Speed of convergence to the Perron-Frobenius stationary distribution

Jiawei Huo[(a),*] and Pierre Magal[(b),†]

[(a)] *School of Mathematical Sciences, Beijing Normal University, Beijing 100875, People's Republic of China.*

[(b)] *Univ. Bordeaux, IMB, UMR 5251, F-33400 Talence, France. CNRS, IMB, UMR 5251, F-33400 Talence, France.*

July 2, 2023

Abstract

In this paper, we study the second eigenvalue problem for non-negative matrices. Using the Hilbert projective metric, we give a new simple proof of the distance between the first and second eigenvalue whenever the spectral radius is a single peripheral eigenvalue. This article focuses on the convergence speed of (finite-dimensional) linear dynamical systems to the Perron-Frobenius stationary distribution. We extend these results to primitive matrices and cooperative ordinary differential equations, and our proof also extends to some non-autonomous discrete-time systems.

Keywords: *Non-negative matrices; Hilbert projective metric; Second eigenvalue problem; Convergence speed; Perron-Frobenius stationary distribution.*

1 Introduction

The Perron Frobenius theorem says that the spectral radius $r(A) > 0$ of a non-negative matrix $A \in M_n(\mathbb{R})$ with $A \gg 0$ (i.e., with all the components strictly positive) is a simple dominant eigenvalue of A. This famous theorem gives a result of convergence to the Perron-Frobenius stationary distribution, that is

$$\lim_{m \to \infty} \frac{A^m x}{r(A)^m} = \alpha(x)\, V_R(A),$$

where $x > 0$ and

$$\alpha(x) := \frac{\langle V_L(A), x \rangle}{\langle V_L(A), V_R(A) \rangle} > 0,$$

where $V_R(A) \gg 0$ and $V_L(A) \gg 0$ are respectively right and left eigenvectors of A, associated with $r(A)$.

[*]The research of this author is supported by China Scholarship Council.

[†]Corresponding author. e-mail: pierre.magal@u-bordeaux.fr

It follows that the solution of the discrete time system

$$u(t+1) = Au(t), \forall t \in \mathbb{N}, \text{ and } u(0) = u_0 \in \mathbb{R}^n_+ \setminus \{0\},$$

converge in direction to the line generated by $V_R(A)$. That is,

$$\lim_{t\to\infty} \frac{u_t}{\|u_t\|} = \frac{V_R(A)}{\|V_R(A)\|},$$

where $\|\cdot\|$ is a norm on $\mathbb{R}^n$.

This property can be extended to non-autonomous linear systems of the form (1.4). In the seventies, Golubitsky, Keeler, and Rothschild [10], and Cohen [4] considered the so-called ergodic theorems in demography by using the Hilbert projective metric. The principle is to extend the Perron-Frobenius theorem to a non-autonomous discrete-time linear model. In the eighties, the Hilbert projective metric to obtain global asymptotic stability result for time-dependent sub-linear difference and differential equations, Inaba [16, 17], Tuljapurkar [27, 28] also used Hilbert projective metric and provided new ergodic theorems for population dynamics models.

Let $A \gg 0$ be a strictly positive n by n matrice. The second eigenvalue problem consists in evaluating the distance between the spectral radius of A

$$r(A) = \lim_{m\to\infty} \|A^m\|^{1/m}_{\mathcal{L}(\mathbb{R}^n)},$$

which turns (for finite dimensional space) to be equal to

$$r(A) = \max\{|\lambda| : \lambda \in \sigma(A)\},$$

and the maximum modulus of the remaining eigenvalues is

$$r^\star(A) = \max\{|\lambda| : \lambda \in \sigma(A) \text{ and } \lambda \neq r(A)\}.$$

The second eigenvalue problem consists in evaluating the ratio

$$\frac{r^\star(A)}{r(A)}.$$

While the Perron-Frobenius theorem informs us about the convergence to the stationary distribution, the second eigenvalue problem tells us about the convergence speed to this stationary distribution (see the formula (1.3)).

The main result of this paper about the speed of convergence is the Theorem 4.3, in which we obtain the following estimation

$$r^\star(A) \leq k(A)\, r(A), \tag{1.1}$$

where $k(A) \in [0,1]$ is the Birkhoff contraction ratio, which is defined as

$$k(A) = \sup_{x,y>0: d_H(x,y)\neq 0} \frac{d_H(Ax, Ay)}{d_H(x,y)},$$

where $d_H(x, y)$ is the so called Hilbert projective metric [15] which is defined only when x and y are comparable (see Section 3 for more details). In the special case, where $x \gg 0$ and $y \gg 0$, we can write

$$d_H(x, y) = \ln \left(\max_{1 \leq i \leq n} \frac{x_i}{y_i} \max_{1 \leq j \leq n} \frac{y_j}{x_j} \right). \tag{1.2}$$

In the sixties and seventies, the inequality (1.1) was proved by Zabreiko, Krasnoselskii, and Pokornyi [29]. This problem was reconsidered by Eveson and Nussbaum [9]. We refer to the books of Seneta [25], Lemmens and Nussbaum [21], and Krause [20] for a nice survey of the Hilbert projective metric and their applications. We also refer to Krasnosel'skii and Sobolev [19] for a short survey paper on this subject.

In the present paper, we propose a new proof for inequality (1.1). In Theorem 4.3, we will prove that we can find a constant $\chi \geq 1$, such that for each integer $m \geq 1$ and each $z \in \mathbb{R}^n$,

$$\| \left(\frac{A}{r(A)} \right)^m z - \Pi z \| \leq \chi \, k(A)^m \|z\|, \tag{1.3}$$

where

$$\Pi(x) = \frac{\langle V_L(A), x \rangle \, V_R(A)}{\langle V_L(A), V_R(A) \rangle}, \forall x \in \mathbb{R}^n,$$

and $V_L(A)$ and $V_R(A)$ are respectively left and right positive eigenvectors of A associated with $r(A)$. One can note that the inequality (1.3) is not a direct consequence of (1.1) (since $r(A)$ are obtain as a limit).

The plane of the paper is as follows. In Section 2, we recall the Perron-Frobenius theorem and introduce some notations that will be used in the rest of the paper. In Section 3, we recall some results about the Hilbert projective metric and extend a well-known estimation in Lemma 3.4. Section 4 is devoted to the second eigenvalue problem for non-negative matrices. We first consider the case of strictly positive matrices in Section 4.1. Then we consider the case of the primitive matrices in Section 4.2. Section 4.3 considers an example of three-dimensional Hahn matrices. At the end of Section 4.3 (see Lemma 4.8), we will prove that the limit

$$\lim_{m \to +\infty} k(A^m)^{\frac{1}{m}} \text{ exists,}$$

and based on the numerical simulations (see Figure 2), we will make the following conjecture

$$\lim_{m \to +\infty} k(A^m)^{\frac{1}{m}} = \frac{r^\star(A)}{r(A)}.$$

Section 5 states a result for cooperative systems, and we provide an example of application the speed of convergence to the Perron-Frobenius stationary distribution during the exponential phase of epidemic wave. We conclude the paper with Section 6, which is devoted to non-autonomous discrete-time linear systems. That is,

$$u(t+1) = A_t u(t), \forall t \in \mathbb{N}, \text{ and } u(0) = u_0 \in \mathbb{R}^n_+ \setminus \{0\}. \tag{1.4}$$

In Section 6, we extend the result mentioned above to a non-autonomous system. We obtain a stronger convergence result (in norm of operator) than the recent result obtained by Pituk, and Pötzsche [24] by imposing an additional condition on the left eigenvector of $A_t A_{t-1} \dots A_0$. This extra condition is automatically satisfied for Markovian matrices. Therefore, we conclude the paper by estimating convergence speed for non-autonomous Markovian matrices.

2 Perron-Frobenius theorem

In this section, we introduce some notations and recall the Perron-Frobenius theorem. We refer to Ducrot et al. [8, Chapter 4] for more results and references on this subject with applications in population dynamics.

Definition 2.1. *Let $x, y \in \mathbb{R}^n$ be given. We will use the following notations*

$$
\begin{aligned}
&x \geq y \Leftrightarrow x_i \geq y_i \text{ for all } i = 1, ..., n,\\
&x > y \Leftrightarrow x \geq y \text{ and } x_{i_0} > y_{i_0} \text{ for some } i_0 \in \{1, ..., n\},\\
&x \gg y \Leftrightarrow x_i > y_i \text{ for all } i = 1, ..., n.
\end{aligned}
$$

We can of course extend the above definition to the real matrices.

Definition 2.2. *Let $A \in M_n(\mathbb{R})$ be given. The* ***spectral radius*** *$r(A)$ of A is defined by*

$$r(A) = \lim_{p \to +\infty} \|A^p\|_{\mathcal{L}(\mathbb{R}^n)}^{\frac{1}{p}}.$$

In addition, the ***peripheral spectrum*** *$\sigma_{per}(A)$ of A is defined by*

$$\sigma_{per}(A) = \{\lambda \in \sigma(A) : |\lambda| = r(A)\}.$$

By Jordan normal form of A, $r(A)$ has the following equality

$$r(A) = \sup \{|\lambda| : \lambda \in \sigma(A)\}.$$

In the follows, we introduce the definitions of primitive and irreducible matrices.

Definition 2.3. *Let $A \in M_n(\mathbb{R})$ be a non-negative matrix. We will say that A is* ***primitive*** *if there exists an integer $m \geq 1$ such that*

$$A^m \gg 0.$$

We will say that a matrix $A \in M_n(\mathbb{R})$ is ***irreducible*** *if there exists an integer $m \geq 1$ such that*

$$I + A + A^2 + \cdots + A^m \gg 0.$$

We recall the classical theorem of Perron–Frobenius.

Theorem 2.4. *(Perron–Frobenius) Let $A \in M_n(\mathbb{R})$ be a non-negative and irreducible matrix. Then A satisfies the following properties:*

(i) $r(A) > 0$.

(ii) *$r(A)$ is a simple eigenvalue of A.*

(iii) *There exists $V_R(A) \gg 0$ a right eigenvector of A such that*

$$AV_R(A) = r(A)\, V_R(A).$$

(iv) *There exists $V_L(A) \gg 0$ a left eigenvector of A (i.e. an eigenvector of A^T) such that*

$$A^T V_L(A) = r(A)\, V_L(A) \Leftrightarrow V_L(A)^T A = r(A)\, V_L(A)^T.$$

Moreover, the matrix A is primitive if and only if

$$\sigma_{per}(A) = \{\lambda \in \sigma(A) : |\lambda| = r(A)\} = \{r(A)\}.$$

3 Hilbert projective metric in $\mathbb{R}^n$

Definition 3.1. *If $x > 0$ and $y > 0$ are two elements of $\mathbb{R}^n_+$. Then x and y are* ***comparable*** *if we can find two real numbers $\alpha > 0$ and $\beta > 0$ such that*

$$\alpha y \leq x \leq \beta y.$$

If x and y are comparable, we can define

$$\max(x/y) = \inf\{\beta > 0 : x \leq \beta y\},$$

and

$$\min(x/y) = \sup\{\alpha > 0 : \alpha y \leq x\},$$

and one has

$$\min(y/x) = \frac{1}{\max(x/y)}.$$

We use the above $\max(x/y)$ and $\min(x/y)$ notations, because

$$\max(x/y) = \max_{i=1,\dots,n} \frac{x_i}{y_i}, \text{ and } \min(x/y) = \min_{i=1,\dots,n} \frac{x_i}{y_i}.$$

whenever $x \gg 0$ and $y \gg 0$.

Definition 3.2. *If $x > 0$ and $y > 0$ are two comparable elements of $\mathbb{R}^n_+$, the* ***Hilbert metric*** *(or Hilbert projective metric) is defined by*

$$d_H(x,y) = \ln\left[\frac{\max(x/y)}{\min(x/y)}\right] = \ln\left[\max(y/x)\max(x/y)\right].$$

We can extend the definition by imposing

$$d_H(x,y) = +\infty,$$

whenever x and y are not comparable.

If $x > 0$ and $y > 0$ are two comparable elements of $\mathbb{R}^n_+$, then we always have

$$\max(x/y) \geq \min(x/y),$$

which implies

$$d_H(x,y) \geq 0.$$

In the special case, where $x \gg 0$ and $y \gg 0$, we can write

$$d_H(x,y) = \ln \frac{\max_{1\leq i\leq n} \frac{x_i}{y_i}}{\min_{1\leq j\leq n} \frac{x_j}{y_j}} = \ln\left(\max_{1\leq i\leq n} \frac{x_i}{y_i} \max_{1\leq j\leq n} \frac{y_j}{x_j}\right). \tag{3.1}$$

The following lemma summarizes some properties of the Hilbert metric. It follows from this lemma that the Hilbert metric is a pseudo-metric.

Lemma 3.3. *If $x > 0$, $y > 0$, and $z > 0$ are pairwise comparable elements of $\mathbb{R}^n_+$. The Hilbert metric satisfies the following properties:*

(i) $d_H(x, y) = 0$ *if and only if there exists some* $\lambda > 0$ *such that* $x = \lambda y$.

(ii) $d_H(x, y) = d_H(y, x)$.

(iii) $d_H(x, z) \leq d_H(x, y) + d_H(y, z)$.

(iv) *For any* $\lambda > 0$ *and* $\mu > 0$,
$$d_H(\lambda x, \mu y) = d_H(x, y).$$

(v) *Let* $D \in M_n(\mathbb{R})$ *be a diagonal matrix with strictly positive diagonal elements or a permutation matrix, we have*
$$d_H(Dx, Dy) = d_H(x, y).$$

The following lemma establishes a relationship between the Hilbert metric d_H and the standard metric (induced by the Euclidean norm). The following lemma extends a well known inequality which is true only when $\|x\| = \|y\| = 1$. We refer to the book of Lemmens and Nussbaum [21, inequality (2.21) p.48], or the book of Krause [20, Lemma 2.1.10, p. 24].

Lemma 3.4. *If $x > 0$ and $y > 0$ are two comparable elements of $\mathbb{R}^n_+ \setminus \{0\}$, then*

$$\|x - y\| \leq |\|x\| - e^{-d_H(x,y)}\|y\|| + 2|\|y\| - e^{-d_H(x,y)}\|x\||.$$

If we assume in addition that $\|x\| = \|y\| = M$, we obtain

$$\|x - y\| \leq 3M \left(1 - e^{-d_H(x,y)}\right). \tag{3.2}$$

Proof. Let x and y be two comparable elements of $\mathbb{R}^n_+ \setminus \{0\}$. We can define a, b as follows

$$a = \min(x/y) > 0 \text{ and } b = \min(y/x) > 0,$$

and we deduce

$$d_H(x, y) = -\ln(ab).$$

For simplicity, we use $c = d_H(x, y)$. Then we have

$$ay \leq x \text{ and } bx \leq y.$$

Assuming that the norm is monotone on $\mathbb{R}^n_+$, we obtain

$$\|ay\| = a\|y\| \leq \|x\| \Rightarrow a \leq \frac{\|x\|}{\|y\|},$$

and

$$\|bx\| = b\|x\| \leq \|y\| \Rightarrow b \leq \frac{\|y\|}{\|x\|}.$$

It follows that

$$e^{-c} = a\frac{\|y\|}{\|x\|}b\frac{\|x\|}{\|y\|} \leq a\frac{\|y\|}{\|x\|} \leq 1 \text{ and } e^{c} = (a\frac{\|y\|}{\|x\|}b\frac{\|x\|}{\|y\|})^{-1} \geq \left(b\frac{\|x\|}{\|y\|}\right)^{-1} \geq 1.$$

Then

$$ay \leq x \text{ and } e^{-c}\frac{\|x\|}{\|y\|} \leq a \Rightarrow e^{-c}\frac{\|x\|}{\|y\|}y \leq x, \tag{3.3}$$

$$bx \leq y \text{ and } e^{-c}\frac{\|y\|}{\|x\|} \leq b \Rightarrow e^{-c}\frac{\|y\|}{\|x\|}x \leq y. \tag{3.4}$$

We deduce from the above inequalities that

$$0 \leq x - ye^{-c}\frac{\|x\|}{\|y\|} = (x-y) + y(1 - e^{-c}\frac{\|x\|}{\|y\|}) \leq x(1 - e^{-c}\frac{\|y\|}{\|x\|}) + y(1 - e^{-c}\frac{\|x\|}{\|y\|}).$$

Assuming that the norm is monotone on $\mathbb{R}^n_+$, and using the triangle inequality, we obtain

$$\|(x-y) + y(1 - e^{-c}\frac{\|x\|}{\|y\|})\| \leq \|x\||1 - e^{-c}\frac{\|y\|}{\|x\|}| + \|y\||1 - e^{-c}\frac{\|x\|}{\|y\|}|.$$

By the triangle inequality, we obtain

$$\|x-y\| \leq \|(x-y) + y(1 - e^{-c}\frac{\|x\|}{\|y\|})\| + \|y(1 - e^{-c}\frac{\|x\|}{\|y\|})\| \leq |\|x\| - e^{-c}\|y\|| + 2|\|y\| - e^{-c}\|x\||.$$

□

Definition 3.5. *If $x > 0$ and $y > 0$, and $\max(y/x) < +\infty$, we define the* ***Hopf oscillation***

$$\operatorname{osc}(y/x) = \max(y/x) - \min(y/x). \tag{3.5}$$

In the case that $\max(y/x) = +\infty$, we define

$$\operatorname{osc}(y/x) = +\infty.$$

Definition 3.6. *Let $A \in M_n(\mathbb{R})$ with $A \gg 0$. We define the* ***projective diameter*** *as*

$$\Delta(A) := \sup\left\{d_H(Ax, Ay) : (x,y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+, Ax \neq 0 \text{ and } Ay \neq 0\right\},$$

the ***Birkhoff contraction ratio*** *as*

$$k(A) := \inf\{\lambda > 0 : d_H(Ax, Ay) \leq \lambda\, d_H(x,y), \forall x, y \in \mathbb{R}^n_+ \setminus \{0\}\},$$

and ***Hopf oscillation ratio*** *as*

$$N(A) := \inf\{\gamma > 0 : \operatorname{osc}(Ax, Ay) \leq \gamma \operatorname{osc}(y,x), \forall x, y \in \mathbb{R}^n_+ \setminus \{0\}\}.$$

The Birkhoff contraction ratio can be equivalently defined as

$$k(A) = \sup_{x,y>0: d_H(x,y)\neq 0} \frac{d_H(Ax, Ay)}{d_H(x,y)}. \tag{3.6}$$

The following result combines Birkhoff's [1, 2], and Hopf's [12] results. A short proof of this theorem was obtained by Eveson and Nussbaum [9].

Theorem 3.7 (**Birkhoff-Hopf**). *Suppose that $A \in M_n(\mathbb{R})$ with $A \gg 0$. It follows that*

$$k(A) = N(A) = \tanh\left[\frac{\Delta(A)}{4}\right].$$

By Theorem 3.7 we have

$$k(A) = \tanh\left[\frac{\Delta(A)}{4}\right] = \frac{e^{\frac{\Delta(A)}{4}} - e^{-\frac{\Delta(A)}{4}}}{e^{\frac{\Delta(A)}{4}} + e^{-\frac{\Delta(A)}{4}}} = \frac{1-\sqrt{e^{-\Delta(A)}}}{1+\sqrt{e^{-\Delta(A)}}},$$

therefore, Theorem 3.7 implies that

$$k(A) < 1.$$

The following result gives an explicit formula for $k(A)$. This result was proved by Seneta [25, Theorem 3.12, p. 108]. The arguments leading to the two inequalities which comprise the proof of this theorem are due to Ostrowski [23] and Brushell [3].

Theorem 3.8 (**Seneta**). *Let $A = (a_{ij}) \in M_n(\mathbb{R})$, and assume that $A \gg 0$. Then the Birkhoff contraction ratio is strictly less than 1, and explicitly given by*

$$k(A) = \frac{1-\sqrt{\Phi(A)}}{1+\sqrt{\Phi(A)}} < 1, \tag{3.7}$$

with

$$\Phi(A) = \min_{1\le i,j,k,l\le n} \frac{a_{ik}a_{jl}}{a_{jk}a_{il}}. \tag{3.8}$$

By comparing Theorems 3.7 and 3.8, we have

$$\Phi(A) = e^{-\Delta(A)}$$

whenever $A \gg 0$.

Lemma 3.9. *Let $A, B \in M_n(\mathbb{R})$, and assume that $A \gg 0$ and $B \gg 0$. Then we have the following properties*

(i)

$$d_H(Ax, Ay) \le k(A)\, d_H(x, y), \forall x > 0, \forall y > 0. \tag{3.9}$$

(ii)

$$k(AB) \le k(A)\, k(B). \tag{3.10}$$

(iii) *For each integer $p \ge 1$, $q \ge 1$, and $m = p + q$,*

$$k(A^m) \le k(A^p)\, k(A^q) \le k(A)^m. \tag{3.11}$$

Proof. Proof of (ii). Let $x > 0$ and $y > 0$. Inequality (3.9) is direct consequence of the definition (3.6) of Birkhoff contraction ratio, and we have

$$d_H(A\,Bx, ABy) \le k(A)\, d_H(Bx, By) \le k(A)\, k(B)\, d_H(x, y),$$

and (3.10) follows. $\square$

In the following lemma, we give an estimate of the supper bound of $k(A)$.

Lemma 3.10. *Let $A = (a_{ij}) \in M_n(\mathbb{R})$, and assume that $A \gg 0$. Then*

$$k(A) \leq \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}},$$

where

$$A_{\max} = \max_{i,j=1,\ldots,n} a_{ij}, \text{ and } A_{\min} = \min_{i,j=1,\ldots,n} a_{ij}.$$

Proof. By Theorem 3.8, we have

$$\Phi(A) = \min_{1 \leq i,j,k,l \leq n} \frac{a_{ik} a_{jl}}{a_{jk} a_{il}} \geq \frac{A_{\min}^2}{A_{\max}^2},$$

which implies

$$\sqrt{\Phi(A)} \geq \sqrt{\frac{A_{\min}^2}{A_{\max}^2}} = \frac{A_{\min}}{A_{\max}},$$

and we obtain

$$k(A) = \frac{1 - \sqrt{\Phi(A)}}{1 + \sqrt{\Phi(A)}} \leq \frac{1 - \dfrac{A_{\min}}{A_{\max}}}{1 + \dfrac{A_{\min}}{A_{\max}}} = \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}}.$$

□

4 The second eigenvalue problem for non negative matrices

4.1 Main result

Let $A = (a_{ij}) \in M_n(\mathbb{R})$, and assume that $A \gg 0$.

Definition 4.1. *We define the **second spectral radius** of A as*

$$r^\star(A) := \sup \{|\lambda| : \sigma(A) \setminus \{r(A)\}\}.$$

Definition 4.2. *Let $A = (a_{ij}) \in M_n(\mathbb{R})$, and assume that $A \gg 0$. Let $\Pi \in \mathcal{L}(\mathbb{R}^n)$, the linear map defined by*

$$\Pi(x) = \frac{\langle V_L(A), x \rangle V_R(A)}{\langle V_L(A), V_R(A) \rangle}, \forall x \in \mathbb{R}^n. \tag{4.1}$$

Then Π a projector on the eigenspace associated with $r(A)$, which is the unique projector satisfying

$$\Pi A = A \Pi = r(A) \Pi.$$

Define

$$B = A(I - \Pi) = A - r(A)\Pi,$$

then we observe that

$$r^\star(A) = \lim_{p\to+\infty} \|B^p\|_{\mathcal{L}(\mathbb{R}^n)}^{\frac{1}{p}} = r(B),$$

and by the Perron-Frobenius theorem we have

$$\sigma(B) = (\sigma(A) \setminus \{r(A)\}) \cup \{0\}.$$

The main result of this article is the following theorem.

Theorem 4.3 (**Speed of convergence**). *Let $A = (a_{ij}) \in M_n(\mathbb{R})$, and assume that $A \gg 0$. Then*

$$r^\star(A) \le k(A)\, r(A), \tag{4.2}$$

and we can find a constant $\chi \ge 1$, such that for each integer $m \ge 1$ and each $z \in \mathbb{R}^n$,

$$\| \left(\frac{A}{r(A)}\right)^m z - \Pi z\| \le \chi\, k(A)^m \|z\|, \tag{4.3}$$

where $0 \le k(A) < 1$ is the Birkhoff contraction ratio.

Proof. Replacing A by $A/r(A)$, we can always assume that $r(A) = 1$. Now, by the Perron-Frobenius theorem, we can decompose A into

$$A = \Pi + B,$$

where $\Pi \gg 0$ is a projector matrix ($\Pi^2 = \Pi$ and $\Pi A = A\Pi = \Pi$) associated with the linear operator defined in (4.1), and since $r(A) = 1$ it follows that $r(B) < 1$. Moreover

$$\Pi\, B = B\, \Pi = 0,$$

hence

$$A^n = \Pi + B^n, \forall n \in \mathbb{N}.$$

Since $V_L(A) \gg 0$, the map

$$|||x||| = \langle V_L(A), |x| \rangle,$$

defines a norm on $\mathbb{R}^n$ which satisfies

$$|||A^n x||| = |||x|||, \forall n \ge 1, \text{ and } \forall x \in \mathbb{R}^n_+,$$

which follows from the definition of the left eigenvector (i.e., from the fact $V_L(A)^T A = r(A) V_L(A)^T$), and since by assumption $r(A) = 1$.

We can choose $x \gg 0$ large enough (with $\|x\| = M$) so that

$$x + B_{\mathbb{R}^n}(0,1) \subset (0, +\infty)^n.$$

We deduce that by choosing $z \in \mathbb{R}^n$ with $|||z||| = \rho > 0$ (small enough so that $|||(I-\Pi)|||_{\mathcal{L}(\mathbb{R}^n)}\rho \le 1$), and since $(I-\Pi)z \in B_{\mathbb{R}^n}(0,1)$, we deduce that

$$y = x + (I - \Pi)z \gg 0,$$

and since $y \ge 0$ we obtain

$$|||y||| = |||A^n y|||, \forall n \ge 1,$$

and

$$|||y||| = |||x|||.$$

Dropping the three bars norm notation, by Lemma 3.4, and since $\Pi(x-y) = 0$,

$$\|B^n x - B^n y\| = \|A^n x - A^n y\| \leq 3M(1 - e^{-d_H(A^n x, A^n y)}).$$

and since $1 - e^{-x} \leq x, \forall x \geq 0$, we obtain

$$\|B^n x - B^n y\| = \|A^n x - A^n y\| \leq 3\,M\,d_H(A^n x, A^n y).$$

By Birkhoff theorem and (3.11), we have

$$\|B^n x - B^n y\| = \|A^n x - A^n y\| \leq 3\,M\,k(A^n) d_H(x,y) \leq 3\,M\,k(A)^n d_H(x,y).$$

Next

$$\|B^n z\| = \|B^n(I - \Pi + \Pi)z\| = \|B^n(I - \Pi)z\| = \|A^n x - A^n y\| \leq 3M\,k(A)^n d_H(x,y),$$

and by taking the supremum over all $z \in \mathbb{R}^n$ with $\|z\| = \rho > 0$, we obtain

$$\|B^n\|_{\mathcal{L}(\mathbb{R}^n)} \leq \frac{3M}{\rho}\,k(A)^n \sup_{z \in \mathbb{R}^n : \|z\| \leq 1} d_H(x, x+z),$$

and (4.3) follows.

Moreover we deduce that

$$\frac{\ln\left(\|B^n\|_{\mathcal{L}(\mathbb{R}^n)}\right)}{n} \leq \frac{\ln\left(\dfrac{3M}{\rho}\right)}{n} + \ln\left(k(A)\right) + \frac{\ln\left(\sup_{z \in B_{\mathbb{R}^n}(0,1)} d_H(x, x+z)\right)}{n}.$$

We deduce that

$$\ln\left(r(B)\right) = \lim_{n \to \infty} \frac{\ln\left(\|B^n\|_{\mathcal{L}(\mathbb{R}^n)}\right)}{n} \leq \ln\left(k(A)\right),$$

the proof is completed. □

Example 4.4. *The above result could be optimal since for the example*

$$A_0 = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}.$$

Then the spectral radius of A_0 is 1, and since A_0 is a projector, and the second eigenvalue is 0. Moreover, by Lemma 3.10, we have

$$k(A_0) = 0.$$

Therefore, in this example, Theorem 4.3 gives an exact estimation of the distance to the second eigenvalue.

Example 4.5. *In the follows, we give a numerical comparison of second eigenvalue and Birkhoff contraction of the strict positive matrices.*

We consider the following 2-dimensional matrix A_1

$$A_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

By using the function eig *of MATLAB to compute the eigenvalues of* A_1*, we obtain*

$$r^\star(A_1) = 0.37, r(A_1) = 5.37, \text{ and } k(A_1) = 0.1.$$

Then we have

$$r^\star(A_1) \leq k(A_1)\, r(A_1) = 0.54.$$

This example shows that our estimate of the second eigenvalue of matrix A_1 *is pretty good compared to the spectral radius of* A_1*.*

Example 4.6. *We consider the following 3-dimensional matrix* A_2

$$A_2 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

By using the function eig *of MATLAB to compute the eigenvalues of* A_2*, we obtain*

$$r^\star(A_2) = 1.12, r(A_2) = 16.12, \text{ and } k(A_2) = 0.21.$$

Then we have

$$r^\star(A_2) \leq k(A_2)\, r(A_2) = 3.36.$$

4.2 The case of primitive matrices

In order to extend Theorem 4.3 to non-negative irreducible matrix $A \geq 0$, one must remember that

$$\sigma_{per}(A) = \{\lambda \in \sigma(A) : |\lambda| = r(A)\} = \{r(A)\}$$

if and only if A is primitive. Therefore it makes sense to consider now only the case of primitive matrices.

It is well known that a non-negative matrix $A \geq 0$ is primitive if and only if

$$A^{n^2} \gg 0,$$

where n is the dimension of A. We refer for example to the book of Horn and Johnson [13, Corollary 8.6.9, p. 520] for a proof of this result.

Now since

$$\sigma(A^{n^2}) = \left\{\lambda^{n^2} : \lambda \in \sigma(A)\right\},$$

by applying Theorem 4.3 to A^{n^2}, we obtain the following corollary.

Corollary 4.7. *Let $A = (a_{ij}) \in M_n(\mathbb{R})$. Assume that $A \geq 0$ and A is primitive. Then*

$$r^\star(A) \leq k(A^{n^2})^{\frac{1}{n^2}}\, r(A),$$

and we can find a constant $\chi \geq 1$, such that for each integer $m \geq n^2$ and each $z \in \mathbb{R}^n$,

$$\| \left(\frac{A}{r(A)} \right)^m z - \Pi z \| \leq \chi\, k(A^{n^2})^{\frac{m}{n^2}}\, \|z\|,$$

where $0 \leq k(A^{n^2}) < 1$ is the Birkhoff contraction ratio of A^{n^2}.

4.3 Numerical illustration for Hahn matrix model

The Hahn matrix model was used to describe cell metabolism. This model was introduced in the early sixties and developed in the seventies by Hahn [11] to model cell metabolism. Such matrices have a particular form that allows computing their full spectrum (see Appendix A). This property was observed simultaneously by Demongeot [6, 7], and Thames and White [26]. This class of matrices called *quasi-circulant matrices* have the following form

$$A = \alpha_0 I + \alpha_1 P + \ldots + \alpha_m P^m,$$

where $\alpha_0, \alpha_1, \ldots, \alpha_m \in \mathbb{R}$, and

$$P = \begin{bmatrix} 0 & 0 & \cdots & 0 & p_n \\ p_1 & 0 & \cdots & 0 & 0 \\ 0 & p_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & p_{n-1} & 0 \end{bmatrix},$$

where $p_1, \ldots, p_n \in (0, +\infty)$.

Let us consider the 3 by 3 example of quasi-circulant matrix

$$N(n+1) = AN(n), \forall n \geq 0, N(0) = N_0 \in \mathbb{R}^3_+, \tag{4.4}$$

where

$$A = \begin{bmatrix} \alpha_0 & 2D\alpha_2 & 2D\alpha_1 \\ \alpha_1 & \alpha_0 & 2D\alpha_2 \\ \alpha_2 & \alpha_1 & \alpha_0 \end{bmatrix}. \tag{4.5}$$

Consider the matrix

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

Circulant matrices use a permutation P only. That is $2D = 1$ in the above example (see the book of Davis [5] for more results). Then it is straightforward

$$P^2 = \begin{bmatrix} 0 & 2D & 0 \\ 0 & 0 & 2D \\ 1 & 0 & 0 \end{bmatrix},$$

and we deduce

$$A = \alpha_0 I + \alpha_1 \begin{bmatrix} 0 & 0 & 2D \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 2D & 0 \\ 0 & 0 & 2D \\ 1 & 0 & 0 \end{bmatrix} = \alpha_0 I + \alpha_1 P + \alpha_2 P^2.$$

Let $\lambda_{P,1}, \lambda_{P,2}, \lambda_{P,3}$ be the eigenvalues of matrix P. By Lemma A.1, it follows that

$$\lambda^3 = 2D.$$

and

$$\lambda_{P,1} = \sqrt[3]{2D},$$

$$\lambda_{P,2} = \sqrt[3]{2D}\, e^{i\dfrac{2\pi}{3}} = \sqrt[3]{2D}\, e^{i\left(\pi - \dfrac{\pi}{3}\right)},$$

and

$$\lambda_{P,3} = \sqrt[3]{2D} e^{i\dfrac{4\pi}{3}} = \sqrt[3]{2D} e^{i\left(\pi + \dfrac{\pi}{3}\right)},$$

so we obtain

$$\lambda_{P,1} = \sqrt[3]{2D}, \lambda_{P,2} = \sqrt[3]{2D}\left[-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right], \text{ and } \lambda_{P,2} = \sqrt[3]{2D}\left[-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right].$$

Moreover, by Lemma A.2, the eigenvalues of the matrix A are given from the eigenvalues of P by

$$\lambda_{A,i} = \alpha_0 + \alpha_1 \lambda_{P,i} + \alpha_2 \lambda_{P,i}^2.$$

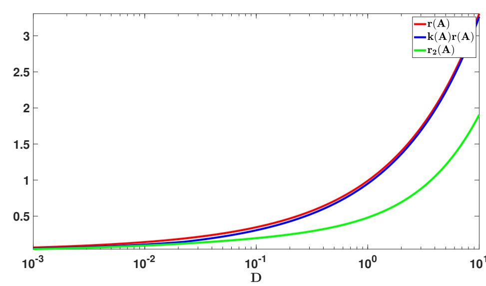

Figure 1: *In this figure we plot $r(A)$, $k(A)\, r(A)$, and $r^\star(A)$ in function of D. We take $\alpha_0 = 0.01$ (very small), $\alpha_1 = 0.4$ and $\alpha_2 = 0.3$. We observe that $k(A)$ is close to 1.*

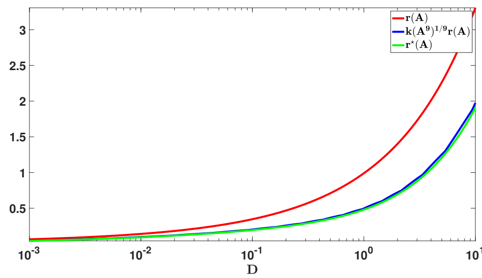

Figure 2: *In this figure we plot $r(A)$, $k(A)\,r(A)$, and $r^\star(A)$ in function of D. We take $\alpha_0 = 0.01$ (very small), $\alpha_1 = 0.4$ and $\alpha_2 = 0.3$. We observe that $k(A^9)^{1/9}$ (where $9 = 3^2$ corresponds to the Corollary 4.7) is close to* $\dfrac{r^\star(A)}{r(A)}$.

Lemma 4.8. *Let $A = (a_{ij}) \in M_n(\mathbb{R})$. Assume that $A \geq 0$ and A is primitive. Then the limit*

$$\lim_{m \to +\infty} k(A^m)^{\frac{1}{m}} \text{ exists,}$$

and satisfies

$$\lim_{m \to +\infty} k(A^m)^{\frac{1}{m}} = e^{\inf_{m \geq 0} \frac{\ln(k(A^m))}{m}}.$$

The proof of this result given below is taken from the book of Kato [18, pp. 27–28].

Proof. Define $a_p := \ln(k(A^p))$ for $p \geq 0$. Here $A^0 := I$, which implies that $\|A^0\| = 1$, so that $a_0 = 0$. Let us prove that

$$\lim_{p \to +\infty} \frac{a_p}{p} = \inf_{m > 0} \frac{a_m}{m}.$$

Since we have

$$k\left(A^{p+q}\right) \leq k(A^p)\, k(A^q),\ \forall p, q \in \mathbb{N},$$

we deduce that the sequence $(a_p)_{p \geq 0}$ is sub-additive, namely it satisfies

$$a_{p+q} \leq a_p + a_q,\ \forall p, q \in \mathbb{N},$$

and the result follows from the classical argument from the book of Kato [18, pp. 27–28]. □

Conjecture: The above numerical simulations suggest the following property

$$\lim_{m \to +\infty} k(A^m)^{\frac{1}{m}} = \frac{r^\star(A)}{r(A)}.$$

5 Application to cooperative system of ordinary differential equations

Let $A = (a_{ij}) \in M_n(\mathbb{R})$, and assume that the off-diagonal elements of A are non-negative, and $A + \delta I$ is non-negative irreducible whenever $\delta > 0$ is large enough. Then

$$e^{At} \gg 0,\ \forall t > 0.$$

Recall that $s(A)$ the **spectral bound of** A is defined by

$$\mathrm{s}(A) = \sup \{\mathrm{Re}\ \lambda : \lambda \in \sigma(A)\}.$$

Then we can find $V_R(A) \gg 0$ a Right eigenvector of A such that

$$AV_R(A) = \mathrm{s}(A)\, V_R(A),$$

and we can find $V_L(A) \gg 0$ a Left eigenvector of A (i.e. an eigenvector of A^T) such that

$$A^T V_L(A) = \mathrm{s}(A)\, V_L(A) \Leftrightarrow V_L(A)^T A = \mathrm{s}(A)\, V_L(A)^T.$$

We define the projector

$$\Pi(x) = \frac{\langle V_L(A), x\rangle\, V_R(A)}{\langle V_L(A), V_R(A)\rangle}, \ \forall x \in \mathbb{R}^n.$$

Then Π a projector on the eigenspace associated with $\mathrm{s}(A)$, which is the unique projector satisfying

$$\Pi\, A = A\, \Pi = \mathrm{s}(A)\, \Pi.$$

Definition 5.1. *We define the* ***second spectral bound***

$$\mathrm{s}^\star(A) = \sup\{\mathrm{Re}\ \lambda : \sigma(A) \setminus \{\mathrm{s}(A)\}\},$$

and the ***Birkhoff contraction bound*** *as*

$$\beta(A) = \begin{cases} \ln\left(k\left(e^A\right)\right) < 0, & \text{if } k\left(e^A\right) > 0, \\ -\infty, & \text{if } k\left(e^A\right) = 0. \end{cases}$$

Now by applying Theorem 4.3 to e^A we obtain that following result.

Corollary 5.2. *Let $A = (a_{ij}) \in M_n(\mathbb{R})$, and assume that the off-diagonal elements of A are non-negative, and $A + \delta I$ is non-negative irreducible whenever $\delta > 0$ is large enough. Then*

$$\mathrm{s}^\star(A) \leq \beta(A) + \mathrm{s}(A) < \mathrm{s}(A),$$

and we can find a constant $\chi \geq 1$, such that for each real number $t \geq 0$ and each $z \in \mathbb{R}^n$,

$$\|e^{(A - s(A)I)t}\, z - \Pi\, z\| \leq \chi\, e^{\beta(A)\, t}\, \|z\|.$$

Proof. Due to the assumption that $A + \delta I$ is non-negative irreducible whenever $\delta > 0$ is large enough, we deduce that $e^{At} \gg 0$ (that is all the components of e^{At} are strictly positive) for each $t > 0$. By applying Theorem 4.3 to $B = e^A$, we deduce that

$$r^\star(e^A) \leq k(e^A) r(e^A), \tag{5.1}$$

and we deduce that

$$\begin{aligned} \mathrm{s}^\star(A) &= \mathrm{s}(A\,(I - \Pi)) = \ln(r(e^A\,(I - \Pi))) \\ &= \ln(r^\star(e^A)) \\ &\leq \ln(k(e^A)) + \ln(r(e^A)) \\ &= \beta(A) + \mathrm{s}(A). \end{aligned} \tag{5.2}$$

By applying the second part of Theorem 4.3 to $B = e^A$, we deduce that we can find a constant $\chi \geq 1$, such that for each integer $m \geq 1$ and each $z \in \mathbb{R}^n$,

$$\|e^{(A-s(A)I)m} z - \Pi z\| \quad = \| \left(\frac{e^A}{r(e^A)} \right)^m z - \Pi z\| \leq \chi \, k(e^A)^m \|z\| = \chi \, e^{\beta(A)\, m} \, \|z\|.$$

By definition of Π, we have

$$\Pi e^{(A-s(A)I)t} z = e^{(A-s(A)I)t} \Pi z = \Pi z, \forall t \geq 0.$$

Therefore, for $t \in [m, m+1]$ (for some positive integer $m \in \mathbb{N}$) we obtain

$$\begin{aligned} \|e^{(A-s(A)I)t} z - \Pi z\| \quad &= \|e^{(A-s(A)I)m} e^{(A-s(A)I)(t-m)} z - \Pi e^{(A-s(A)I)(t-m)} z\| \\ &\leq \chi \, e^{\beta(A)\, m} \, \|e^{(A-s(A)I)(t-m)} z\|, \end{aligned}$$

hence

$$\|e^{(A-s(A)I)t} z - \Pi z\| \leq \widehat{\chi} \, e^{\beta(A)\, t} \, \|z\|,$$

where

$$\widehat{\chi} = \chi \sup_{s \in [0,1]} e^{-\beta(A)\, s} \sup_{\delta \in [0,1]} \sup_{\|z\|=1} \|e^{(A-s(A)I)\delta} z\|,$$

and the proof is completed. □

At the early stage of the epidemic of COVID-19 in China (see Liu et al. [22]), it was reasonable to assume that the number of infectious and unreported satisfy the following system for $t \geq t_0$,

$$\begin{cases} I'(t) = \tau_0 S_0 [I(t) + U(t)] - \nu I(t), \\ U'(t) = \nu(1-f) I(t) - \eta U(t). \end{cases}$$

This system is supplemented by initial data

$$I(t_0) = I_0 > 0, \text{ and } U(t_0) = U_0 \geq 0.$$

Here $t \geq t_0$ is time in days, t_0 is the beginning date of the epidemic, $I(t)$ is the number of asymptomatic infectious individuals at time t, $U(t)$ is the number of unreported symptomatic infectious individuals (i.e. symptomatic infectious with mild symptoms) at time t. In addition, S_0 denotes the number of susceptible at time t_0, τ_0 denotes transmission rate, f denotes the fraction of asymptomatic infectious that become reported symptomatic, $1/\eta$ denotes the average time symptomatic infectious have symptoms, and $1/\nu$ denotes the average time during which asymptomatic infectious are asymptomatic.

As in [22, See caption of Figure 3], we use

$$f = 0.8, \, \eta = 1/7, \, \nu = 1/7, \text{ and } S_0 = 11.081 \times 10^6. \tag{5.3}$$

and

$$\tau_0 = 4.44 \times 10^{-8}. \tag{5.4}$$

the value estimated from the data in [22, See the caption of Figure 3].

So we consider the matrix

$$A = \begin{bmatrix} \tau_0 S_0 - \nu & \tau_0 S_0 \\ \nu(1-f) & -\eta \end{bmatrix}.$$

In Figures 2 and 3 we plot

$$s(A) = \ln\left(r(e^A)\right),$$

$$\beta(A) + s(A) = \ln\left(k\left(e^A\right)\right) + s(A),$$

and

$$s^\star(A) := \sup\{\mathrm{Re}\ \lambda : \sigma(A) \setminus \{\mathrm{s}(A)\}\},$$

in function of τ_0.

In these figures, we first compute e^A by using the MATLAB function expm(A) which is based on the algorythm proposed by Higham [14].

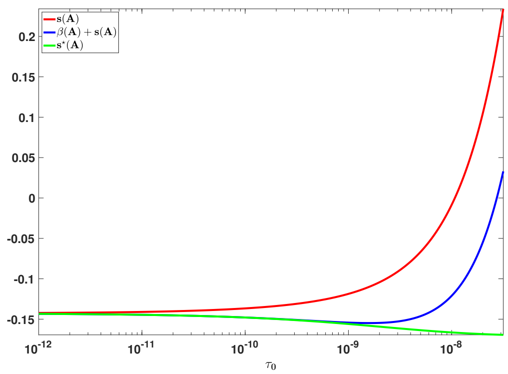

Figure 3: *In this figure we plot $s(A)$, $\beta(A) + s(A)$, and $s^\star(A)$ in function of τ_0. We vary τ_0 from 10^{-12} to 10^{-7} and we use the parameters values in (5.3) and (5.4). Here, we use a log scale for the horizontal axis.*

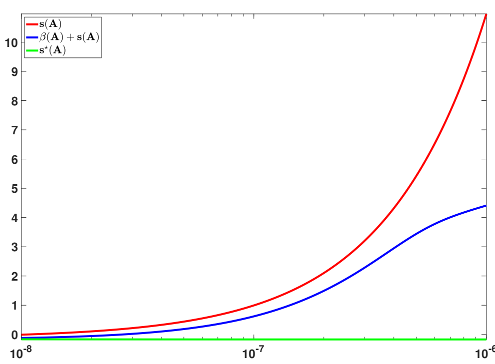

Figure 4: *In this figure we plot $s(A)$, $\beta(A) + s(A)$, and $s^\star(A)$ in function of τ_0. We vary τ_0 from 10^{-8} to 10^{-6} and we use the parameters values in (5.3) and (5.4). Here, we use a log scale for the horizontal axis.*

In order to visualize the speed of convergence to the Perron-Frobenius stationary distribution, we define

$$I_{s(A)}(t) = e^{-s(A)(t-t_0)} I(t), \text{ and } U_{s(A)}(t) = e^{-s(A)(t-t_0)} U(t).$$

Then $I_{s(A)}(t)$ and $U_{s(A)}(t)$ satisfy the following system

$$\begin{pmatrix} I'_{s(A)}(t) \\ U'_{s(A)}(t) \end{pmatrix} = (A - s(A)I) \begin{pmatrix} I_{s(A)}(t) \\ U_{s(A)}(t) \end{pmatrix} \tag{5.5}$$

with the initial value

$$I_{s(A)}(t_0) = I_0 > 0, \text{ and } U_{s(A)}(t_0) = U_0 \geq 0.$$

and

$$I_0 + U_0 = 1.$$

From Figures 5 and 6, we observe that changing the transmission τ from 4×10^{-9} to 4×10^{-8} (by one order to magnitude) will have a very significant effect on the speed of convergence to the Perron-Frobenius stationary distribution. We also observe that the relative proportion of unreported individual is much smaller for 4×10^{-8} than from 4×10^{-9}.

In Figure 5, we use $\tau_0 = 4 \times 10^{-9}$, and we obtain a period of 40-50 days for the solutions to converge to the Perron-Frobenius stationary distribution.

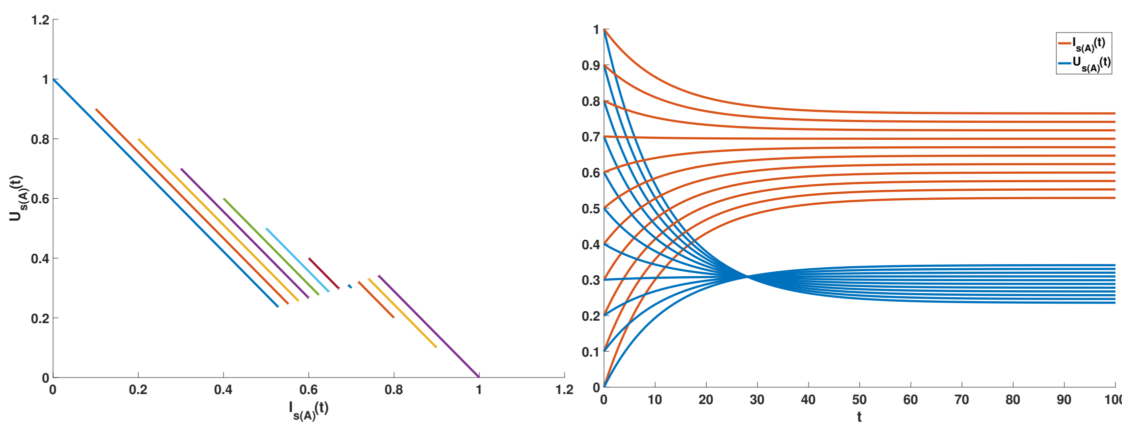

Figure 5: *In this figure, we plot $I_{s(A)}(t)$ and $U_{s(A)}(t)$ in the phase plane (left) and as a function of time (right). In the figure, we use $\tau_0 = 4 \times 10^{-9}$ and we use the parameters values in (5.3) and (5.4). Here $s(A) = 0.16$, and $s^\star(A) = 0.08$, and $s^\star(A) - s(A) = -0.08$ gives us the speed of convergence to the Perron-Frobenius stationary distribution.*

In Figure 6, we use $\tau_0 = 4 \times 10^{-8}$, and we obtain a period of 5-10 days for the solutions to converge to the Perron-Frobenius stationary distribution.

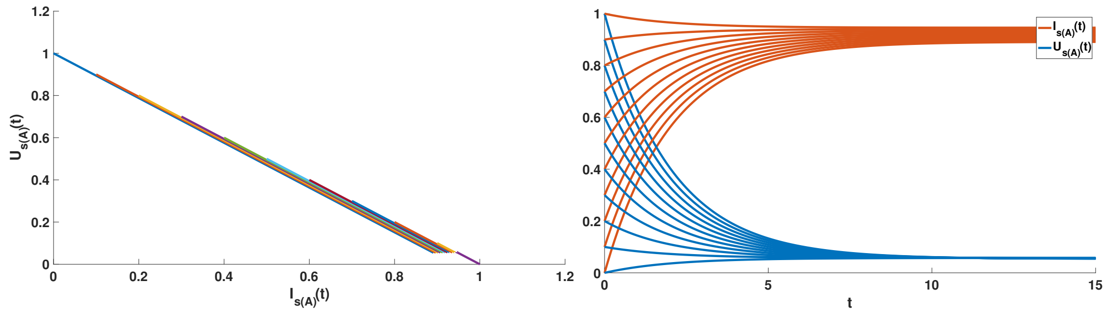

Figure 6: *In this figure, we plot $I_{s(A)}(t)$ and $U_{s(A)}(t)$ in the phase plane (left) and as a function of time (right). We use $\tau_0 = 4 \times 10^{-8}$ and we use the parameters values in (5.3) and (5.4). Here $s(A) = 0.33$, and $s^\star(A) = 0.17$, and $s^\star(A) - s(A) = -0.16$ gives us the speed of convergence to the Perron-Frobenius stationary distribution.*

6 Non-autonomous discrete time system

We consider

$$u(t+1) = A_t u(t), \forall t \in \mathbb{N}, \text{ with } u(0) = u_0 \in \mathbb{R}^n_+ \setminus \{0\}, \tag{6.1}$$

where $A_t \in M_n(\mathbb{R})$ for each $t \in \mathbb{N}$.

Assumption 6.1. *We assume that there exist two non-negative matrices A_- and A_+ (with all their components strictly positive) such that*

$$0 \ll A_- \leq A_t \leq A_+, \forall t \in \mathbb{N}.$$

Lemma 6.2. *Let Assumption 6.1 be satisfied. Then*

$$r(A_-) \leq r(A_t) \leq r(A_+).$$

Proof. Due to $0 \ll A_- \leq A_t \leq A_+, \forall t \in \mathbb{N}$, then we have

$$A_- V_R(A_t) \leq A_t V_R(A_t) = r(A_t) V_R(A_t).$$

Moreover, we apply $V_L(A_-)^T$ on the above inequality and then obtain

$$r(A_-) V_L(A_-)^T V_R(A_t) = V_L(A_-)^T A_- V_R(A_t) \leq V_L(A_-)^T A_t V_R(A_t) = r(A_t) V_L(A_-)^T V_R(A_t).$$

Therefore, we deduce that $r(A_-) \leq r(A_t)$. The proof for $r(A_t) \leq r(A_+)$ is similar. □

Lemma 6.3. *Let Assumption 6.1 be satisfied. Then there exist two real numbers $\alpha > 0$ and $\beta > 0$ (satisfying $\beta \geq 1 \geq \alpha > 0$) such that*

$$\alpha \mathbb{1} \leq V_R(A_t) \leq \beta \mathbb{1}, \quad \forall t \in \mathbb{Z}, \tag{6.2}$$

where $V_R(A_t)$ is the normalized right eigenvector of A_t (i.e. $\|V_R(A_t)\| = 1$) and $\mathbb{1} = (1, 1, ..., 1)^T$.

Proof. Since all the norms are equivalent in $\mathbb{R}^n$ we can find a constant $\gamma_+ > \gamma_- > 0$ such that

$$\gamma_- \|x\|_\infty \leq \|x\| \leq \gamma_+ \|x\|_\infty, \forall x \in \mathbb{R}^n,$$

where $\|\cdot\|_\infty$ is the supremum norm $\|x\|_\infty = \max_{i=1,\dots,n} x_i$. Since $\|V_R(A_t)\| = 1$, we have

$$\|V_R(A_t)\|_\infty = \max_{i=1,\dots,n} V_R(A_t)_i \geq \gamma_+^{-1}.$$

So we can find $i_0 \in \{1, \dots, n\}$, satisfying

$$V_R(A_t)_{i_0} \geq \gamma_+^{-1}.$$

Next, we have

$$A_t V_R(A_t) = r(A_t) V_R(A_t), \quad \forall t \in \mathbb{Z},$$

hence

$$V_R(A_t) = \frac{A_t}{r(A_t)} V_R(A_t) \geq \frac{A_-}{r(A_+)} V_R(A_t) \geq V_R(A_t)_{i_0} \frac{\min A_-}{r(A_+)} \mathbb{1} \geq \gamma_+^{-1} \frac{\min A_-}{r(A_+)} \mathbb{1},$$

where $\min A_- = \min\limits_{1 \leq i,j \leq n} (A_-)_{ij}$.

To prove the second inequality of (6.2), it is sufficient to observe that

$$V_R(A_t) \leq \|V_R(A_t)\|_\infty \mathbb{1} \leq \gamma_-^{-1} \|V_R(A_t)\| \mathbb{1},$$

and the proof is completed. □

By applying the Lemma 6.3 to A_t^T, we obtain the following lemma.

Lemma 6.4. *Let Assumption 6.1 be satisfied. Then there exist two real numbers $\alpha > 0$ and $\beta > 0$ (satisfying $\beta \geq 1 \geq \alpha > 0$) such that*

$$\gamma \mathbb{1} \leq V_L(A_t) \leq \beta \mathbb{1}, \quad \forall t \in \mathbb{Z}, \tag{6.3}$$

where $V_L(A_t)$ is the normalized left eigenvector of A_t (i.e. $\|V_L(A_t)\| = 1$) and $\mathbb{1} = (1, 1, ..., 1)^T$.

We define

$$U(t,s) = A_{t-1} \dots A_s, \text{ if } t > s, \text{ and } U(s,s) = I.$$

Thus, we have the semigroup property for each $t \geq r \geq s$,

$$U(t,r)U(r,s) = (A_{t-1} \dots A_r)(A_{r-1} \dots A_s) = U(t,s).$$

For each $t > s$, we define

$$\Pi(t,s)x = \frac{\langle V_L(U(t,s)), x\rangle V_R(U(t,s))}{\langle V_L(U(t,s)), V_R(U(t,s))\rangle} \tag{6.4}$$

where $V_L(U(t,s)) \gg 0$ and $V_R(U(t,s)) \gg 0$ are respectively a left and a right eigenvectors of $U(t,s)$ associated with $r(U(t,s))$ (the spectral radius of $U(t,s)$). This means that

$$V_L(U(t,s))^T U(t,s) = r(U(t,s)) V_L(U(t,s))^T \text{ and } U(t,s) V_R(U(t,s)) = r(U(t,s)) V_R(U(t,s)),$$

therefore we have

$$\Pi(t,s)U(t,s) = U(t,s)\Pi(t,s) = r(U(t,s))\Pi(t,s).$$

The following lemma is a direct result from Lemma 3.9.

Lemma 6.5. *Let Assumption 6.1 be satisfied. Then there exists a constant $\delta \in [0,1)$ such that*

$$k(U(t,s)) \leq \delta^{t-s}, \forall t > s. \tag{6.5}$$

Proof. By Lemma 3.9, we have

$$k(U(t,s)) \leq \Pi_{j=s,\ldots,t-1} k(A_j), \tag{6.6}$$

and by using (3.7)-(3.8), we deduce that $k(A)$ is continuous on the compact subset $A_- \leq A \leq A_+$, hence

$$\delta = \sup_{A_- \leq A \leq A_+} k(A) < 1,$$

and (6.5) follows from (6.6). $\square$

In addition, we make the following assumption.

Assumption 6.6. *For each $t > s$, we assume that there exist two positive numbers $\gamma > 0$ and $\beta > 0$ (with $\gamma < \beta$) such that*

$$\gamma \mathbb{1} \leq V_L\left(U\left(t,s\right)\right) \leq \beta \mathbb{1},$$

where $V_L\left(U\left(t,s\right)\right)$ is a positive left eigenvector of $U(t,s)$ associated with $r(U(t,s))$.

The main result of this section is the following theorem.

Theorem 6.7. *Let Assumptions 6.1 and 6.6 be satisfied. Then there exists a constant $\chi \geq 1$, such that for each $z \in \mathbb{R}^n$, and each $t > s$,*

$$\left\| \frac{U(t,s)z}{r\left(U(t,s)\right)} - \Pi(t,s)z \right\| \leq \chi\, \delta^{t-s} \, \|z\|,$$

where

$$\delta = \sup_{A_- \leq A \leq A_+} k(A) < 1.$$

Proof. We define

$$\mathbb{U}(t,s) = \frac{U(t,s)}{r\left(U(t,s)\right)}.$$

By the Perron-Frobenius theorem, we decompose $\mathbb{U}(t,s)$ into

$$\mathbb{U}(t,s) = \Pi(t,s) + B(t,s),$$

where $\Pi(t,s) \gg 0$ is the projector defined above, and we have

$$\Pi(t,s)\mathbb{U}(t,s) = \mathbb{U}(t,s)\Pi(t,s) = \Pi(t,s),$$

$$r(B(t,s)) < 1,$$

and

$$\Pi(t,s)\, B(t,s) = B(t,s)\, \Pi(t,s) = 0.$$

Since $V_L\left(U\left(t,s\right)\right) \geq \gamma \mathbb{1}$ is the left eigenvector, the map

$$|||x|||_{(t,s)} = \left\langle V_L\left(U\left(t,s\right)\right), |x| \right\rangle,$$

defines a norm on $\mathbb{R}^n$ which satisfies

$$|||\mathbb{U}(t,s)x|||_{(t,s)} = |||x|||_{(t,s)}, \forall t > s.$$

We reconsider following norm defined in Euclidean space $\mathbb{R}^n$,

$$\|x\| = \langle \mathbb{1}, |x| \rangle = \sum_{i=1}^{n} |x_i|, \quad \forall x \in \mathbb{R}^n.$$

By using Assumption 6.6, we deduce that

$$\gamma \|x\| \leq |||x|||_{(t,s)} \leq \beta \|x\|, \quad \forall x \in \mathbb{R}^n, \tag{6.7}$$

which is equivalent to

$$\frac{1}{\beta} |||x|||_{(t,s)} \leq \|x\| \leq \frac{1}{\gamma} |||x|||_{(t,s)}, \quad \forall x \in \mathbb{R}^n.$$

Moreover

$$\begin{aligned} \|\Pi(t,s)x\| &= \frac{\langle V_L(U(t,s)), |x| \rangle \|V_R(U(t,s))\|}{\langle V_L(U(t,s)), V_R(U(t,s)) \rangle} \\ &= \frac{|||x|||_{(t,s)} \|V_R(U(t,s))\|}{|||V_R(U(t,s))|||_{(t,s)}} \\ &\leq \frac{\beta \|x\| \|V_R(U(t,s))\|}{\gamma \|V_R(U(t,s))\|} \end{aligned}$$

hence

$$\|\Pi(t,s)\|_{\mathcal{L}(\mathbb{R}^n)} \leq \frac{\beta}{\gamma},$$

and

$$\|I - \Pi(t,s)\|_{\mathcal{L}(\mathbb{R}^n)} \leq 1 + \frac{\beta}{\gamma}. \tag{6.8}$$

We can choose $x \gg 0$ large enough (with $\|x\| = M$) so that

$$x + B_{\mathbb{R}^n}(0,1) \subset (0, +\infty)^n.$$

By using (6.8), we can find $\rho > 0$ small enough such that

$$\|I - \Pi(t,s)\|_{\mathcal{L}(\mathbb{R}^n)} \rho \leq 1.$$

Let $z \in \mathbb{R}^n$ with $\|z\| = \rho$. Then $(I - \Pi(t,s))z \in B_{\mathbb{R}^n}(0,1)$, and we deduce that

$$y = x + (I - \Pi(t,s))z \gg 0$$

and since $x \geq 0$ and $y \geq 0$, we obtain

$$|||x|||_{(t,s)} = |||\mathbb{U}(t,s)x|||_{(t,s)},$$

and

$$|||y|||_{(t,s)} = |||\mathbb{U}(t,s)y|||_{(t,s)},$$

and since $\|x\| = M$, the inequality (6.7) gives

$$|||y|||_{(t,s)} = |||x|||_{(t,s)} \leq \beta M.$$

Since $\Pi(t,s)(x-y)=0$, we have by Lemma 3.4

$$\begin{aligned}\|B(t,s)x-B(t,s)y\| &\leq \tfrac{1}{\gamma}\,|||\,B(t,s)x-B(t,s)y|||_{(t,s)}\\ &= \tfrac{1}{\gamma}\,|||\,\mathbb{U}(t,s)x-\mathbb{U}(t,s)y|||_{(t,s)}\\ &\leq \frac{3\,\beta\,M}{\gamma}\left(1-e^{-d_H(\mathbb{U}(t,s)x,\mathbb{U}(t,s)y)}\right),\end{aligned}$$

and since $1-e^{-x}\leq x, \forall x\geq 0$, we obtain

$$\|B(t,s)z\|=\|B(t,s)x-B(t,s)y\|\leq \frac{3\,\beta\,M}{\gamma}d_H\left(\mathbb{U}(t,s)x,\mathbb{U}(t,s)y\right)\leq \frac{3\,\beta\,M}{\gamma}k\left(\mathbb{U}(t,s)\right)d_H\left(x,y\right).$$

By Lemma 6.5, we deduce that

$$\|B(t,s)\|_{\mathcal{L}(\mathbb{R}^n)}\leq \chi\,\delta^{t-s}, \forall t>s,$$

with

$$\chi=\frac{3\,\beta\,M}{\gamma\rho}\sup_{\|z\|=\rho>0} d_H(x,x+z),$$

and the proof is completed. $\square$

Definition 6.8. *A non-negative matrix $A\in M_n(\mathbb{R})$ is called* ***Markovian*** *if the sum of each column of A is equal to 1. That is,*

$$\mathbb{1}^T A=\mathbb{1}^T,$$

where $\mathbb{1}=(1,\dots,1)^T$.

Now, assuming that each matrix A_t is Markovian, we deduce that $U(t,s)$ is also Markovian and

$$\mathbb{1}^T U(t,s)=\mathbb{1}^T, \forall t\geq s.$$

Together with Assumption 6.1, it follows that

$$r(U(t,s))=1,$$

and we obtain the following corollary.

Corollary 6.9. *Let Assumption 6.1 be satisfied. Assume in addition that each matrix A_t is Markovian. Then there exists a constant $\chi\geq 1$, such that for each $z\in\mathbb{R}^n$, and each $t>s$,*

$$\left\|U(t,s)z-\frac{\sum_{i=1}^n z_i}{\sum_{i=1}^n V_R\left(U(t,s)\right)_i}V_R\left(U(t,s)\right)\right\|\leq \chi\,\delta^{t-s}\,\|z\|,$$

where

$$\delta=\sup_{A_-\leq A\leq A_+} k(A)<1.$$

Appendix

A Quasi-circulant matrices

We consider the following matrix

$$A = \alpha_0 I + \alpha_1 P + ... + \alpha_n P^n, \tag{A.1}$$

where

$$P = \begin{bmatrix} 0 & 0 & \cdots & 0 & p_n \\ p_1 & 0 & \cdots & 0 & 0 \\ 0 & p_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & p_{n-1} & 0 \end{bmatrix}. \tag{A.2}$$

Lemma A.1. *Let $\lambda_1, ..., \lambda_n$ be the eigenvalues of P, then for each $\lambda_i, i = 1, ..., n$, we have*

$$\lambda_i^n = p_1 p_2 ... p_n.$$

Proof. For each $\lambda_i, i \leq i \leq n$, we have

$$\det(\lambda_i I - P) = 0. \tag{A.3}$$

(A.3) reads as

$$\det(\lambda_i I - P) = \begin{vmatrix} \lambda_i & 0 & \cdots & 0 & -p_n \\ -p_1 & \lambda_i & \cdots & 0 & 0 \\ 0 & -p_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & -p_{n-1} & \lambda_i \end{vmatrix} = \begin{vmatrix} \lambda_i & 0 & \cdots & 0 & -p_n \\ 0 & \lambda_i & \cdots & 0 & -\frac{p_1 p_n}{\lambda_i} \\ 0 & -p_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & -p_{n-1} & \lambda_i \end{vmatrix}$$

$$= \begin{vmatrix} \lambda_i & 0 & \cdots & 0 & -p_n \\ 0 & \lambda_i & \cdots & 0 & -\frac{p_1 p_n}{\lambda_i} \\ 0 & 0 & \cdots & 0 & \vdots \\ \vdots & \vdots & \ddots & \vdots & -\frac{p_1 p_2 ... p_{n-2} p_n}{\lambda_i^{n-2}} \\ 0 & 0 & \cdots & 0 & \lambda_i - \frac{p_1 p_2 ... p_n}{\lambda_i^{n-1}} \end{vmatrix} = \lambda_i^n - p_1 p_2 ... p_n = 0$$

□

Lemma A.2. *Let $\xi_1, \xi_2, ..., \xi_n$ be some eigenvectors of matrix P corresponding to the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$. Then $\xi_1, \xi_2, ..., \xi_n$ are also the eigenvectors of matrix A.*

Proof. Note that $\xi_1, \xi_2, ..., \xi_n$ are the eigenvectors of matrix P corresponding to the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, we have

$$P\xi_i = \lambda_i \xi_i, \quad 1 \leq \xi \leq n.$$

By (A.1), we obtain

$$\begin{aligned} A\xi_i &= \alpha_0 I\xi_i + \alpha_1 P\xi_i + ... + \alpha_n P^n \xi_i \\ &= (\alpha_0 + \alpha_1\lambda_i + \alpha_2\lambda_i^2 + ... + \alpha_n\lambda_i^n)\xi_i. \end{aligned}$$

Therefore, ξ_i is the eigenvector of matrix A corresponding to eigenvalue

$$\alpha_0 + \alpha_1\lambda_i + \alpha_2\lambda_i^2 + ... + \alpha_n\lambda_i^n.$$

The proof is completed. □

References

[1] G. Birkhoff, Extensions of Jentzsch's theorem, *Trans. Amer. Math. Soc.*, **85(1)** (1957), 219-227.

[2] G. Birkhoff, Uniformly semi-primitive multiplicative processes, *Trans. Amer. Math. Soc.*, **104** (1962), 37-51.

[3] P. J. Bushell, On projective contraction ratio for positive linear mappings, *J. London Math. Soc.*, **6** (1973), 256-258.

[4] J. E. Cohen, Ergodic theorems in demography, *Bull. Amer. Math. Soc.*, **1(2)** (1979), 275-296.

[5] P. J. Davis, *Circulant matrices*, Wiley, New York, 1979.

[6] J. Demongeot, A stochastic model for the cellular metabolism, *Recent developments in statistics* (1977), 655-662.

[7] J. Demongeot, Biological boundaries and biological age, *Acta Biotheoretica*, **57** (2009), 397-419.

[8] A. Ducrot, Q. Griette, Z. Liu, and P. Magal, *Differential Equations and Population Dynamics I: Introductory Approaches*. Springer Nature (2022).

[9] S. P. Eveson, and R. D. Nussbaum, An elementary proof of the Birkhoff-Hopf theorem, *Math. Proc. Cambridge Philos. Soc.*, **117(1)** (1995), 31-56.

[10] M. Golubitsky, E. B. Keeler, M. Rothschild, Convergence of the age structure: application of projective metric, *Theor. Popul. Biol.*, **7** (1975), 84-93.

[11] G. M. Hahn, A formalism describing the kinetics of some mammalian cell populations, *Math. Biosci.*, **6** (1970), 295-304.

[12] E. Hopf, An inequality for positive linear integral operators, *J. Math. Mech.*, **12(5)** (1963), 683-692.

[13] R. A. Horn and C. R Johnson, *Matrix analysis*. Cambridge University Press (2012).

[14] N. J. Higham, The Scaling and Squaring Method for the Matrix Exponential Revisited, *SIAM J. Matrix Anal. Appl.*, **26(4)** (2005), 1179-1193.

[15] D. Hilbert, über die gerade linie als kürzeste verbindung zweier punkte: Aus einem an herrn f. klein gerichteten briefe, *Mathematische Annalen*, **46(1)** (1895), 91-96.

[16] H. Inaba, A semigroup approach to the strong ergodic theorem of the multistate stable population process, *Math. Population Stud.*, **1(1)** (1988), 49-77.

[17] H. Inaba, Weak ergodicity of population evolution processes, *Math. Biosci.*, **96(2)** (1989), 195-219.

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

[19] M. A. Krasnosel'skii, and A. V. Sobolev, Spectral clearance of a focusing operator, *Funktsional'nyi Analiz i ego Prilozheniya*, **17(1)** (1983), 73-74.

[20] U. Krause, *Positive Dynamical Systems in Discrete Time*, Studies in Mathematics, vol. 62, de Gruyter, Berlin, (2015).

[21] B. Lemmens, and R. Nussbaum, *Nonlinear Perron-Frobenius Theory*, (Vol. 189). Cambridge University Press (2012).

[22] Z. Liu, P. Magal, O. Seydi, and G. Webb, Understanding unreported cases in the 2019-nCov epidemic outbreak in Wuhan, China, and the importance of major public health interventions, *Biology*, **9(3), 50** (2020).

[23] A. M. Ostrowski, Positive matrices and functional analysis, Recent Advances in Matrix Theory, H .Schneider, editor, Univ. of Wisconsin Press, Madison, Wisconsin, (1964), 81-101.

[24] M. Pituk, and C. Pötzsche, Ergodicity in nonautonomous linear ordinary differential equations, *J. Math. Anal. Appl.*, **479(2)** (2019), 1441-1456.

[25] E. Seneta, *Non-negative Matrices and Markow Chains*, Series in Statistics, Springer, 1973.

[26] Jr, H. D. Thames, and R. A. White, State-vector models of the cell cycle 1. Parametrization and fits to labeled mitoses data, *J. Theor. Biol.* , **67(4)** (1977), 733-756.

[27] S. Tuljapurkar, Population dynamics in variable environments. IV. Weak ergodicity in the Lotka equation, *J. Math. Biol.*, **14** (1982) 221-230.

[28] S. Tuljapurkar, An uncertain life: demography in random environments, *Theor. Popul. Biol.* , **35(3**) (1989), 227-294.

[29] P. P. Zabreiko, M. A. Krasnosel'skii, and Yu. V. Pokornyi , On a class of linear positive operators, *Functional Analysis and Its Applications*, **5** (1971), 272-279.