf{Loop}_i^{\mathsf{T}}\, \mathsf{T}_{i,i+1}$, for $1 \leq j \leq i$,

$$\mathsf{L}_{i,j} = \left[\prod_{b=j}^{i} \mathsf{Loop}_b^{\mathsf{T}} \mathsf{T}_{b,b-1} \right] \mathsf{Loop}_{j-1}^{\mathsf{T}}\, \mathsf{T}_{j-1,j}.$$

Remark 4.2. *The matrices* U *(respectively* L*) are transition matrices when* T *is recurrent, since, in this case, with probability 1, starting from any position* i*, a Markov chain* X *with transition matrix* T *will have some decreasing steps (resp. increasing steps) which ensures that, for any* i*,* $\sum_j \mathsf{U}_{i,j} = 1$ *(resp.* $\sum_j \mathsf{L}_{i,j} = 1$*). It is however possible for a BDP to have globally a.s. a finite number of steps* -1 *in the transient case, so that* U *is not always a transition matrix. Since starting from any point* i*, a Markov chain with transition matrix* T *will have a* $+1$ *step with probability 1 (by irreducibility),* L *is well defined-even when* T *is transient.*

Set

$$\begin{cases}\pi^{\mathsf{U}}_a &= \pi^{\mathsf{T}}_{a+1}\,\mathsf{T}_{a+1,a} \quad \text{for } a\geq 0,\\ \pi^{\mathsf{L}}_a &= \pi^{\mathsf{T}}_{a-1}\,\mathsf{T}_{a-1,a} \quad \text{for } a\geq 1.\end{cases} \tag{4.17}$$

Proposition 4.3. *Assume that* T *is irreducible.*

(i) T *is recurrent (resp. positive recurrent) iff* L *is irreducible and recurrent (resp. positive recurrent). If* L *is a transition matrix then,* L *admits* π^{L} *as invariant measure (in all cases, including* L *transient).*

(ii) If T *is recurrent and* U *is a transition matrix, then* U *is irreducible and recurrent.* T *is positive recurrent iff* U *is positive recurrent. The measure* π^{U} *is invariant by* U *(in all cases, including* U *transient, and even, when* U *is not a transition matrix!).*

Proof. In the proof, we treat simultaneously (i) and (ii). First, the fact that the recurrence of T is equivalent to irreducibility and recurrence of U (resp. of L) is clear. By irreducibility of T, the recurrence of T implies that each edge $(a,a+1)$ and $(a,a-1)$ are traversed infinitely often by a Markov chain with transition matrix T so that U and L are recurrent. The converse uses the same type of argument.

If T is positive recurrent, then by the ergodic theorem, the proportion of time spent at a by a Markov chain with transition matrix T converges to π^{T}_a, and then, the proportion of time spent at an increasing step $(a,a+1)$ is $\pi^{\mathsf{T}}_a\mathsf{T}_{a,a+1}$, and the proportion of time spent at a decreasing step $(a,a-1)$ is $\pi^{\mathsf{T}}_a\mathsf{T}_{a,a-1}$. A simple consequence of that is that (4.17) holds in the positive recurrent case (since $\sum_a \pi^{\mathsf{T}}_{a+1}$ converges, $\sum_a \pi^{\mathsf{T}}_{a+1}\mathsf{T}_{a+1,a}$ and $\sum_a \pi^{\mathsf{T}}_{a+1}\mathsf{T}_{a,a+1}$ converge too).

Now, let us check the statements concerning the invariant measures.

By (1.5),

$$\pi^{\mathsf{U}}_a = \pi^{\mathsf{T}}_{a+1}\mathsf{T}_{a+1,a} = \frac{\prod_{j=1}^{a+1}\mathsf{T}_{j-1,j}}{\prod_{j=1}^{a}\mathsf{T}_{j,j-1}}.$$

We want to prove that π^{U} is invariant by U.

$$\begin{aligned}\pi^{\mathsf{U}}_a\mathsf{U}_{a,b} &= \frac{\mathsf{T}_{0,1}\dots\mathsf{T}_{a,a+1}}{\mathsf{T}_{1,0}\dots\mathsf{T}_{a,a-1}}\mathsf{T}_{b+1,b}\left(\prod_{i=a}^{b}\mathsf{T}_{i,i+1}\right)\left(\prod_{i=a}^{b+1}\mathsf{Loop}^{\mathsf{T}}_i\right)\\ &= \frac{\mathsf{T}_{0,1}\dots\mathsf{T}_{b,b+1}}{\mathsf{T}_{1,0}\dots\mathsf{T}_{b,b-1}}\left[\mathsf{T}_{a,a+1}\mathsf{Loop}^{\mathsf{T}}_a\left(\prod_{i=a+1}^{b+1}\mathsf{T}_{i,i-1}\mathsf{Loop}^{\mathsf{T}}_i\right)\right] = \pi^{\mathsf{U}}_b\mathsf{L}_{b+1,a+1}\end{aligned}$$

since the last bracket is $\mathsf{L}_{b+1,a+1}$; this allows us to conclude the invariance of π^{U} by U:

$$\sum_{a:a\leq b+1}\pi^{\mathsf{U}}_a\mathsf{U}_{a,b} = \pi^{\mathsf{U}}_b\sum_{a:a\leq b+1}\mathsf{L}_{b+1,a+1} = \pi^{\mathsf{U}}_b\sum_{j:j\leq b+2}\mathsf{L}_{b+1,j} = \pi^{\mathsf{U}}_b.$$

Now, we want to prove that π^{L} is invariant by L. Write $\pi^{\mathsf{L}}_a = \pi^{\mathsf{T}}_{a-1}\mathsf{T}_{a-1,a} = \frac{\prod_{j=1}^{a-1}\mathsf{T}_{j-1,j}}{\prod_{j=1}^{a-1}\mathsf{T}_{j,j-1}}\mathsf{T}_{a-1,a}$ $= \frac{\prod_{j=1}^{a}\mathsf{T}_{j-1,j}}{\prod_{j=1}^{a-1}\mathsf{T}_{j,j-1}}$, so that

$$\begin{aligned}
\pi^{\mathsf{L}}_a\mathsf{L}_{a,b} &= \frac{\prod_{j=1}^{a}\mathsf{T}_{j-1,j}}{\prod_{j=1}^{a-1}\mathsf{T}_{j,j-1}}\left[\prod_{k=b}^{a}\mathsf{Loop}^{\mathsf{T}}_k\mathsf{T}_{k,k-1}\right]\mathsf{Loop}^{\mathsf{T}}_{b-1}\,\mathsf{T}_{b-1,b} \\
&= \left[\frac{\prod_{j=1}^{b-1}\mathsf{T}_{j-1,j}}{\prod_{j=1}^{b-1}\mathsf{T}_{j,j-1}}\mathsf{T}_{b-1,b}\right]\mathsf{M}_{a,a-1}\left[\prod_{k=b}^{a}\mathsf{T}_{k-1,k}\mathsf{Loop}^{\mathsf{T}}_k\right]\mathsf{Loop}^{\mathsf{T}}_{b-1} \\
&= \pi^{\mathsf{L}}_{b-1}\left[\prod_{k=b-1}^{a-1}\mathsf{Loop}^{\mathsf{T}}_k\mathsf{T}_{k,k+1}\right]\left(\mathsf{T}_{a,a-1}\mathsf{Loop}^{\mathsf{T}}_a\right) = \pi^{\mathsf{L}}_{b-1}\mathsf{U}_{b-1,a-1}.
\end{aligned}$$

From here, $\sum_a \pi^{\mathsf{L}}_a\mathsf{L}_{a,b} = \sum_a \pi^{\mathsf{L}}_{b-1}\mathsf{U}_{b-1}, a = \pi^{\mathsf{L}}_{b-1}$ if U is a transition matrix. □

Remark 4.4. *This example is one of the simplest* ◹ *and* ◺ *transition matrices one can construct using stopping times of a BD Markov chain. One can construct many other transition matrices by designing other stopping times: for example, for a* ◹ *transition matrix: starting at* k*, start the trajectory when it hits* $k-1$*, or at the first time where the 5 last steps are* $(+1,+0,+1,-1,+1)$*.*

4.4 Repair shop Markov chain

The following is a well-known Markov chain that is present in different textbooks. It is an integrable system, where recurrence, positive recurrence and transience have been characterized. Most of the results about this chain can be found in Brémaud's Book [3] under the tag *repair shop*. Nevertheless, the methods that we will use to obtain the same results are based on Theorem 2.1 and are therefore of different nature.

The repair shop chain is defined as the Markov chain X_n given by:

$$X_{n+1} = (X_n - 1)_+ + Z_{n+1}$$

where $(Z_n, n \geq 0)$ is a sequence of i.i.d. random variables with distribution $(a_k, k \geq 0)$, meaning that $\mathbb{P}(Z_n = k) = a_k$ for every $k \geq 0$. This chain models the number of broken machines in a repair shop, where each day one broken machine is repaired (when there is at least one available to repair), and where the number of new machines that need to be repaired day $n+1$ is Z_{n+1}. The transition matrix A associated with this chain is

$$\mathsf{A} = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & a_0 & a_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

Note 4.5. *Some generalizations of the repair shop Markov chain appear in the literature, notably, in relation to queueing theory; see e.g. Abolnikov & Dukhovny [1] and references therein.*

4.4.1 Positive recurrence criterion

Set $m := \sum_s s a_s$ the mean of Z_1, i.e. of the distribution $(a_0, a_1, \ldots)$.

Proposition 4.6. *The transition matrix* A *is positive recurrent if and only if* $m < 1$*.*

Proof. By Lemma 1.1, equation (1.12)

$$C_N := \frac{\det(\mathsf{Id} - \mathsf{A}_{[0,N]})}{\prod_{j=1}^{N+1} \mathsf{A}_{j,j-1}} = \sum_{s\in S^N} \left(\prod_{j=1}^{\ell(s)} (\mathsf{Id} - \mathsf{A})_{s_{j-1}+1,s_j}\right) \left(\prod_{\substack{j\in[0,N-1]\\ j\neq s}} \mathsf{A}_{j+1,j}\right) a_0^{-N-1} \quad (4.18)$$

$$= \sum_{s\in S^N} \prod_{j=1}^{\ell(s)} \frac{(\mathsf{Id} - \mathsf{A})_{s_{j-1}+1,s_j}}{a_0}. \quad (4.19)$$

Since $(\mathsf{Id} - \mathsf{A})_{x,y} = 1_{x=y} - a_{y-x+1-1_{x=0}}$, we have

$$C_N = \sum_{s\in S^N} \prod_{j=1}^{\ell(s)} \frac{1_{s_j - s_{j-1}=1} - a_{s_j-(s_{j-1}+1)+1-1_{s_{j-1}=0}}}{a_0}.$$

This is a kind of product of transitions that measures the passage from s_{j-1} to s_j. For the increments except for the first one, define

$$t_\delta = \frac{1_{\delta=1} - a_\delta}{a_0}, \quad \delta \geq 1$$

and for the first increment, define

$$t_\delta^f = \frac{(\mathsf{Id} - \mathsf{A})_{0,\delta}}{a_0} = \frac{1_{\delta=0} - a_\delta}{a_0}, \delta \geq 0.$$

Consider the generating functions which encode the increments of the sequence $(s_j, j \geq 0)$

$$G_f(x) = \sum_{\delta\geq 0} t_\delta^f x^\delta = \frac{1-a_0}{a_0} - \sum_{s>0} \frac{a_s}{a_0} x^s,$$

$$G(x) = \sum_{\delta\geq 1} t_\delta x^\delta = \frac{1-a_1}{a_0} x - \sum_{s>1} \frac{a_s}{a_0} x^s.$$

Now, $C_N = [x^N] G_f(x)/(1-G(x))$ (notation for the extraction of the coefficient of x^N in the generating function $G_f(x)/(1-G(x))$) so that

$$\sum_{N\geq 0} C_N = \lim_{x\to 1} G_f(x)/(1-G(x));$$

the sum of the coefficients of a series is obtained by a simple evaluation at 1 but only when the power series converge at this point. Since $G^f(1) = 0$ and $1 - G(1) = 0$, we apply the L'hôpital rule, which says

$$\sum_{N\geq 0} C_N = \lim_{x\to 1} \frac{G_f(x)}{1-G(x)} = \lim_{x\to 1} \frac{G'_f(x)}{-G'(x)}.$$

We have,

$$G'_f(1) = -\sum_{s>0} \frac{a_s s}{a_0} = -\frac{m}{a_0},$$

$$-G'(1) = -\left[(1-a_1) - \sum_{s>1} s a_s\right] / a_0 = -\frac{1-m}{a_0};$$

therefore $\sum_{N\geq 0} C_N = m/(1-m)$, from what we see that this converges iff $m < 1$, which is then the sufficient and necessary condition for positive recurrence. ☐

4.4.2 Recurrence criterion

Proposition 4.7. *The transition matrix* A *is recurrent if and only if* $m \leq 1$.

Proof. Let $\alpha_N := \det((\mathsf{Id} - \mathsf{J})_{[0,N]})$ for

$$\mathsf{J} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & \cdots \\ a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & a_0 & a_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix},$$

where A is the matrix presented in the previous section. Since $(\mathsf{Id} - \mathsf{A})_{[2,N]} = (\mathsf{Id} - \mathsf{A})_{[1,N-1]} = (\mathsf{Id} - \mathsf{J})_{[0,N-2]}$, by Theorem 2.8

$$\mathsf{A} \text{ is recurrent } \Leftrightarrow \left(\lim_{N\to\infty} a_0 \frac{\alpha_{N-1}}{\alpha_N} = 1 \right). \tag{4.20}$$

By Lemma 1.1

$$\alpha_N = \sum_{s \in S^N} \left(\prod_{j=1}^{\ell(s)} (\mathsf{Id} - \mathsf{J})_{s_{j-1}+1, s_j} \right) \left(\prod_{j \in [0,N-1] \setminus s} \mathsf{J}_{j+1,j} \right) \tag{4.21}$$

$$= \sum_{s \in S^N} \prod_{j=1}^{\ell(s)} (\mathsf{Id} - \mathsf{J})_{s_{j-1}+1, s_j} a_0^{s_j - s_{j-1} - 1}. \tag{4.22}$$

Since $(\mathsf{Id} - \mathsf{J})_{x,y} = 1_{x,y} - a_{y-x+1}$, we have

$$\alpha_N = \sum_{s \in S^N} \prod_{j=1}^{\ell(s)} \left(1_{s_j - s_{j-1} = 1} - a_{s_j - (s_{j-1}+1)+1} \right) a_0^{s_j - s_{j-1} - 1}.$$

Again, this is a kind of product of transitions that weigh the passage from s_{j-1} to s_j. We set

$$t_\delta = (1_{\delta=1} - a_\delta)\, a_0^{\delta-1}, \quad \delta \geq 1.$$

Consider the generating function which encodes the increments of the sequence $(s_j, j \geq 0)$

$$G_{\mathsf{J}}(x) = \sum_{\delta \geq 1} t_\delta x^\delta = \frac{(1-a_1)}{a_0}(a_0 x) - \sum_{s>1} \frac{a_s}{a_0} (a_0 x)^s.$$

Now, $\alpha_N = [x^N]\,(1/(1-G_{\mathsf{J}}(x)))$. By Theorem 2.8, $a_0 \frac{\alpha_{N-1}}{\alpha_N}$ is non-decreasing in N and bounded by 1, so that $\lim_N a_0 \frac{\alpha_{N-1}}{\alpha_N}$ exists, and better than that it is equal to $\sup_N a_0 \frac{\alpha_{N-1}}{\alpha_N} = S \leq 1$. A standard result of calculus (the ratio test) applies: the radius r of convergence of $1/(1-G_{\mathsf{J}})$ satisfies

$$1/r = a_0/S.$$

This allows us to complete a step in our reasoning:

$$\mathsf{A} \text{ is recurrent } \Leftrightarrow \left(\lim_N a_0 \frac{\alpha_{N-1}}{\alpha_N} = 1 \right) \Leftrightarrow (r = 1/a_0).$$

Now notice that the function $1/(1-z)$ has a radius of convergence 1 and since G_{J} has radius of convergence at least $1/a_0$, the function $1/(1-G_{\mathsf{J}}(x))$ has radius of convergence given by the

$$\begin{aligned}\overline{R} &:=_{(a)} (1/a_0) \wedge \inf\{|x|, G_{\mathsf{J}}(x)=1\}\\ &=_{(b)} (1/a_0) \wedge \inf\{x>0, G_{\mathsf{J}}(x)=1\}\\ &=_{(c)} \inf\{x>0, G_{\mathsf{J}}(x)=1\}\end{aligned}$$

Equality (a) follows from the preceding discussion, equality (b) from the fact that the coefficients of G_{J} are non-positive (from a_2/a_s), and equality (c) holds because $G_{\mathsf{J}}(1/a_0)=1$. We infer that

$$\mathsf{A} \text{ is recurrent } \Leftrightarrow \overline{R} = 1/a_0. \tag{4.23}$$

To finish this sequence of equivalent assertions, it is enough to prove that

$$\overline{R} = 1/a_0 \Leftrightarrow m \le 1.$$

Start by noting that $G_{\mathsf{J}}(1/a_0) = (1/a_0)(1-\sum_{s\ge1} a_s) = 1$. Now $G'_{\mathsf{J}}(x) = 1-\sum_{s\ge1} s a_s (a_0 x)^{s-1}$ and therefore $G'_{\mathsf{J}}(1/a_0) = 1-m$.

If $m>1$, then the function G_{J} is locally decreasing at $1/a_0$. Hence, there exists some $\varepsilon>0$ such that $G_{\mathsf{J}}(1/a_0-\varepsilon)>1$, which implies together with $G_{\mathsf{J}}(0)=0$ and the intermediate value theorem, that $G_{\mathsf{J}}=1$ has at least one solution on $[0,1/a)$ which implies that $\overline{R}<1/a_0$ (it is transient).

If $m\le 1$, then $0 \le \sum_{s\ge1} s a_s (a_0x)^{s-1} \le \sum_{s\ge1} s a_s = m$ when $0\le x\le 1/a_0$ so that $G'_{\mathsf{J}}(x) \ge 1-m \ge 0$ on $[0,1/a_0]$ (in fact, $G'_{\mathsf{J}}>0$ on $[0,1/a_0)$). Since $G_{\mathsf{J}}(1/a_0)=1$ and $G_{\mathsf{J}}(0)=0$, and G_{J} is increasing monotone on $[0,1/a_0)$, then $\overline{R}=1/a_0$. $\square$

A Appendix

A.1 Continuous-time counterparts

A continuous Markov process $X=(X_t, t\ge 0)$ is a continuous-time process described by means of a generator $\mathsf{G}=(\mathsf{G}_{i,j} : i,j\in S)$, where S ($\mathbb{N}$ for us) is the state space, and which satisfies $\mathsf{G}_{i,i} = -\sum_{j\ne i}\mathsf{G}_{i,j}$ (each of these sums being finite), so that each row of G sums up to zero. The value $\mathsf{G}_{i,j}$ is a rate (for $i\ne j$), and can be seen as the parameter of an exponential distribution: it is the jump rate for the process when its value is i, at which it jumps at $j\ne i$. For more information on this type of process see [18, 19]. An invariant measure π^C for the continuous Markov process is a non-negative measure satisfying $\pi^C\mathsf{G}=0$.

The jump process associated with X is the discrete-time Markov chain $Y=(Y_k, k\ge 0)$, defined by

$$Y_k = X(\tau_k), \quad \text{for } k\ge 0$$

where $\tau_0=0$, and for $k\ge1$, $\tau_k=\inf\{t : t>\tau_{k-1}, X_t\ne X_{\tau_{k-1}}\}$, that is the k-th jump time of X. The transition matrix of Y is $\mathsf{M}=(\mathsf{M}_{i,j} : i,j\in\mathbb{N})$ defined as

$$\mathsf{M}_{i,j} = -\mathsf{G}_{i,j}/\mathsf{G}_{i,i}, \quad \forall i\ne j \qquad \text{and} \qquad \mathsf{M}_{i,i}=0 \quad \forall i\in\mathbb{N}.$$

The properties of positive recurrence, null recurrence and transience are inherited from the jump chain to the continuous chain under non-explosion assumptions (Theorem 3.4.1 and Theorem 3.5.3 [18]). Also, the knowledge of the (or an) invariant measure of one of these processes (either of Y or of X) allows one to deduce the corresponding invariant measure of the other, by using:

$$\pi_i = -\pi_i^C \mathsf{G}_{i,i}, \text{ for all } i\ge 0,$$

where π is the invariant measure for the discrete process Y.

The last important remark is: G is ◹ (resp. ◺) iff M is ◹ (resp. ◺). For this reason, our results apply to continuous Markov processes with ◺ and ◹ generator matrices G.

A.2 BDP and orthogonal polynomials

Karlin & McGregor approach relies on the study of the spectral properties of the tridiagonal transition matrix T (see notation in (1.4)) and its connection with a family of orthogonal polynomials $(Q_i, i \geq 0)$ defined as follows: set $Q_0(x) = 1$, $p_0 Q_1(x) = x - r_0$, and

$$xQ_j(x) = q_j Q_{j-1}(x) + r_j Q_j(x) + p_j Q_{j+1}(x),\ j \geq 1;$$

and this can be rewritten in the following form

$$Q(x) := \begin{bmatrix} Q_0(x) & Q_1(x) & Q_2(x) & \cdots \end{bmatrix}^t, \quad xQ(x) = \mathsf{T}Q(x). \tag{A.1}$$

Observe that $Q(x)$ is then an eigenvector of T associated with the eigenvalue x.

Karlin & McGregor [13] prove that there exists a unique measure ψ on $[-1, 1]$ for which the family $(Q_i, i \geq 0)$ forms an orthogonal family. More precisely

$$\pi_j \int_{-1}^{1} Q_i(x) Q_j(x) d\psi(x) = \mathbf{1}_{i=j},$$

where π is the invariant measure of T. Further,

$$(\mathsf{T}^n)_{i,j} = \pi_j \int_{-1}^{1} x^n Q_i(x) Q_j(x)\, d\,\psi(x). \tag{A.2}$$

Their approach is somehow more natural in the continuous settings: define the transition rate matrix

$$\mathsf{G} = \begin{bmatrix} -(\lambda_0+\mu_0) & \lambda_0 & 0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}, \tag{A.3}$$

and a second family of orthogonal polynomials $(\tilde{Q}_i, i \geq 0)$ as follows:

$$\begin{cases} -x\tilde{Q}_0(x) & = & -(\lambda_0+\mu_0)\tilde{Q}_0(x) + \lambda_0 \tilde{Q}_1(x), \\ -x\tilde{Q}_n(x) & = & \mu_n \tilde{Q}_{n-1}(x) - (\lambda_n+\mu_n)\tilde{Q}_n(x) + \lambda_n \tilde{Q}_{n+1}(x), \\ \tilde{Q}_0(x) & \equiv & 1 \end{cases} \tag{A.4}$$

so that (A.4) can be rewritten as

$$-x\tilde{Q}(x) = \mathsf{G}\tilde{Q}(x). \tag{A.5}$$

These polynomials are the orthogonal polynomials of a solvable Stieljes moment problem associated with a regular probability measure Ψ on $[0, +\infty)$. There exists a unique measure Ψ on $[0, +\infty)$ for which the $(\tilde{Q}_i, i \geq 0)$ forms an orthogonal family ([13]), more precisely, $\pi_j^C \int_{-1}^{1} \tilde{Q}_i(x)\tilde{Q}_j(x) d\Psi(x) = \mathbf{1}_{i=j}$, where π^C is the explicitly known invariant measure of the continuous-time process.

Set $P'(t) = P(t)\mathsf{G}$, for $t \geq 0$ and $P(0) = \mathsf{Id}$. The matrix $P(t)$ is the transition matrix of the continuous-time BD process, and $P_{i,j}(t)$ is the probability that the state of the chain is j at time t given that it started at time 0 in state i (a detail: μ_0 is not assumed to be 0, the case where absorption at 0 may occur is included). Then Karlin & McGregor defined "formally"

$$f_i(x,t) = \sum_{j\geq 0} P_{i,j}(t)\tilde{Q}_j(x) \tag{A.6}$$

and in vectorial notation

$$f(x,t) = P(t)\tilde{Q}(x) \Rightarrow \partial f(x,t)/\partial t = P'(t)\tilde{Q}(x) = P(t)\mathsf{G}\tilde{Q}(x) = -xf(x,t),$$

subject to the initial condition $f(x,0) = \tilde{Q}(x)$. From here $f(x,t) = \exp(-xt)\tilde{Q}(x)$, so that

$$f_i(x,t) = \exp(-xt)\tilde{Q}_i(x).$$

Now, reinterpret $f_i(x,t)$ on the L^2 space in which we are working in, equipped with its basis of orthogonal polynomials $(\tilde{Q}_j(x), j\geq 0)$. The extraction of $P_{i,j}(t)$ in (A.6) can be done using the orthogonality of the $\tilde{Q}_j$'s,

$$\mathbb{P}_{i,j}(t) = \left[\int f_i(x,t)\tilde{Q}_j(x)d\Psi(x)\right] / \left[\int \tilde{Q}_j(x)^2 d\Psi(x)\right]$$

(they set $\int \tilde{Q}_j(x)^2 d\Psi(x) = 1/\pi^C{}_j$).

The main point in the construction: the orthogonality of the polynomials, means, since $\tilde{Q}_0(x) = 1$, and $\int \tilde{Q}_0 d\Psi = 1$, that $\int \tilde{Q}_j d\Psi = 0$ for $j\geq 1$, the moments $\int x_n d\Psi$ can be expressed in the $\tilde{Q}_n$ (since $\tilde{Q}_n$ has degree n).

Karlin & McGregor constructed their study by establishing a correspondence between the set of matrices G of continuous-time BD processes, and the set of solvable Stieltjes moment problems. From here, the measure Ψ encodes somehow in an indirect way the polynomials (Q_i, $i\geq 0$) (as a transform), which are solution to $-x\tilde{Q}(x) = \mathsf{G}\tilde{Q}(x)$ and then they encode the spectral properties of G, which drives the behaviour of $P(t)$ (by (A.6)). The extraction of the recurrence criterion from here (see [12, pp. 370–376]) is done by expressing the recurrence in terms of a certain property of Ψ, which in turn, is shown to be expressible in terms of the coefficients of G (which provides the criteria given at the beginning of Section 1, in the discrete-time version of the BD process).

The methods developed by Karlin & McGregor are really elegant and satisfying from a theoretical point of view. These methods connect probability theory, algebra, measure theory (specifically the "moment problem") and the theory of orthogonal polynomials. However, the focus made on the map $\mathsf{G}\mapsto\Psi$, which is something that can be compared with the recourse of Fourier transform in other fields of probability theory, has the effect to lock a bit these studies, and this, for two reasons. The first one is that the correspondence is exact, so that, these tools are not simply available for any extension of the class of BD processes. Secondly, the measures Ψ are in general not known, nor computable, so that, Ψ is used as a formal encoding tool, rather than as a computing tool that helps to make some computations: there are just a handful of important cases in which it can be computed (see e.g. Schoutens [21]). The approach we propose is different since it is centred on direct computations of quantities of interests. Limitations exist, but they are not the same at all. The criteria of recurrence/transience we provide, do not rely on the moment's computation of any measure.

References

[1] L. Abolnikov and A. Dukhovny. Markov chains with transition delta-matrix: ergodicity conditions, invariant probability measures and applications. *J. Appl. Math. Stochastic Anal.*, 4(4):333–355, 1991. MR1136432

[2] D. J. Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM Journal on Discrete Mathematics*, 3(4):450–465, 1990. MR1069105

[3] P. Brémaud. *Markov Chains*. Springer International Publishing, 2020. MR1689633

[4] P. J. Brockwell, J. Gani, and S. I. Resnick. Birth, immigration and catastrophe processes. *Adv. in Appl. Probab.*, 14(4):709–731, 1982. MR0677553

[5] A. Z. Broder. Generating random spanning trees. In *FOCS*, vol. 89, pages 442–447, 1989.

[6] P. Cartier and D. Foata. *Problèmes combinatoires de commutation et réarrangements*. Springer-Verlag, Lecture Notes in Mathematics, 1969. MR0239978

[7] P. Flajolet and F. Guillemin. The formal theory of birth-and-death processes, lattice path combinatorics and continued fractions. *Adv. in Appl. Probab.*, 32(3):750–778, 2000. MR1788094

[8] L. Fredes and J.-F. Marckert. Aldous-Broder theorem: extension to the non reversible case and new combinatorial proof, 2021.

[9] T. Helmuth. Loop-weighted walk. *Ann. Inst. Henri Poincaré D*, 3(1):55–119, 2016. MR3462630

[10] Y. Hu, R. Lyons, and P. Tang. A reverse Aldous–Broder algorithm. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 57-2, pages 890–900. Institut Henri Poincaré, 2021. MR4260489

[11] S. Kapodistria, T. Phung-Duc, and J. Resing. Linear birth/immigration-death process with binomial catastrophes. *Probab. Engrg. Inform. Sci.*, 30(1):79–111, 2016. MR3436782

[12] S. Karlin and J. McGregor. The classification of birth and death processes. *Trans. Amer. Math. Soc.*, 86:366–400, 1957. MR0094854

[13] S. Karlin and J. McGregor. The differential equations of birth-and-death processes, and the Stieltjes moment problem. *Trans. Amer. Math. Soc.*, 85:489–546, 1957. MR0091566

[14] C. Krattenthaler. The theory of heaps and the Cartier–Foata monoid, 2006. MR2178686

[15] G. F. Lawler. *Loop-Erased Random Walk*, pages 197–217. Birkhäuser Boston, Boston, MA, 1999. MR1703133

[16] C. R. MacCluer. The many proofs and applications of Perron's theorem. *Siam Review*, 42(3):487–498, 2000. MR1786935

[17] P. Marchal. Loop-erased random walks, spanning trees and Hamiltonian cycles. *Electronic Communications in Probability*, 5:39–50, 2000. MR1736723

[18] J. R. Norris. *Markov Chains*. Cambridge University Press, 1998. MR1600720

[19] E. Pardoux. *Markov Processes and applications: Algorithms, Networks, Genome and Finance*, volume 796. John Wiley & Sons, 2008. MR2488952

[20] P. Pollett, H. Zhang, and B. J. Cairns. A note on extinction times for the general birth, death and catastrophe process. *J. Appl. Probab.*, 44(2):566–569, 2007. MR2340220

[21] W. Schoutens. *Stochastic Processes and Orthogonal Polynomials*, volume 146 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 2000. MR1761401

[22] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer Science & Business Media, 2006. MR2209438

[23] G. X. Viennot. Heaps of pieces, I: Basic definitions and combinatorial lemmas. In *Combinatoire énumérative*, pages 321–350. Springer, 1986. MR0927773

[24] D. B. Wilson. Generating random spanning trees more quickly than the cover time. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, pages 296–303, 1996. MR1427525

[25] D. Zeilberger. A combinatorial approach to matrix algebra. *Discrete Mathematics*, 56(1):61–72, 1985. MR0808086

Acknowledgments. We are grateful to the referees for their comments, suggestions and remarks, that helped us to improve the quality of the paper.