On the complex eigenvalues of finite Markov generators

Laurent Miclo

Toulouse School of Economics
Institut de Mathématiques de Toulouse
CNRS and University of Toulouse

Abstract

Dmitriev and Dynkin characterized the set of possible eigenvalues of Markov generators on a set of N points. We recover this result via decompositions in cycles. This probabilistic approach enables us to get more specific informations about the possible eigenvalues when assumptions are made on the lengths of the cycles entering in the decomposition.

Keywords: Finite Markov generators, finite transition kernels, eigenvalues, decomposition into cycles.

MSC2020: primary: 60J27, secondary: 15A18, 05C38, 05C50, 60J10.

Fundings: the grants ANR-17-EURE-0010 and AFOSR-22IOE016 are acknowledged.

1 Introduction: Dmitriev and Dynkin's result

Kolmogorov asked in 1938 for a characterisation of the set U_N of complex numbers which are eigenvalues of Markov transition kernels on a finite space of cardinal $N \in \mathbb{N}$. The survey talk of Besenyei [1] provides a nice account of the main steps involved in its resolution. The paper of Dmitriev and Dynkin [2] solves this question for $N \leqslant 5$ and contains important ideas for the general case. In particular they introduce a geometrical reformulation of the eigenvalues of finite Markov transition kernels, as well as a minimal angle property. The complete solution was finally given by Karpelevich [4]. The characterisation of the set T_N of complex numbers which are the opposite of eigenvalues of finite Markov generators on a finite space of cardinal $N \in \mathbb{N}$ is simpler and is already contained in Dmitriev and Dynkin [2]. Indeed, it is a consequence of the localisation of U_N around the eigenvalue 1, due to the conical structure of T_N.

Our purpose here is to present another approach to deduce T_N, more direct and more probabilistic in spirit. Furthermore, for any $\lambda \in T_N$, we provide a simple Markov generator admitting $-\lambda$ for eigenvalue (see the end of the proof of Theorem 1 below). This was the initial motivation for this investigation, stemming from the search of simple models for the classification of Markov generators via interweaving relations, see [7]. Our proof is based on the decompositions into cycles of finite Markov generators. It has the advantage to provide more specific informations about the localisation of the eigenvalues when some restrictions are put on such decompositions.

Let us be more precise. Consider V a finite set containing N points, $N \in \mathbb{N}$. Recall that a Markov generator on V is a matrix $L \coloneqq (L(x,y))_{x,y\in V}$ whose off-diagonal entries are non-negative and whose row sums vanish. Denote $\mathcal{L}(V)$ the set of all such matrices, so that T_N is the set of all the opposites of eigenvalues of matrices from $\mathcal{L}(V)$.

Our first task here will be to recover the following result:

Theorem 1 (Dmitriev and Dynkin [2]) *The set T_N is the set S_N of complex numbers whose angle with the positive real axis is smaller (in absolute value) or equal than $\frac{\pi}{2} - \frac{\pi}{N}$, with the convention $S_1 = \{0\}$.*

Recall that a Markov kernel on V is a matrix $K \coloneqq (K(x,y))_{x,y\in V}$ whose entries are non-negative and whose row sums are equal to 1. Denote $\mathcal{K}(V)$ the set of all such matrices, as well as U_N the set of all eigenvalues of all matrices from $\mathcal{K}(V)$. Let A_N be the acute angular sector delimited by the two half-lines starting from 1 and respectively going through $\exp\left(\frac{\mathrm{i}2\pi}{N}\right)$ and $\exp\left(\frac{-\mathrm{i}2\pi}{N}\right)$. The following result is contained in Dmitriev and Dynkin [2] and was the main ingredient in their proof of Theorem 1. We go the reverse way and deduce it from Theorem 1.

Proposition 2 (Dmitriev and Dynkin [2]) *We have*

$$U_N \quad \subset \quad A_N$$

For the full description of U_N, see Karpelevich [4] (or Dmitriev and Dynkin [2] for $N \leqslant 5$), or the survey of Besenyei [1].

The decompositions in cycles of finite irreducible Markov generators are recalled in the next section, where the alternative proofs of Theorem 1 and Proposition 2 are presented. This approach enables us to show in the last section that if a Markov generator L can be approximated by irreducible Markov generators with decompositions in cycles of orders bounded by $n \in [\![2, N]\!]$, then its eigenvalues belong to T_n, independently from N. When L is non-transient, the existence of such approximations can be characterized directly on the irreducible classes of L. In the general case, one has to be more careful, as it will be illustrated on an example.

2 A proof by decompositions in cycles

Here we recover the results by Dmitriev and Dynkin [2] recalled in the introduction, via decompositions in cycles.

A Markov generator $L \in \mathcal{L}(V)$ is said to be irreducible when $\exp(L)$ has only positive entries. Denote $\mathcal{I}(V)$ the set of irreducible Markov generators on V and R_N the set of all opposites of eigenvalues of matrices from $\mathcal{I}(V)$. A first observation is:

Lemma 3 *We have $R_N \subset T_N \subset \mathrm{cl}(R_N)$, the closure of R_N.*

We will check in the proof of Theorem 1 that $T_N = R_N$.

Proof

The first inclusion is obvious, since $\mathcal{I}(V) \subset \mathcal{L}(V)$.

For the second inclusion, let us recall the following result.

Let A and B be two $V \times V$-matrices with complex entries and for any $z \in \mathbb{C}$, consider the matrix $A(z)$ given by

$$A(z) \quad \coloneqq \quad A + zB$$

For $z \in \mathbb{C}$, denote $\sigma(z) \subset \mathbb{C}$ the set of eigenvalues of $A(z)$. Let $\mathfrak{K}$ be the set of non-empty compact subsets of $\mathbb{C}$, endowed with the Hausdorff distance. From Section 1 of Chapter 2 of Kato [5], the mapping $\sigma : \mathbb{C} \to \mathfrak{K}$ is continuous.

Consider $\lambda \in T_N$, as well as $L \in \mathcal{L}(V)$ such that $-\lambda$ is an eigenvalue of L. Define the following Markov generators

$$\forall\, \epsilon \geqslant 0, \qquad L_\epsilon \quad \coloneqq \quad L + \epsilon J \tag{1}$$

where J is the $V \times V$-matrix whose off-diagonal entries are equal to $1/N$ and whose diagonal entries are equal to $-(1 - 1/N)$. For $\epsilon > 0$, we have $L_\epsilon \in \mathcal{I}(V)$. From the above continuity result, λ is approached by eigenvalues of L_ϵ for small $\epsilon > 0$, so that $\lambda \in \mathrm{cl}(R_N)$ as wanted. ■

Let us recall the decompositions of elements from $\mathcal{I}(V)$ into cycles. We need some notations. A cycle C of V is an injective mapping from $\mathbb{Z}_n \coloneqq \mathbb{Z}/(n\mathbb{Z})$ to V, where $n \in [\![N]\!]$ is called the order of C. Two cycles C and C' are said to be equivalent if they coincide up to a translation, i.e., if have the same order $n \in [\![N]\!]$ and if there exists $m \in \mathbb{Z}_n$ such that

$$\forall\, l \in \mathbb{Z}_n, \qquad C'(l) \quad = \quad C(l+m)$$

The set of the corresponding equivalent classes is denoted $\mathcal{C}(V)$. To simplify the notations, such equivalent classes will be identified with representative cycles.

Fix a Markov generator $L \in \mathcal{L}(V)$. By irreducibility, it admits a unique invariant measure π, which furthermore give a positive weight to all points of V. Given $C \in \mathcal{C}(V)$ of order $n \in [\![2, N]\!]$, we introduce the Markov generator $\Lambda_{\pi,C}$ via

$$\forall\, x \neq y \in V, \qquad \Lambda_{\pi,C}(x,y) \quad \coloneqq \quad \begin{cases} \dfrac{1}{\pi(x)} & \text{, if } \exists\, m \in \mathbb{Z}_n \text{ such that } x = C(m) \text{ and } y = C(m+1) \\ 0 & \text{, otherwise} \end{cases}$$

(the diagonal entries are such that the row sums vanishes: for any $x \in V$, $L(x,x) = -1/\pi(x)$). Note that $\Lambda_{\pi,C}$ does not depend on the representative cycle C of the equivalence class from $\mathcal{C}(V)$. When C is of order 1, by convention we take $\Lambda_{\pi,C} = 0$.

A decomposition of L into cycles is the writing

$$L \quad = \quad \sum_{C \in \mathcal{C}(V)} a(C)\Lambda_{\pi,C} \tag{2}$$

where $(a(C))_{C\in\mathcal{C}(V)}$ is a family of non-negative numbers. The existence of such a decomposition can be found in [6], see also the book of Kalpazidou [3] for an extensive discussion. In general such a decomposition is not unique.

When looking for a decomposition in cycles, it seems we first need to know the invariant measure π. On the contrary, it provides a way to find this invariant probability:

Lemma 4 *Assume that a decomposition such as* (2) *holds for some $L\in\mathcal{L}(V)$ and some probability π (giving a positive weight to the images of all cycles C with $a(C)>0$). Then π is invariant for L.*

Proof

Indeed, note that π is invariant for any Markov generator of the form $\Lambda_{\pi,C}$ with $C\in\mathcal{C}(V)$. ∎

Illustrations of this result will be provided by Examples 7 and **??** in the sequel.

The main ingredient of the proof of Theorem 1 is the following result.

Lemma 5 *For any $C\in\mathcal{C}(V)$ and $\varphi\in\mathbb{C}^V$, we have*

$$\pi[\bar\varphi\Lambda_{\pi,C}[\varphi]] \quad\in\quad S_N$$

where $\bar\varphi$ is the function from $\mathbb{C}^V$ taking the conjugate values of φ.

Proof

For any $\varphi\in\mathbb{C}^V$, we compute

$$\begin{aligned}
\pi[\bar\varphi\Lambda_{\pi,C}[\varphi]] &= \sum_{x,y\in V}\pi(x)\Lambda_{\pi,C}(x,y)\bar\varphi(x)(\varphi(y)-\varphi(x))\\
&= \sum_{x\in C(\mathbb{Z}_n)}\pi(x)\Lambda_{\pi,C}(x,C(C^{-1}(x)+1))\bar\varphi(x)(\varphi(C(C^{-1}(x)+1))-\varphi(x))\\
&= \sum_{x\in C(\mathbb{Z}_n)}\bar\varphi(x)(\varphi(C(C^{-1}(x)+1))-\varphi(x))\\
&= \sum_{z\in\mathbb{Z}_n}\bar\phi(z)(\phi(z+1)-\phi(z))
\end{aligned}$$

where n is the order of C and $\phi\coloneqq\varphi\circ C$.

For any function $\phi:\mathbb{Z}_n\to\mathbb{C}$, there exists a family of complex coefficients $(b(k))_{k\in\mathbb{Z}_N}$ such that

$$\forall\ z\in\mathbb{Z}_n,\qquad \phi(z) \quad=\quad \sum_{k\in\mathbb{Z}_n} b(k)\exp\left(\frac{2\pi \mathrm{i} kz}{N}\right)$$

It follows that

$$\begin{aligned}
\sum_{z\in\mathbb{Z}_n}\bar\phi(z)(\phi(z+1)-\phi(z)) &= \sum_{z\in\mathbb{Z}_n}\sum_{k,l\in\mathbb{Z}_n}\bar b(k)b(l)\exp\left(-\frac{2\pi\mathrm{i}kz}{N}\right)\left(\exp\left(\frac{2\pi\mathrm{i}l(z+1)}{N}\right)-\exp\left(\frac{2\pi\mathrm{i}lz}{N}\right)\right)\\
&= \sum_{z\in\mathbb{Z}_n}\sum_{k,l\in\mathbb{Z}_n}\bar b(k)b(l)\left(\exp\left(\frac{2\pi\mathrm{i}l}{N}\right)-1\right)\exp\left(\frac{2\pi\mathrm{i}(l-k)z}{N}\right)\\
&= \sum_{k,l\in\mathbb{Z}_n}\bar b(k)b(l)\left(\exp\left(\frac{2\pi\mathrm{i}l}{N}\right)-1\right)\sum_{z\in\mathbb{Z}_n}\exp\left(\frac{2\pi\mathrm{i}(l-k)z}{N}\right)\\
&= n\sum_{k\in\mathbb{Z}_n}\bar b(k)b(k)\left(\exp\left(\frac{2\pi\mathrm{i}k}{N}\right)-1\right)\\
&= n\sum_{k\in\mathbb{Z}_n}|b(k)|^2\left(\exp\left(\frac{2\pi\mathrm{i}k}{N}\right)-1\right)
\end{aligned}$$

Note that for any $k \in \mathbb{Z}_n$, we have

$$\exp\left(\frac{2\pi \mathrm{i}k}{N}\right) - 1 \;\in\; S_n \;\subset\; S_N$$

so taking into account that S_N is a cone centered at 0, we get the desired belonging. ■

We can now come to the

Proof of Theorem 1

It is sufficient to show that $R_N = S_N$. Indeed, from Lemma 3 and the fact that S_N is clearly closed, we conclude that $T_N = R_N = S_N$.

We start with the inclusion $R_N \subset S_N$.

Consider $\lambda \in R_N$ and $L \in \mathcal{I}(V)$ and $\varphi \in \mathbb{C}^V \backslash \{0\}$ such that $L[\varphi] = -\lambda\varphi$. Decompose L as in (2), so we can write

$$\pi[\bar{\varphi}L[\varphi]] \;=\; \sum_{C \in \mathcal{C}(V)} a(C)\pi[\bar{\varphi}\Lambda_{\pi,C}[\varphi]]$$

It follows from Lemma 5 and the fact that S_N is a cone centered at 0 that

$$\pi[\bar{\varphi}L[\varphi]] \;\in\; S_N$$

It remains to write

$$\lambda \;=\; \frac{\pi[\bar{\varphi}L[\varphi]]}{\pi[\bar{\varphi}\varphi]} \;\in\; S_N$$

to get the wanted inclusion.

Conversely, consider $\lambda \in S_N$. There exist two non-negative real numbers r, s such that either

$$\lambda \;=\; r\left(1 - \exp\left(\frac{2\pi \mathrm{i}}{N}\right)\right) + s \tag{3}$$

or

$$\lambda \;=\; r\left(1 - \exp\left(\frac{-2\pi \mathrm{i}}{N}\right)\right) + s \tag{4}$$

To construct a Markov generator admitting $-\lambda$ as eigenvalue in the first case, identify V with $\mathbb{Z}_N$ and introduce the Markov generator $\widehat{L}$ given by

$$\forall\ x \neq y \in \mathbb{Z}_N, \qquad \widehat{L}(x,y) \;\coloneqq\; \begin{cases} 1 & \text{, if } y = x+1 \\ 0 & \text{, otherwise} \end{cases}$$

We compute that

$$\begin{aligned} \widehat{L}[\varphi] &= -\left(1 - \exp\left(\frac{2\pi \mathrm{i}}{N}\right)\right)\varphi \\ J[\varphi] &= -\varphi \end{aligned}$$

where J was defined after (1) and φ is the function defined by

$$\forall\ x \in \mathbb{Z}_N, \qquad \varphi(x) \;\coloneqq\; \exp\left(\frac{2\pi \mathrm{i}x}{N}\right)$$

It follows that $-\lambda$ is an eigenvalue of the Markov generator $L \coloneqq r\widehat{L} + sJ$.

For (4), it is enough to replace $\widehat{L}$ by $\check{L}$ given by

$$\forall\ x \neq y \in \mathbb{Z}_N, \qquad \check{L}(x,y) \ := \ \begin{cases} 1 & \text{, if } y = x-1 \\ 0 & \text{, otherwise} \end{cases}$$

■

Let us end this section with the very short

Proof of Proposition 2

For any $K \in \mathcal{K}(V)$, the matrix $K - I$ is a Markov generator, where I is the $V \times V$ identity matrix. It follows that for any $\theta \in U_N$, we have $-(\theta - 1) \in T_N$. It remains to note that A_N is the image of S_N by the mapping $\mathbb{C} \ni \theta \mapsto 1 - \theta$. ■

3 Extensions

Here we present some improvement of Theorem 1.

For any $L \in \mathcal{I}(V)$, define the order $\nu(L)$ as the minimal $n \in [\![N]\!]$, such that L can be decomposed under the form (2) where $a(C) = 0$ for all cycles C of order strictly larger than n. Define

$$\mathcal{I}_n(V) \ := \ \{L \in \mathcal{I}(V) \,:\, \nu(L) \leqslant n\} \tag{5}$$

Note that $\mathcal{I}_1(V) = \varnothing$ as soon as $N \geqslant 2$. In the trivial case where V is a singleton, we have $\mathcal{I}_1(V) = \mathcal{L}(V) = \{0\}$. The set $\mathcal{I}_2(V)$ corresponds to irreducible and reversible Markov generators.

Define $\mathcal{L}_n(V)$ as the closure of $\mathcal{I}_n(V)$ in the set of $V \times V$ matrices. We have $\mathcal{L}_n(V) \subset \mathcal{L}(V)$ by closure of $\mathcal{L}(V)$. Denote $T_{N,n}$ the set of all the opposites of eigenvalues of matrices from $\mathcal{L}_n(V)$. From a slight modification of the proof of Theorem 1, we get:

Theorem 6 *For $n \in [\![2, N]\!]$, $T_{N,n}$ coincides with S_n.*

This result is not true for $n = 1$ and $N \geqslant 2$, since $T_{N,1} = \varnothing$ and $S_1 = \{0\}$.

Proof

Denote $R_{N,n}$ the set of all the opposites of eigenvalues of matrices from $\mathcal{I}_n(V)$. Lemma 5 extends into $R_{N,n} \subset T_{N,n} \subset \mathrm{cl}(R_{N,n})$. This is now a consequence of Theorem 5.1 page 107 of Kato's book [5] showing the continuity of the eigenvalues in terms of the matrix.

Following the first part of the proof of Theorem 1, we get that

$$R_{N,n} \ \subset \ S_n \tag{6}$$

Let us marginally modify the second part of the proof of Theorem 1. We begin by remarking that any $\lambda \in S_n$ can be written under the form (3) or (4), with N replaced by n and still $r, s \geqslant 0$.

Consider the first case. Choose a cycle C of order n and introduce the Markov generator $\widehat{L}$ via

$$\forall\ x \neq y \in V, \qquad \widehat{L}(x,y) \ := \ \begin{cases} 1 & \text{, if } x, y \in C(\mathbb{Z}_n) \text{ and } y = C(C^{-1}(x)+1) \\ 0 & \text{, otherwise} \end{cases}$$

Define the function φ by

$$\forall\ x \in \mathbb{Z}_N, \qquad \varphi(x) \ := \ \begin{cases} \exp\left(\frac{2\pi \mathrm{i} C^{-1}(x)}{n}\right) & \text{, if } x \in C(\mathbb{Z}_n) \\ 0 & \text{, otherwise} \end{cases}$$

we compute that

$$\widehat{L}[\varphi] \ = \ -\left(1 - \exp\left(\frac{2\pi\mathrm{i}}{n}\right)\right)\varphi$$

Furthermore, we still have

$$J[\varphi] \quad = \quad -\varphi$$

where J was defined after (1).

It follows that $-\lambda$ is an eigenvalue of the Markov generator $L \coloneqq r\widehat{L} + sJ$.

Note that $J \in \mathcal{I}_2(V)$, as an irreducible Markov generator reversible with respect to the uniform probability. The uniform probability is also invariant for $\widehat{L}$ and thus for L. It follows that when $s > 0$, $L \in \mathcal{I}_n(V)$ (recall that $n \geqslant 2$), so that $\lambda \in R_{N,n}$. When $s = 0$, we have

$$\begin{aligned} \lambda &= r\left(1 - \exp\left(\frac{2\pi \mathrm{i}}{n}\right)\right) \\ &= \lim_{s \to 0_+} r\left(1 - \exp\left(\frac{2\pi \mathrm{i}}{n}\right)\right) + s \\ &\in \mathrm{cl}(R_{N,n}) \end{aligned}$$

We deduce $S_n \subset \mathrm{cl}(R_{N,n})$, and in connexion with (6), that $R_{N,n} = S_n$ and finally $T_{N,n} = S_n$.

The case (4), with N replaced by n, is treated similarly. ■

Let us give an example of kind of improvements one can expect from this result.

Example 7 For $k \in \mathbb{N}$, consider the state space $V \coloneqq \mathbb{Z}_{2k}^2$. An element $x \coloneqq (x_1, x_2) \in V$ is said to be even (respectively odd) if $x_1 + x_2 = 0$ (resp. $x_1 + x_2 = 1$) in $\mathbb{Z}_2$ seen as $\mathbb{Z}_{2k}/(2\mathbb{Z}_k)$.

We endow V with the generator L given by

$$\forall\ x \coloneqq (x_1, x_2) \neq y \coloneqq (y_1, y_2) \in V, \qquad L(x,y) \quad \coloneqq \quad \begin{cases} 1 & \text{, if } x \text{ is even, } y_1 = x_1 \pm 1 \text{ and } y_2 = x_2 \\ 1 & \text{, if } x \text{ is odd, } y_1 = x_1 \text{ and } y_2 = x_2 \pm 1 \\ 0 & \text{, otherwise} \end{cases}$$

Denote $e_1 \coloneqq (1,0)$, $e_2 \coloneqq (0,1)$ and for any $x \in V$ even, let $C_{x,1}$ (respectively $C_{x,-1}$) the cycle $(x, x+e_1, x+e_1+e_2, x+e_2)$ (resp. $(x, x+e_1, x+e_1-e_2, x-e_2)$), identified with its sequential values on $\mathbb{Z}_4$. Let π be the uniform probability on V.

We have the cycle decomposition

$$L \quad = \quad \frac{1}{2^{2k+1}} \sum_{x \in V, \text{even},\, s \in \{\pm 1\}} \Lambda_{\pi, C_{x,s}}$$

showing first that π is invariant for L, via Lemma 4, and second that $L \in \mathcal{I}_4(V)$.

Thus Theorem 6 implies that the spectrum of L is included into S_4, contrary to Theorem 1, which only enables us to get that the spectrum of L is included into S_{4k^2}.

While these considerations can be extended to different factors $V \coloneqq Z_{2k} \times Z_{2l}$, with $k \neq l \in \mathbb{N}$, it is not so clear how to construct an analogue example in higher dimensions.

□

Our next goal is to introduce a subset $\mathcal{J}_n(V)$ of $\mathcal{L}_n(V)$ generalizing $\mathcal{I}_n(V)$.

Let $L \in \mathcal{L}(V)$ be given. A non-empty subset $A \subset V$ is said to be **irreducible** (relatively to L) when the $A \times A$ matrix $\exp(L_A)$ has only positive entries, where $L_A \coloneqq (L(x,y))_{x,y \in A}$. In particular, the singletons are irreducible. The irreducible subset A is furthermore said to be an **irreducible class**, when it is maximal for the inclusion among all irreducible subsets. Denote $\mathcal{A}$ the set of irreducible classes. The elements of $\mathcal{A}$ form a partition of the state space:

$$V \quad = \quad \bigsqcup_{A \in \mathcal{A}} A$$

For $A \in \mathcal{A}$, consider the Markov generator L'_A on A obtained from L_A by possibly modifying the diagonal entries to make the row sums vanish. Since $L'_A \in \mathcal{I}(A)$, its order $\nu(L'_A)$ was already defined. We extend this definition to L via:

$$\nu(L) \quad \coloneqq \quad \max\{\nu(L'_A)\,:\, A \in \mathcal{A}\}$$

For $L \in \mathcal{I}(V)$, we have $\mathcal{A} = \{V\}$ and this new notion of order $\nu(L)$ coincides with the previous one for irreducible Markov generators.

For any $n \in [\![N]\!]$, define

$$\mathcal{J}_n(V) \quad \coloneqq \quad \{L \in \mathcal{L}(V)\,:\, \nu(L) \leqslant n\}$$

In particular we have $\mathcal{I}_n(V) \subset \mathcal{J}_n(V)$. Our purpose here is to show the following result:

Theorem 8 *For any $n \in [\![2, N]\!]$, we have $\mathcal{J}_n(V) \subset \mathcal{L}_n(V)$.*

From the observation following (5), this result is trivially true for $n = 1$.

Before proving this inclusion, let us present a preliminary result.

Introduce a relation $\rightarrowtail$ on $\mathcal{A}$ via

$$\forall\ A \neq A' \in \mathcal{A}, \qquad A \rightarrowtail A' \quad \Leftrightarrow \quad \exists\ x \in A, \exists\ y \in A'\,:\, L(x, y) > 0$$

We have for any fixed $L \in \mathcal{L}(V)$:

Lemma 9 *There exists a bijection F between $\mathcal{A}$ and $[\![|\mathcal{A}|]\!]$ (recall that $|\mathcal{A}| \in [\![N]\!]$ stands for the cardinal of $\mathcal{A}$) such that*

$$\forall\ A, A' \in \mathcal{A}, \qquad A \rightarrowtail A' \quad \Rightarrow \quad F(A) > F(A')$$

Proof

An irreducible class $A \in \mathcal{A}$ is said to be **recurrent**, when L_A is a Markov generator on A, i.e. when $L'_A = L_A$. Or equivalently, if there is no $A' \in \mathcal{A}\backslash\{A\}$ such that $A \rightarrowtail A'$. An irreducible class $A \in \mathcal{A}$ which is not recurrent is said to be **transient**. The set of recurrent and irreducible classes are respectively written $\mathcal{R}$ and $\mathcal{T}$. Note that we always have $\mathcal{R} \neq \varnothing$ but it may happen that $\mathcal{T} = \varnothing$. In the latter case L is said to be **non-transient**.

When L is non-transient, there is no $\rightarrowtail$ between the elements of $\mathcal{A}$, so we can choose for F any bijection between $\mathcal{A}$ and $[\![|\mathcal{A}|]\!]$. Otherwise, for any $A \in \mathcal{T}$, consider the longest finite family $(A_1, ..., A_{\ell(A)})$ of irreducible classes with $A_1 = A$, $A_{\ell(A)} \in \mathcal{R}$ and for any $l \in [\![\ell(A) - 1]\!]$, $A_l \rightarrowtail A_{l+1}$. All the elements of this family are distinct. Indeed, otherwise, if $A_k = A_l$ with $l < k \in [\![|A|]\!]$, then $A_k \sqcup A_{k+1} \sqcup \cdots \sqcup A_l$ is an irreducible subset, a contradiction. Next choose $A_1 \in \mathcal{T}$ with the largest $\ell(A_1)$ among the elements of $\mathcal{T}$. There is no $A' \in \mathcal{T}$ such that $A' \rightarrowtail A_1$, otherwise we would have $\ell(A') \geqslant \ell(A_1) + 1$. Then we define $F(A_1) \coloneqq |\mathcal{A}|$. Next we choose A_2 in $\mathcal{T}\backslash\{A_1\}$ with largest possible $\ell(A_2)$ among the elements of $\mathcal{T}\backslash\{A_1\}$. There is no $A' \in \mathcal{T}\backslash\{A_1\}$ such that $A' \rightarrowtail A_2$, otherwise we would have $\ell(A') \geqslant \ell(A_2) + 1$. We take $F(A_2) \coloneqq |\mathcal{A}| - 1$. We keep following this procedure until $\mathcal{T}$ is exhausted: $\mathcal{T} = \{A_1, A_2, ..., A_{|\mathcal{T}|}\}$ and

$$\forall\ l \in [\![|\mathcal{T}|]\!], \qquad F(A_l) \quad = \quad |\mathcal{A}| - l + 1$$

It remains to choose arbitrary the values $F(A)$ in $[\![|\mathcal{A}| - |\mathcal{T}|]\!]$ for $A \in \mathcal{R}$, to get a wanted function F. ■

We can now come to the

Proof of Theorem 8

Fix $n \in [\![2, N]\!]$ and $L \in \mathcal{J}_n(V)$. To prove the wanted inclusion, it is sufficient to construct a family $(L_\epsilon)_{\epsilon\in(0,1)}$ from $\mathcal{I}_n(V)$ such that

$$\lim_{\epsilon\to 0_+} L_\epsilon = L \tag{7}$$

For any $A \in \mathcal{A}$, denote π_A the invariant probability on A associated to L'_A.

For any $\epsilon \in (0,1)$, introduce the probability π_ϵ given on V by

$$\forall\, x \in V, \qquad \pi_\epsilon(x) \coloneqq \frac{\epsilon^{F(A)}\pi_A(x)}{Z_\epsilon}$$

where $A \in \mathcal{A}$ is such that $x \in A$, F is a mapping as in Lemma 9, and Z_ϵ is the normalising constant.

Define the Markov generator L_ϵ via

$$\forall\, x \neq y \in V, \qquad L_\epsilon(x,y) \coloneqq \begin{cases} L(x,y) & \text{, if } L(x,y) > 0 \\ L(y,x)\pi_\epsilon(y)/\pi_\epsilon(x) & \text{, if } (x,y) \in \mathcal{E} \\ 0 & \text{, otherwise} \end{cases}$$

where

$$\mathcal{E} \coloneqq \{(x,y) \in V^2 \,:\, L(y,x) > 0 \text{ and } x, y \text{ belong to different irreducible classes}\}$$

Consider $(x,y) \in \mathcal{E}$ and let $A \neq A' \in \mathcal{A}$ be such that $x \in A$ and $y \in A'$. We have $A' \rightarrowtail A$, so that $F(A') > F(A)$ and thus

$$\lim_{\epsilon\to 0_+} \frac{\pi_\epsilon(y)}{\pi_\epsilon(x)} = 0$$

It follows that

$$\lim_{\epsilon\to 0_+} L_\epsilon(x,y) = 0$$

and (7) holds.

For $(x,y) \in \mathcal{E}$, introduce the Markov generator $L_{\epsilon,(x,y)}$ defined by

$$\forall\, x' \neq y' \in V, \qquad L_{\epsilon,(x,y)}(x',y') \coloneqq \begin{cases} L_\epsilon(x',y') & \text{, if } \{x',y'\} = \{x,y\} \\ 0 & \text{, otherwise} \end{cases}$$

Note that π_ϵ is reversible for this Markov generator and thus invariant.

Furthermore π_ϵ is also invariant for the Markov generators L'_A (seen as operators on V by completing this matrix by null entries outside A), for any $A \in \mathcal{A}$. Since we can write

$$L_\epsilon = \sum_{A\in\mathcal{A}} L'_A + \sum_{(x,y)\in\mathcal{E}} L_{\epsilon,(x,y)}$$

it follows that π_ϵ is invariant for L_ϵ and even its unique invariant probability by irreducibility.

For any $A \in \mathcal{A}$, since $L \in \mathcal{J}_n(V)$, we have $L'_A \in \mathcal{I}_n(A)$, so we can find coefficients $(a(C))_{C\in\mathcal{C}_n(A)}$ such that

$$\begin{aligned} L'_A &= \sum_{C\in\mathcal{C}_n(A)} a(C)\Lambda_{\pi_A,C} \\ &= \sum_{C\in\mathcal{C}_n(A)} \frac{a(C)\epsilon^{F(A)}}{Z_\epsilon}\Lambda_{\pi_\epsilon,C} \end{aligned}$$

Furthermore, for $(x,y)\in\mathcal{E}$, seeing (x,y) as a cycle, we can write

$$L_{\epsilon,(x,y)} \;=\; \pi_\epsilon(y)L(y,x)\Lambda_{\pi_\epsilon,(x,y)}$$

so that

$$L_\epsilon \;=\; \sum_{C\in\mathcal{C}_n(A)} \frac{a(C)\epsilon^{F(A)}}{Z_\epsilon}\Lambda_{\pi_\epsilon,C} + \sum_{(x,y)\in\mathcal{E}} \pi_\epsilon(y)L(y,x)\Lambda_{\pi_\epsilon,(x,y)}$$

corresponds a decomposition in cycles of L_ϵ. These cycles have orders bounded above by $2\vee n = n$, since it is assumed to belong to $\mathbb{N}\backslash\{1\}$. Thus L can be approximated by elements of $\mathcal{I}_n(V)$ and by consequence belongs to $\mathcal{L}_n(V)$. ■

The following example shows that in general we have $\mathcal{J}_n(V)\neq\mathcal{L}_n(V)$.

Example 10 Fix $l\in\mathbb{N}$ and consider the finite state space $V:=\mathbb{Z}_{2l}\sqcup\{\infty\}$. For any $\epsilon\geqslant 0$, introduce the Markov generator L_ϵ given by

$$\forall\; x\neq y\in V,\qquad L_\epsilon(x,y) \;:=\; \begin{cases} 1 & \text{, if } x,y\in\mathbb{Z}_{2l} \text{ and } y=x+1\\ 1 & \text{, if } x\in\{0,l\} \text{ and } y=\infty\\ \epsilon & \text{, if } x=\infty \text{ and } y\in\{0,l\}\\ 0 & \text{, otherwise}\end{cases}$$

where the elements of $\mathbb{Z}_{2l}$ where also identified with their representatives in $[\![0,2l-1]\!]$.

For any $\epsilon>0$, we have $L_\epsilon\in\mathcal{I}(V)$ and

$$\lim_{\epsilon\to 0_+} L_\epsilon \;=\; L_0 \;\notin\; \mathcal{I}(V) \tag{8}$$

Write $L:=L_0$. Taking into account the terminology introduced in the proof of Lemma 9, relatively to L, the points of $\mathbb{Z}_{2l}$ form a transient irreducible class and the only recurrent point is ∞. We have

$$\begin{aligned}\nu(L'_{\mathbb{Z}_{2l}}) &= 2l\\ \nu(L_{\{\infty\}}) &= 1\end{aligned}$$

It follows that $\nu(L)=2l$ and $L\in\mathcal{J}_{2l}(V)$, but $L\notin\mathcal{J}_{2l-1}(V)$ and thus $L\notin\mathcal{J}_{l+2}(V)$ as soon as $l\geqslant 3$. Let us check that $L\in\mathcal{L}_{l+2}(V)$. We will infer that for $l\geqslant 3$, we have $L\in\mathcal{L}_{l+2}(V)\backslash\mathcal{J}_{l+2}(V)$.

For $\epsilon>0$, consider the probability π_ϵ given on V by

$$\forall\; x\in V,\qquad \pi_\epsilon(x) \;:=\; \frac{1}{1+2l\epsilon}\begin{cases}\epsilon & \text{, if } x\in\mathbb{Z}_{2l}\\ 1 & \text{, if } x=\infty\end{cases}$$

Consider the two cycles $C_1:=(\infty,0,1,...,l)$ and $C_2:=(\infty,l,l+1,\cdots 2l-1,0)$. It appears that

$$L_\epsilon \;=\; \frac{\epsilon}{1+2l\epsilon}\Lambda_{\pi_\epsilon,C_1} + \frac{\epsilon}{1+2l\epsilon}\Lambda_{\pi_\epsilon,C_2}$$

From Lemma 4, we deduce that π_ϵ is the invariant probability of L_ϵ. Furthermore, $\nu(L_\epsilon)$ is bounded above by $l+2$, the order of C_1 and C_2 (for $l\geqslant 3$, $\nu(L_\epsilon)=l+2$). From the convergence (8), we get $L\in\mathcal{L}_{l+2}(V)$.

□

The previous example leaves open the problem of a more concrete description of $\mathcal{L}_n(V)$. Nevertheless if our attention is restricted to non-transient Markov generators, the situation is clearer. Denote $\mathcal{N}(V)$ the set of non-transient Markov generators.

Proposition 11 *For any $n\in[\![2,N]\!]$, we have $\mathcal{N}(V)\cap\mathcal{J}_n(V)=\mathcal{N}(V)\cap\mathcal{L}_n(V)$.*

For $n = 1$ and $N \geqslant 2$, this equality does not hold: $\mathcal{N}(V) \cap \mathcal{J}_1(V) = \{0\}$ while $\mathcal{N}(V) \cap \mathcal{L}_1(V) = \varnothing$.

Proof

Due to Theorem 8, it is sufficient to show that $\mathcal{N}(V) \cap \mathcal{L}_n(V) \subset \mathcal{N}(V) \cap \mathcal{J}_n(V)$, for given $n \in [\![2, N]\!]$.

Note that $0 \in \mathcal{N}(V) \cap \mathcal{J}_n(V)$, so consider $L \in \mathcal{N}(V) \cap \mathcal{L}_n(V) \setminus \{0\}$ and a sequence $(L_k)_{k \in \mathbb{Z}_+}$ from $\mathcal{I}_n(V)$ converging to L. Denote $(\pi_k)_{k \in \mathbb{Z}_+}$ the corresponding sequence of invariant probability measures and for $k \in \mathbb{Z}_+$,

$$L_k \quad = \quad \sum_{C \in \mathcal{C}(V)} a_k(C) \Lambda_{\pi_k, C} \tag{9}$$

a cycle decomposition with $a_k(C) = 0$ for any cycle C of order strictly larger than n.

Let $A_1, ..., A_\ell$ be the recurrent irreducible classes of L: since $L \in \mathcal{N}(V)$, we have $V = A_1 \sqcup A_2 \sqcup \cdots \sqcup A_\ell$.

We will need the following observation:

Lemma 12 *There exists $\eta \in (0,1)$ and $k_1 \in \mathbb{Z}_+$ such that*

$$\forall\ k \geqslant k_1,\ \forall\ l \in [\![\ell]\!],\ \forall\ x, y \in A_l, \qquad \eta \pi_k(x) \ \leqslant\ \pi_k(y) \ \leqslant\ \frac{1}{\eta} \pi_k(x)$$

Proof

Introduce

$$\begin{aligned} \lambda &\coloneqq \max\{|L(x,x)| \ :\ x \in V\} \\ K &\coloneqq \frac{L}{\lambda} + I \end{aligned}$$

as well as for $k \in \mathbb{Z}_+$,

$$\begin{aligned} \lambda_k &\coloneqq \max\{|L_k(x,x)| \ :\ x \in V\} \\ K_k &\coloneqq \frac{L_k}{\lambda_k} + I \end{aligned}$$

The matrix K_k is a Markov kernel on V and π_k is an invariant probability for K_k, i.e. $\pi_k = \pi K_k$. Furthermore, we have (since $L \neq 0$),

$$\lim_{k \to \infty} K_k \quad = \quad K$$

Consider

$$\begin{aligned} E &\coloneqq \{(x,y) \in V \ :\ K(x,y) > 0\} \ = \ \{(x,y) \in V \ :\ L(x,y) > 0\} \\ \epsilon &\coloneqq \min\{K(x,y) \ :\ (x,y) \in E\} \ = \ \frac{1}{\lambda} \min\{L(x,y) \ :\ (x,y) \in E\} \end{aligned}$$

We prove the above lemma with $\eta \coloneqq (\epsilon/2)^{|E|-1}$.

Define

$$k_1 \quad \coloneqq \quad \min\{k \in \mathbb{Z}_+ \ :\ \forall\ j \geqslant k,\ \forall\ (x,y) \in E,\ K_j(x,y) \geqslant \epsilon/2\}$$

Fix $l \in [\![\ell]\!]$ and $x, y \in A_l$. Consider an injective sequence $(x_0, x_1, ..., x_m)$ with $x_0 = x$, $x_m = y$ and $(x_j, x_{j+1}) \in E$ for all $j \in [\![0, m-1]\!]$. We have for $k \geqslant k_1$,

$$\begin{aligned} \pi_k(y) &\geqslant \pi_k(x) K_k(x, x_1) K_k(x_1, x_2) \cdots K_k(x_{m-1}, y) \\ &\geqslant \pi_k(x)(\epsilon/2)^m \end{aligned}$$

$$\geqslant \quad \eta \pi_k(x)$$

which is the wanted bound. ■

Consider $C : \mathbb{Z}_m \to V$ a cycle whose image is not included into one of the irreducible classes of L: we can find $j \in \mathbb{Z}_m$ such that $x \coloneqq C(j) \in A_l$ and $y \coloneqq C(j+1) \in A_{l'}$ with $l \neq l' \in [\![\ell]\!]$. Since $L(x,y) = 0$, we get from (9) that

$$\lim_{k\to\infty} \frac{a_k(C)}{\pi_k(x)} \quad = \quad 0$$

Taking Lemma 12 into account, we get that

$$\forall\ x' \in A_l, \qquad \lim_{k\to\infty} a_k(C)|\Lambda_{\pi_k,C}(x',x')| \quad = \quad 0$$

Since this is true for all the irreducible classes crossed by C, we deduce

$$\lim_{k\to\infty} a_k(C)\Lambda_{\pi_k,C} \quad = \quad 0$$

It follows that

$$L \quad = \quad \lim_{k\to\infty} \sum_{l\in[\![\ell]\!]} \sum_{C\in\mathcal{C}(A_l)} a_k(C)\Lambda_{\pi_k,C}$$

(where $\mathcal{C}(A_l)$ is identified with the set of cycles taking values in A_l, for $l \in [\![\ell]\!]$).

We infer that for any $l \in [\![\ell]\!]$,

$$L_{A_l} \quad = \quad \lim_{k\to\infty} \sum_{C\in\mathcal{C}(A_l)} a_{k,l}(C)\Lambda_{\pi_{k,l},C} \tag{10}$$

with, for any $k \in \mathbb{Z}_+$ and $l \in [\![\ell]\!]$, $\pi_{k,l}$ the probability on A_l defined by

$$\forall\ x \in A_l, \qquad \pi_{k,l}(x) \quad \coloneqq \quad \frac{\pi_k(x)}{\pi_k(A_l)}$$

and

$$\forall\ C \in \mathcal{C}(A_l), \qquad a_{k,l}(C) \quad \coloneqq \quad \frac{a_k(C)}{\pi_k(A_l)}$$

Note that for any given $l \in [\![\ell]\!]$, the sequences $((a_{k,l}(C))_{C\in\mathcal{C}(A_l)})_{k\in\mathbb{Z}_+}$ and $(\pi_{k,l})_{k\in\mathbb{Z}_+}$ are relatively compact (respectively in $\mathbb{R}_+^{\mathcal{C}(A_l)}$ and on the set of probability measures on A_l), so we can extract a subsequence such that both the coefficients and the probability measures converge toward $(\widetilde{a}_l(C))_{C\in\mathcal{C}(A_l)}$ and $\widetilde{\pi}_l$. We deduce from (10) that

$$L_{A_l} \quad = \quad \sum_{C\in\mathcal{C}(A_l)} \widetilde{a}_l(C)\Lambda_{\widetilde{\pi}_l,C}$$

where $\widetilde{a}_l(C) = 0$ if the order of C is strictly larger than n.

Since this is true for all the irreducible classes of L, we get that $\nu(L) \leqslant n$ and finally $L \in \mathcal{N}(V) \cap \mathcal{J}_n(V)$. ■

Note that $\mathcal{N}(V) \cap \mathcal{J}_2(V)$ is the set of Markov generators decomposable into cycles of order 2 and admitting an invariant probability charging all the points of V, which amounts to say they are reversible with respect to a positive probability. Since $S_2 = \mathbb{R}_+$, we recover from Theorem 6 the well-known fact that the eigenvalues of such matrices are real valued.

References

[1] Adam Besenyei. Stochastic matrices and geometry: a gem from Dmitriev and Dynkin. Talk available at `https://abesenyei.web.elte.hu/publications/dynkin.pdf`, May 2014.

[2] N. Dmitriev and E. Dynkin. On the characteristic numbers of a stochastic matrix. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 49:159–162, 1945.

[3] Sophia L. Kalpazidou. *Cycle representations of Markov processes*, volume 28 of *Applications of Mathematics (New York)*. Springer, New York, second edition, 2006. Stochastic Modelling and Applied Probability.

[4] F. I. Karpelevich. On the characteristic roots of matrices with nonnegative elements. *Izv. Akad. Nauk SSSR, Ser. Mat.*, 15:361–383, 1951.

[5] Tosio Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[6] Laurent Miclo. Recuit simulé sans potentiel sur un ensemble fini. In *Séminaire de Probabilités, XXVI*, volume 1526 of *Lecture Notes in Math.*, pages 47–60. Springer, Berlin, 1992.

[7] Laurent Miclo. On finite interweaving relations. Working paper or preprint available at `https://hal.science/hal-04525348`, March 2024.

miclo@math.cnrs.fr

Toulouse School of Economics,
1, Esplanade de l'université,
31080 Toulouse cedex 6, France.
Institut de Mathématiques de Toulouse,
Université Paul Sabatier, 118, route de Narbonne,
31062 Toulouse cedex 9, France.