Convergence rate of the powers of an operator. Applications to stochastic systems

Bernard Delyon

▶ **To cite this version:**

Bernard Delyon. Convergence rate of the powers of an operator. Applications to stochastic systems. Bernoulli, 2017, 23 (4A), pp.2129-2180. 10.3150/15-BEJ778 . hal-01001705

HAL Id: hal-01001705
https://hal.science/hal-01001705

Submitted on 4 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Convergence rate of the powers of an operator. Applications to stochastic systems

Bernard Delyon *

June 4, 2014

Abstract

We extend the traditional operator theoretic approach for the study of dynamical systems in order to handle the problem of non-geometric convergence. We show that the probabilistic treatment developed and popularized under Richard Tweedie's impulsion, can be placed into an operator framework in the spirit of Yosida-Kakutani's approach. General Theorems as well as specific results for Markov chains are given. Application examples to general classes of Markov chains and dynamical systems are presented.

Keywords: Markov chains
AMS subject classification: 60J05

1 Introduction

This paper is mainly concerned with the asymptotical behaviour of homogeneous Markov chains, i.e. processes of the form

$$X_{n+1} = \varphi(X_n, U_n) \tag{1}$$

where U_n is an i.i.d. sequence and φ a certain function; the initial condition X_0 is deterministic or random. There exist essentially two different approaches for the analysis of the asymptotic behaviour of such systems: the operator theoretic approach and the probabilistic approach. The first approach focuses on the properties of the transition operator T defined as

$$Tf(x) = E[f(X_{n+1})|X_n = x] = E[f(\varphi(x, U_n))] = \int f(\varphi(x,u))\mu(du) \tag{2}$$

where μ is the distribution of U_n. The second is based on the fine study of the trajectories of X_n, especially the recurrence properties. The most typical objective is to understand how the function

$$T^n f(x) = E[f(X_n)|X_0 = x] \tag{3}$$

possibly converges to some limit, and at which rate. We know that this limit would be $\pi(f)$, where π is the invariant measure. This in particular allows to study arbitrary correlations

$$E[f(X_n)g(X_0)] = E[g(X_0)T^n f(X_0)]. \tag{4}$$

*IRMAR, Campus de Beaulieu, 35042 Rennes cedex, France

In the case of dynamical systems, i.e. $U_n = 0$, there is no pointwise limit to (3) whereas (4) may converge; this means that (3) converges actually in some weak sense. The problem can thus be summarized as: In which sense does $T^n f$ converges, for which functions f, and at which rate?

This paper deals with the first approach, although some fruitful ideas have been borrowed from the second one, especially concerning the case where there is no spectral gap.

A huge amount of literature is concerned with both of these points of views. In this section, we shall first give a sketch of the main ideas with typical examples of the simplest situations, and then we shall present our plan of action.

1.1 The Yosida–Kakutani Theorem and the Ionescu-Tulcea–Marinescu Theorem for quasi-compactness

It is well known that in the finite case (i.e. when X_n takes values in a finite state space), T is actually a matrix, and when T^n converges, the rate is always geometric. It is given by ρ^n, where ρ is the modulus of the second eigenvalue T. The gap between 1 (first eigenvalue) and ρ, is the spectral gap. If the eigenvalue 1 is not simple or if other eigenvalues of modulus one are present, the asymptotic behaviour of T^n is still given by the eigenvalues of modulus one and their eigenspaces, up to a remainder of order ρ^n. We present now the classical operator approach in the case of a general state space, which may be seen as the infinite dimensional extension of this matrix treatment. What is expected here is that for some norm $|.|$ and any function f with $|f| < \infty$

$$|T^n f - \pi(f)| \leq C\rho^n |f| \tag{5}$$

for some $C > 0$ and $0 < \rho < 1$. Examples are given below.

An operator T on a Banach space $(E, |.|)$ is said quasi-compact if some power of T can be written as

$$T^n = K + V \tag{6}$$

where V has spectral radius < 1 and K is compact. Quasi-compactness has been extensively studied [7]. The Yosida-Kakutani Theorem [18] says that if in addition the sequence $|T^k|$ is bounded then E splits as $E = E_c \bigoplus E_0$ where E_c is the finite dimensional space generated by the eigenvectors with eigenvalues of modulus 1, E_0 is closed with $TE_0 \subset E_0$ and the restriction of T to E_0 has spectral radius < 1. Denoting by λ_i, $i = 1, ...p$ the eigenvalues of T with modulus one, by E_i the corresponding eigenspaces, by P_i the projection on E_i parallel $\bigoplus_{j \neq i} E_j$, one has the equivalent formulation

$$T = \sum_{i=1}^{p} \lambda_i P_i + Q, \quad Q = TP_0 = P_0 T \tag{7}$$

where

$$|\lambda_1| = \cdots = |\lambda_p| = 1 \tag{8}$$

$$\text{each } P_i \text{ is a } |.|\text{-continuous projection, with finite rank if } i > 0 \tag{9}$$

$$\sum_{i=0}^{p} P_i = Id \tag{10}$$

$$P_i P_j = P_j P_i = 0, \quad 0 \leq i < j \leq p \tag{11}$$

$$|Q^n| \to 0. \tag{12}$$

The last equation implies of course that $|Q^n| \leq C\rho^n$ for some $C > 0$, $0 < \rho < 1$; another consequence of these equations is that for any $k \geq 1$

$$T^k = \sum_{i=1}^{p} \lambda_i^k P_i + Q^k. \tag{13}$$

A decade later, Ionescu-Tulcea and Marinescu provided a Theorem [8, 10, 9] useful for checking that quasi-compactness holds when $|T^n|$ is bounded[1]: it is assumed that there exists a weaker norm $\|.\|$ on E, for which $\{Tf : f \in E, |f| \leq 1\}$ is $\|.\|$-compact and in addition, for some $\gamma < 1$, $c \geq 0$ and $k > 0$, and all $f \in E$

$$|T^k f| \leq \gamma |f| + c\|f\|. \tag{14}$$

It turns out that conditions (6) and (14) have different natural domains of applications. For an illustrative purpose, we give below two simple but typical examples concerning Markov chains. Namely, we show that (6) is well suited for dealing with Harris chain with convergence in total variation of the distribution of the variable, whereas (14) is more adapted to non necessarily irreducible chains where, on the other hand, the transition has a contraction effect on the variable.

Example 1. We consider here a Markov chain on a measured space S, which satisfies a Doeblin condition in the sense that there exists a positive measure $\nu(dx)$ such that its transition kernel satisfies for all $x \in S$

$$p(x, dy) \geq \nu(dy).$$

Then one can write

$$Tf(x) = \int f(y)\nu(dy) + \int f(y)(p(x, dy) - \nu(dy))$$

and (6) applies with $|f| = \|f\|_\infty$, on the space E of bounded measurable functions, $\|V\| \leq 1 - \nu(S)$. With some extra effort, one can show that E_c is the one-dimensional space of constant functions. If π is the invariant measure, one gets

$$|T^n f - \pi(f)| \leq C\rho^n |f| \tag{15}$$

or by duality, for any initial measure μ

$$\|T^{*n}\mu - \pi\|_{TV} \leq C\rho^n \|\mu\|_{TV}$$

where $\|.\|_{TV}$ is the total variation norm.

Example 2. Let us consider now a chain with the form

$$X_{n+1} = \varphi(X_n, U_n)$$

where U_n is an i.i.d. sequence with distribution μ. Hence

$$Tf(x) = \int f(\varphi(x, u))\mu(du).$$

The function φ is supposed to satisfy adequate measurability assumptions and a uniform contraction property on the metric space (S, d). Specifically, we assume that

$$|\varphi(x, u)) - \varphi(y, u)| \leq \gamma d(x, y)$$

for some $\gamma < 1$ and all x, y, u. In this case it is easy to check that (14) applies with

$$\begin{aligned} &\|f\| = \|f\|_\infty \\ &|f| = \|f\| + [f] \\ &[f] = \sup_{x \neq y} \frac{f(x) - f(y)}{d(x, y)}. \end{aligned} \tag{16}$$

In order to have the $\|.\|$-compactness of $B = \{Tf : f \in E, |f| \leq 1\}$, we assume that the state space is compact. Convergence in total variation will not hold in general (e.g. the chain $X_{n+1} = (X_n + U_n)/2$, $U_n \sim \mathcal{B}(1, 1/2)$). But we still have geometric convergence for the stronger norm $f \mapsto |f|$.

[1] The point had been actually introduced much sooner by Doeblin and Fortet in [4], Eq.(2) and (3) p. 143, but in a more specific context.

1.2 The probabilistic approach

Let us consider an irreducible aperiodic Markov chain with invariant measure π. Interestingly, it appears that in many situations, geometric convergence like (15) does not occur, but nevertheless for many $f \in E$, $T^n f - \pi(f)$ converges exponentially fast to 0. In other words the convergence is not $|\!|.|\!|$-uniform, and sometimes this convergence does not follow an exponential rate, but is slower. This situation has been treated quite successfully with a very probabilistic approach, where the speed of convergence is related to the integrability of recurrence times. The reference [13], and more specifically [11], deals with these situations. Two key concepts are used: the ψ-irreducibility, and a drift condition for controlling moments of recurrence times. A simple illustrative example of this absence of spectral gap is the following operator on $(\mathbb{R}^{\mathbb{N}}, \|.\|_\infty)$:

$$Tf(x) = \frac{1}{2}(f(x) + f((x-1)_+), \quad x \in \mathbb{N}.$$

corresponding to the following chain on $\mathbb{N}$

$$X_{n+1} = (X_n - U_{n+1})_+, \qquad P(U_n = 0) = P(U_n = 1) = \tfrac{1}{2}.$$

The pointwise convergence $T^n f(x) \to \pi(f) = f(0)$ is very fast, but this convergence is not uniform. In particular this makes (15) impossible to occur with $|\!|f|\!| = \|f\|_\infty$. A possible operator theoretical approach is that one has for some weaker norm $\|.\|$

$$\|T^n f - \pi(f)\| < \rho_n |\!|f|\!| \tag{17}$$

for any $f \in E$, and some fixed decreasing sequence ρ_n. The norm $\|.\|$ introduced here has actually strong connections with the one involved in the Ionescu Tulcea-Marinescu approach. The rate of decrease of ρ_n depends on the choice of $\|.\|$. Notice that if in (17) the norms were equal, the convergence of ρ_n to zero would imply the geometric convergence; however, this is not the case any more when the norms are different.

1.3 Aim of the paper

The aim of this paper is to show that these ideas can be combined successfully and that they lead to an operator theoretic approach where non geometric convergence is considered. The main feature of this theory is to work simultaneously with two norms and to use this for measuring non geometric rates of convergence.

Our approach has essentially two steps: we first give conditions under which (7) to (11) hold with

$$\|Q^n f\| \le \rho_n |\!|f|\!|, \quad \rho_n \to 0 \tag{18}$$

instead of (12). This is the main objective of Section 2 (see Theorem 1). Notice that in this decomposition the Banach space is $(E, |\!|.|\!|)$, and the norm $\|.\|$ only appears in (18); in particular nothing guarantees that $|\!|Q^n f|\!|$ tends to zero.

Section 3 is concerned with geometric convergence, i.e. $\rho_n = C\rho^n$. Specifically Theorem 3 shows how the Yoshida-Kakutani and Ionescu Tulcea-Marinescu approaches can be combined into a single statement. This allows an easy treatment of chains having an irreducible component and another component behaving like Example 2 above.

Section 4 is concerned with sub-geometric convergence. Theorem 7 proposes a way to estimate the decay rate of the sequence ρ_n.

General theorems concerning Markov chains and examples are given throughout the paper in order to point out that this approach is very versatile for the study of a large class of dynamical systems, in particular for irreducible as well as for non-irreducible Markov chains.

2 General results

In the whole paper, we shall consider an operator T on a vector space $(E, |\!|.|\!|)$ endowed with another norm $\|.\|$. We shall denote by B_0, B the unit balls for these norms:

$$B_0 = \{f \in E : \|f\| \leq 1\}, \tag{19}$$
$$B = \{f \in E : |\!|f|\!| \leq 1\}. \tag{20}$$

We shall work under the following assumptions

(A0) $(E, |\!|.|\!|)$ is a Banach space, B is complete for the metric induced by $\|.\|$, and for some C_0

$$\forall f \in E, \quad \|f\| \leq C_0 |\!|f|\!|. \tag{21}$$

(A1) The number $C_T = \sup_n |\!|T^n|\!|$ is finite.

$(E, \|.\|)$ is typically not complete. For instance one can have $E = C_b(\mathbb{R})$, $|\!|f|\!| = \|f\|_\infty$ and $\|f\| = \sup_x \frac{|f(x)|}{1+|x|^2}$.

Theorem 1. *If in addition to (A0) and (A1), T is a sum of two operators*

$$T = K + V \tag{22}$$

both $|\!|.|\!|$-continuous and $\|.\|$-continuous, which satisfy for some $C_K > 0$ and for any n and any $f \in E$

$$KT^n KB \text{ is } \|.\|\text{-totally bounded}, \tag{23}$$
$$\|V^n f\| \leq \varepsilon'_n |\!|f|\!|, \qquad \varepsilon'_n \to 0, \tag{24}$$
$$\sum_{k\geq 0} |\!|KV^k|\!| < \infty, \tag{25}$$
$$|\!|Kf|\!| \leq C_K \|f\|, \tag{26}$$

then (7) to (11) and (18) hold true.

If T is $|\!|.|\!|$-continuous and $\|.\|$-continuous, and T^k satisfies the assumptions above for some $k > 0$, then (7) to (11) and (18) hold true.

The proof is postponed to Appendix A. This proof utilizes the more general Theorem 10 stated in Section A.1, and is based on an extensive use of the identity:

$$T^n = \sum_{i=1}^{n} T^{n-i}(T-V)V^{i-1} + V^n = \sum_{i=1}^{n} T^{n-i}KV^{i-1} + V^n. \tag{27}$$

Very coarsely the assumptions imply that for any sequence $f_k \in B$, the sequence $T^k f_k$ is $\|.\|$-totally bounded. This allows us to prove that E is the direct sum of two $|\!|.|\!|$-closed, T-stable spaces

$$E = \{f : \|T^n f\| \to 0\} \oplus \{f : \liminf_n \|f - T^n f\| = 0\} = E_0 \oplus E_c. \tag{28}$$

Next we prove that E_c is finite dimensional (by proving that its unit ball is compact) with a basis of eigenvectors. The projection P_0 of Equation (7) is then the projection on E_0 parallel to E_c.

Application to Markov chains. We shall consider a measurable space $(S, \mathscr{F})$ with a measurable weight function $v \geq 1$ and we adopt the following notation

$$\|f\|_v = \|f/v\|_\infty. \tag{29}$$

We denote by E the Banach space of bounded measurable functions on $(S, \mathscr{F})$. We recall that a transition operator on $(S, \mathscr{F})$ is a function $(x, A) \mapsto T(x, A)$ such that for any $x \in S$, $A \to T(x, A)$ is a probability measure, and for any $A \in \mathscr{F}$, $x \to T(x, A)$ is measurable.

Theorem 2. *Let T be a Markov transition operator:*

$$(Tf)(x) = \int_y f(y)T(x,dy).$$

Assume that for some set K_0 and some $c_v > 0$

$$Tv(x) \leq v(x) - c_v, \qquad \forall x \notin K_0 \tag{30}$$
$$Tv \text{ is bounded on } K_0 \tag{31}$$

and that there exists another kernel $K(x,dy)$ such that $0 \leq K(x,dy) \leq T(x,dy)$, and such that for some $\varepsilon > 0$, and some non-negative measure ν one has

$$K(x,S) \geq \varepsilon, \quad \forall x \in K_0 \tag{32}$$
$$K(x,S) = 0, \quad \forall x \notin K_0 \tag{33}$$
$$\|Kf\|_\infty \leq \nu(|f|), \quad \forall f \in E \tag{34}$$
$$\nu(v) < \infty. \tag{35}$$

Set

$$|\!|\!| f |\!|\!| = \|f\|_\infty \tag{36}$$
$$\|f\| = \|f\|_v. \tag{37}$$

Then Theorem 1 applies with K and $V = T - K$. In particular Equations (7) to (11) and (18) hold true.

If in addition there is no measurable set A such that $x \mapsto T(x,A)$ is a non-trivial indicator function[2] *then there exist a measure π and a sequence $\rho_n \to 0$ such that for any $f \in E$*

$$\|T^n f - \pi(f)\|_v \leq \rho_n \|f\|_\infty. \tag{38}$$

The proof of this consequence of Theorem 1 is postponed to Appendix B.

REMARK. Equation (30) is known as the "drift condition" (cf Theorem 11.0.1 of [13] or Proposition 5.10 in [14]). Equations (32) to (35) are reminiscent of the T-chain property (cf [13] Theorem 6.0.1), used to check the irreducibility assumption (cf. [13] p. 87). However, the Feller property is not required here. Equation (33) is not a restriction, since cancelling K outside K_0 does not affect the other assumptions. The essential difficulty with the present assumptions is that the set K_0 has to be the same in (30) and in (32). Notice however that the sets K_0 satisfying assumptions (32) and (33) are stable by finite union.

Example. Consider the Markov chain on $\mathbb{R}_+$ defined by

$$X_{n+1} = X_n + 1 + X_n^\alpha W_{n+1}$$

where W_n is an i.i.d. sequence of non-zero centred random variables with values in $[-1,1]$, with a non zero absolutely continuous component. In addition, we assume that

$$1/2 < \alpha < 1.$$

Take

$$v(x) = x^p + 1$$

for some $p \leq 1$ which will be chosen later as $2(1-\alpha)$. Then

$$Tv(x) = 1 + E[(x + 1 + x^\alpha W_1)^p].$$

[2] This would mean that $1_{X_1 \in A}$ would be a deterministic non-constant function of the initial state X_0.

By the second order Taylor formula applied to the function v in the neighbourhood of $x+1$, there exist a random number $0<\theta<1$ such that

$$\begin{aligned} Tv(x) &= 1+(x+1)^p - \frac{p(1-p)}{2}x^{2\alpha}E[(x+1+\theta x^\alpha W_1)^{p-2}W_1^2] \\ &\leq 1+(x+1)^p - \frac{p(1-p)}{2}x^{2\alpha}(x+1-x^\alpha)^{p-2}\sigma^2 \end{aligned}$$

where σ^2 is the variance of W_1. Taking $p=2(1-\alpha)$, we have $0<p<1$ and

$$\begin{aligned} Tv(x) &\leq 1+(x+1)^p - \frac{p(1-p)}{2}\Big(\frac{x}{x+1-x^\alpha}\Big)^{2\alpha}\sigma^2 \\ &\leq 1+x^p - \frac{p(1-p)}{3}\sigma^2 \quad \text{for } x \text{ large enough.} \end{aligned}$$

Equation (30) is satisfied for some interval $K_0=[0,M]$. Equation (31) is obvious. In order to check Equations (32) to (35), notice that if the absolutely continuous component of W_1 has a density $\geq\varepsilon$ on a subset A of $[-1,1]$ with positive measure, $K(x,dy)$ can be taken as $\varepsilon\lambda(A)$ times the distribution of $x+1+x^\alpha\tilde{W}_1$, where $\tilde{W}_1$ has density $1_A/\lambda(A)$, ν is some multiple of the uniform measure on $[0,M+M^\alpha+1]$. Therefore theorem applies. In order to get (38) it remains to prove that $T1_A=1_B$ is impossible unless $B=\mathbb{R}_+$ or $B=\emptyset$. If B is non trivial one can find two sequences x_n and y_n having the same limit such that $x_n\in B$ and $y_n\notin B$. The relation $T1_A=1_B$ would mean that for each n, the distributions of $x_n+1+x_n^\alpha W_1$ and $y_n+1+y_n^\alpha W_1$ are mutually singular (supported on A and A^c), which is impossible for n large because W_1 has an absolutely continuous component. As a consequence, B is necessarily trivial and (38) holds.

Notice that nevertheless $E[X_n]=E[X_0]+n$.

3 Geometric convergence: Quasi-compactness

In this section, we give a theorem which encompasses both Yosida-Kakutani and Ionescu-Tulcea-Marinescu Theorems, and present an application to Markov chains which mixes both kinds of situations presented above. A specific application to autoregressive processes with Markov switching is finally studied. We recall that B denotes the unit closed ball for the norm $|\!|\!|.|\!|\!|$.

Theorem 3. *Let T be an operator on $(E,|\!|\!|.|\!|\!|)$ satisfying (A0), (A1) and*

(A2) *T is $\|.\|$-continuous. For some $\|.\|$-totally bounded set K_B, $\gamma<1$, $c>0$ and $q>0$*

$$T^qB \subset \gamma B + K_B \tag{39}$$

$$|\!|\!|T^q f|\!|\!| \leq \gamma |\!|\!| f |\!|\!| + c\|f\|. \tag{40}$$

Then Equations (7) to (12) hold.

Like Theorem 1, this theorem is a consequence of the general Theorem 10 stated in Section A.1; its proof is postponed to Appendix C.

The following theorem may seem very general and unclear for the applications. We should point out that we intend to bridge a continuum over two extreme cases: the convergence of the Markov chain in Wasserstein distance and the convergence in total variation. This will be exemplified in Theorem 6.

Let us just mention that [.] below is typically a Lipschitz norm like in Equation (16), and that the restriction to sets S_x in (45) to (47) allows to get more local assumptions (i.e. less uniformity). The theorem says that if locally T can be lower bounded by an operator with nice properties, then quasi-compactness holds.

Theorem 4. *Let (S,d) be a metric space and $\mathscr{B}$ its Borel σ-field. We assume that is given a continuous function $v(x) \geq 1$ on S such that for any $A > 0$, $\{x : v(x) \leq A\}$ is compact. Consider a vector space E of $\mathscr{B}$-measurable functions defined on S containing compactly supported Lipschitz functions. On E is defined a semi-norm $f \mapsto [f]$ and we set for any function f on S:*

$$\left|\!\left|\!\left| f \right|\!\right|\!\right| = \|f\| + [f], \tag{41}$$

$$\|f\| = \sup_x \frac{|f(x)|}{v(x)}. \tag{42}$$

We assume that $(E, \left|\!\left|\!\left| . \right|\!\right|\!\right|)$ is a Banach space and that (A0) holds.

Let T be a Markov transition operator defined on E. We assume the existence of $0 < \gamma_b, \gamma_v < 1$ and $c_v > 0$ such that

$$[Tf] \leq \gamma_b [f], \quad f \in E, \tag{43}$$

$$Tv(x) \leq \gamma_v v(x) + c_v. \tag{44}$$

We assume the existence of functions $\varepsilon_d > 0$ and $c_d > 0$, and for any $x \in S$, the existence of a non-negative kernel $K_x(y, dz)$ and a neighbourhood S_x of x, such that for any $y \in S_x$ and $f \in E$,

$$K_x(y, dz) \leq T(y, dz), \tag{45}$$

$$K_x(y, S) \geq \varepsilon_d(x), \tag{46}$$

$$|K_x f(y) - K_x f(x)| \leq c_d(x, y)[f]. \tag{47}$$

Moreover the function $c_d(.,.)$ is assumed to be bounded on compact subsets of $S \times S$, and $\varepsilon_d(x)$ is positive and satisfies

$$\lim_{v(x) \to \infty} \varepsilon_d(x) v(x) = +\infty, \tag{48}$$

$$\forall\ A, \quad \min_{v(x) \leq A} \varepsilon_d(x) > 0. \tag{49}$$

Then Theorem 3 holds with a pair of norms $(\left|\!\left|\!\left| . \right|\!\right|\!\right|', \|.\|')$ respectively equivalent to $\left|\!\left|\!\left| . \right|\!\right|\!\right|$ and $\|.\|$. In particular, if the constant functions are the only eigenvectors of T with an eigenvalue of modulus 1, there exist $C > 0$, $0 < \rho < 1$ and a probability measure π such that for any $f \in E$,

$$\left|\!\left|\!\left| \pi(f) 1 - T^n f \right|\!\right|\!\right| \leq C \rho^n \left|\!\left|\!\left| f \right|\!\right|\!\right| \tag{50}$$

and $\pi(v) < \infty$.

The proof is postponed to Appendix D. We use Theorem 3 with $q = 1$. The idea is to set

$$Kf(x) = \sum_{i=1}^{n} \theta_i(x) K_{x_i} f(x_i)$$

where $\theta_1, \ldots \theta_n$ is a partition of the unity of a large portion of the space, each x_i being a point of the support of θ_i. Clearly $K(B)$ is compact. It remains to prove that $\|(T - K)f\| \leq \gamma \|f\|$ (which implies (39)) and that (40) holds true.

We shall consider two examples, one where $[f] = 0$ and we get geometric convergence in $\|.\|$ norm (which, by duality, corresponds to geometric weighted total variation convergence for the distribution of the Markov chain), and another case where $[.]$ plays an important role.

Application to geometric total variation convergence. In the case $[f] = 0$, the kernel $K_x(y, dz)$ should not depend on y, and we get the following corollary:

Corollary 5. (Weighted local Doeblin condition) *Let (S, d) be a metric space and $\mathscr{B}$ its Borel σ-field. We assume that is given a continuous function $v(x) \geq 1$ on S such that for any $A > 0$, $\{x : v(x) \leq A\}$ is compact. Consider a vector space E of $\mathscr{B}$-measurable functions defined on S containing compactly supported Lipschitz functions. We set for any function f on S:*

$$\left|\!\left|\!\left| f \right|\!\right|\!\right| = \sup_x \frac{|f(x)|}{v(x)}. \tag{51}$$

We assume that $(E, \left|\!\left|\!\left| . \right|\!\right|\!\right|)$ is a Banach space and that (A0) holds.

Let T be a Markov transition operator defined on E. We assume the existence of $0 < \gamma_v < 1$ and $c_v > 0$ such that

$$Tv(x) \leq \gamma_v v(x) + c_v. \tag{52}$$

We assume the existence of a function $\varepsilon_d > 0$, and for any $x \in S$ of a non-negative kernel $K_x(dz)$ and a neighbourhood S_x of x, such that

$$K_x(dz) \leq T(y, dz), \quad y \in S_x \tag{53}$$

$$K_x(S) \geq \varepsilon_d(x). \tag{54}$$

The function $\varepsilon_d(x)$ is positive and satisfies

$$\lim_{v(x)\to\infty} \varepsilon_d(x) v(x) = +\infty, \tag{55}$$

$$\forall\ A, \quad \min_{v(x)\leq A} \varepsilon_d(x) > 0. \tag{56}$$

Then Equations (7) to (12) hold. In particular, if the constant functions are the only eigenvectors of T with an eigenvalue of modulus 1, there exist $C > 0$, $0 < \rho < 1$ and a probability measure π such that for any $f \in E$,

$$\left|\!\left|\!\left| \pi(f) - T^n f \right|\!\right|\!\right| \leq C\rho^n \left|\!\left|\!\left| f \right|\!\right|\!\right| \tag{57}$$

and $\pi(v) < \infty$.

In many cases $\varepsilon_d(x) = 1/2$ will do the job, but in the following example the situation is more complicated:

$$X_{n+1} = \begin{cases} \frac{1}{2} X_n & \text{with probability } 1 - p(X_n) \\ V_n & \text{with probability } p(X_n) \end{cases}$$

where V_n is an i.i.d. sequence and p is a positive continuous function of x; V_n can be constant. We see that only the second type of transition contributes to the convergence in total variation, even if $V_n \equiv 0$, this is why we shall need $p(x)$ not to be too small. Let us assume that for some $0 < \alpha < 1$

$$E[|V_n|^\alpha] < \infty$$
$$\lim_{x\to\infty} p(x)|x|^\alpha = +\infty$$
$$p \text{ is positive and continuous}$$

then Equations (52) to (56) are clearly satisfied with

$$v(x) = |x|^\alpha + 1$$
$$K_x(f) = \tfrac{1}{2} p(x) E[f(V_1)]$$
$$\varepsilon_d(x) = \tfrac{1}{2} p(x)$$
$$S_x = \{y : p(y) > p(x)/2\}$$

and the exponential convergence holds. Notice that the continuity of p is far from being necessary since what we need is only that S_x is a neighbourhood of x, and that the minimum of p on any bounded set is > 0.

Application to functional autoregressive processes with Markov switching. We consider the following mixed Markov process $(I_n, X_n) \in S$ where $S = \{1, \dots s\} \times \mathbb{R}^d$:

$$P(I_{n+1} = j | I_n = i) = p_{ij}, \quad 1 \leq i, j \leq s \tag{58}$$

$$X_{n+1} = \alpha(I_n)\varphi(X_n) + \psi(I_n, V_n) \tag{59}$$

where α is a matrix valued measurable function, φ and ψ are vector valued measurable functions, and V_n is an independent i.i.d. sequence. In other words

$$Tf(i,x) = \sum_k p_{ik} E[f(k, \alpha(i)\varphi(x) + \psi(i, V_1))].$$

If for all i the variable $\psi(i, V_1)$ has a density, we can apply Corollary 5 at the price of extra reasonable assumptions because (53) and (54) would be satisfied for some kernel K_x independent of y (the continuity of φ is important here); our point is to deal with singular measures. As in [2], Th.1.4, we have made efforts to have conditions which allow for non-contracting values for α, as can one see in Equation (61):

Theorem 6. *Consider the Markov chain defined by (58) and (59). We assume that the chain I_n is irreducible and aperiodic with invariant measure π on its finite state space, and that for some $q > 0$*

$$|\varphi(y) - \varphi(z)| \leq |y - z| \tag{60}$$

$$\sum_i \pi_i \log(\|\alpha(i)\|) < 0 \tag{61}$$

$$\sup_i E[|\psi(i, V_1)|^q] < +\infty \tag{62}$$

where $\|.\|$ is the usual matrix norm and $|.|$ the euclidean norm. Then Theorem 4 applies and (50) holds with the norm

$$\mathbf{|}f\mathbf{|}' = \sup_{i,x,x'} \frac{|f(i,x) - f(i,x')|}{|x - x'|^\eta} + \sup_{i,x} \frac{|f(i,x)|}{|x|^\eta + 1}$$

for η small enough. This implies that for any realization (I_n, X_n) of the chain at time n with an arbitrary initial distribution, one can find a coupling with a pair (I', X') having the stationary distribution, such that

$$P(I_n \neq I') + E[|X_n - X'|^\eta] < C\rho^n(1 + E[|X_0|^\eta]).$$

Proof. We will choose

$$[f] = \sum_i \nu_i [f]_i, \quad [f]_i = \sup_{x,y} \frac{|f(i,x) - f(i,y)|}{|x - y|^\eta}$$

$$v(i,x) = |x|^\varepsilon e^{\varepsilon\lambda(i)} + 1$$

$$d((i,x),(j,y)) = 1_{i \neq j} + |x - y|^\eta.$$

for some constants ν_i and $\lambda(i)$ which will be specified later. Concerning K and the neighbourhoods $S_{i,x}$ we simply set:

$$K = T$$

$$S_{i,x} = \{i\} \times \mathbb{R}^d$$

In that case (45) and (46) are obvious ($\varepsilon_d = 1$), and (47) will be a consequence of (43). The technical part is to prove that (43) and (44) hold true. We now focus on (44). We first note that since

$$X_{n+1} = \alpha(I_n)(\varphi(X_n) - \varphi(0)) + \Big(\alpha(I_n)\varphi(0) + \psi(I_n, V_n)\Big)$$

we can assume that $\varphi(0) = 0$. Unsurprisingly, the contraction property (44) is related to the rate at which the product of $\alpha(I_k)$'s converges to zero, this one being itself controlled by the speed at which the law of large numbers acts on the sums of $\log(\|\alpha(I_k)\|)$'s. This uses classically the Poisson equation: Since the chain I_n is irreducible aperiodic on a finite state space, there exists a unique (up to a constant) solution λ to the Poisson equation

$$E[\lambda(I_1)|I_0 = i] = \lambda(i) - l(i) + \pi(l), \quad l(i) = \log(\|\alpha(i)\|)$$

(it is simply $\lambda = \sum_{k=0}^{\infty}(T_0^k - \pi)l$ where $T_0 = (p_{ij})_{1\le i,j\le s}$ is the transition operator of the chain I_n). The process

$$Z_n = |X_n|^\varepsilon e^{\varepsilon\lambda(I_n)}$$

satisfies, thanks to (59) ,(60), and $\varphi(0) = 0$:

$$\begin{aligned} Z_{n+1} &\le (\|\alpha(I_n)\|\,|\varphi(X_n)| + |\psi(I_n, V_n)|)^\varepsilon e^{\varepsilon\lambda(I_{n+1})} \\ &\le \|\alpha(I_n)\|^\varepsilon\,|X_n|^\varepsilon e^{\varepsilon\lambda(I_{n+1})} + |\psi(I_n, V_n)|^\varepsilon e^{\varepsilon\lambda(I_{n+1})} \\ &= Z_n e^{\varepsilon\{\log\|\alpha(I_n)\| + \lambda(I_{n+1}) - \lambda(I_n)\}} + e^{\varepsilon\lambda(I_{n+1})}|\psi(I_n, V_n)|^\varepsilon. \end{aligned}$$

And since the factor of ε is bounded, we have for some c

$$\begin{aligned} Z_{n+1} &\le Z_n\Big(1 + \varepsilon(\lambda(I_{n+1}) - \lambda(I_n) + \log\|\alpha(I_n)\|) + c\varepsilon^2\Big) + e^{\varepsilon\lambda(I_{n+1})}|\psi(I_n, V_n)|^\varepsilon \\ E[Z_{n+1}|\mathscr{F}_n] &\le Z_n\Big(1 + \varepsilon\pi(l) + c\varepsilon^2\Big) + e^{\varepsilon\sup_i\lambda(i)}\sup_i E[|\psi(i, V_1)|^\varepsilon], \end{aligned}$$

where $\mathscr{F}_n$ stand for the σ-field $\sigma(I_i, X_i, 0 \le i \le n)$. Hence, if we take $\varepsilon \le q$ such that $\varepsilon\pi(l) + c\varepsilon^2 < 0$, we obtain (44). Concerning (43):

$$\begin{aligned} |Tf(i,y) - Tf(i,x)| &\le \sum_k p_{ik}E[|f(k, \alpha(i)\varphi(y) + \psi(i, V_1)) - f(k, \alpha(i)\varphi(x) + \psi(i, V_1))|] \\ &\le |\varphi(y) - \varphi(x)|^\eta\|\alpha(i)\|^\eta\sum_k p_{ik}[f]_k \\ [Tf]_i &\le \|\alpha(i)\|^\eta\sum_k p_{ik}[f]_k \\ \sum_i \nu_i[Tf]_i &\le \sum_{i,k}\nu_i\|\alpha(i)\|^\eta p_{ik}[f]_k. \end{aligned}$$

We see that if we can find ν such that

$$\forall\, k, \quad \sum_i \|\alpha(i)\|^\eta\nu_i p_{ik} < \nu_k \tag{63}$$

then Equation (43) will be satisfied. To this aim, we define

$$\nu = \pi + \eta\sum_{k\ge 1}(\pi.l - \pi(l)\pi)P^k$$

with the notation

$$(\pi.l)(i) = \pi_i l(i).$$

Set $\tau_i = \|\alpha(i)\|^\eta$; since

$$\tau_i = 1 + \eta l(i) + O(\eta^2)$$

we get

$$\nu.\tau = \pi + \eta \sum_{k\geq 1} (\pi.l - \pi(l)\pi)P^k + \eta\pi.l + O(\eta^2)$$

hence

$$\begin{aligned}\nu.\tau P &= \pi + \eta \sum_{k\geq 1} (\pi.l - \pi(l)\pi)P^k + \eta\pi(l) + O(\eta^2) \\ &= \nu + \eta\pi(l) + O(\eta^2).\end{aligned}$$

This equation implies that for η small enough, Equation (63) is satisfied. In particular we shall impose $\eta \leq \varepsilon$. We have now proved (43) to (48).

As a byproduct, Equation (43) implies that any eigenfunction f, with associated eigenvalue $|\lambda| = 1$, does not depend on x, and consequently, since I_n is irreducible, f is necessarily constant.

Theorem 4 applies and (50) holds with

$$\left|\!\left|\!\left| f \right|\!\right|\!\right| = \sup_{x,i} \frac{|f(i,x)|}{|x|^\varepsilon e^{\varepsilon\lambda(i)} + 1} + \sum_i \nu_i [f]_i.$$

Since by irreducibility, $\nu_i > 0$ for all i, this norm is equivalent to

$$N(f) = \sup_{i,x} \frac{|f(i,x)|}{|x|^\varepsilon + 1} + \sup_{i,x,y} \frac{|f(i,x) - f(i,y)|}{|x-y|^\eta}.$$

This norm is also equivalent to $\left|\!\left|\!\left| f \right|\!\right|\!\right|'$ because, on the one hand, $\eta \leq \varepsilon$, and on the other hand

$$\begin{aligned}\sup_{i,x} \frac{|f(i,x)|}{|x|^\eta + 1} &\leq \sup_{i,x} \frac{|f(i,x) - f(i,0)| + |f(i,0)|}{|x|^\eta + 1} \\ &\leq \sup_{i,x} \frac{|f(i,x) - f(i,0)|}{|x|^\eta} + \sup_i |f(i,0)| \\ &\leq N(f).\end{aligned}$$

By the duality properties of the Wasserstein distance (cf [17] Theorem 5.10 and Equation (5.11))

$$\inf_{I_n, I', X_n, X'} P(I_n \neq I') + E[|X_n - X'|^\eta] = \sup_{f \text{ Lipschitz}} E[f(I_n, X_n) - f(I', X')]$$

where the infimum is taken over all the pairs of random variables (I', X') and (I_n, X_n) having respectively the stationary distribution and the chain distribution at time n, and f is 1-Lipschitz w.r.t. the distance d. The expectation in the right-hand side is just $E[(Q^n f)(I_0, X_0)]$, which is smaller than $C\rho^n(1 + E[|X_0|^\eta])$. □

4 Subgeometric rates

In the rest of the paper, we shall find conditions under which the rate of convergence of V^n to 0 will give us an insight about the rate of convergence of Q^n to 0. We set for any operator S on E

$$\begin{aligned}\|S\|_{E0} &= \sup_{\left|\!\left|\!\left| f \right|\!\right|\!\right| \leq 1} \|Sf\| \\ \|S\|_{0E} &= \sup_{\|f\| \leq 1} \left|\!\left|\!\left| Sf \right|\!\right|\!\right|.\end{aligned}$$

With this convention, one has

$$\|UV\| \le \|U\|_{E0}\|V\|_{0E}$$
$$|\!|\!|UV|\!|\!| \le \|U\|_{0E}\|V\|_{E0}.$$

We shall consider positive rate sequences α_n, $n \ge 1$, satisfying the conditions (R1) to (R3) below. For instance, sequences like $\alpha_n = (n+1)^{-p}$, $p > 1$, or $\alpha_n = \exp -\sqrt{n}$, or $\alpha_n = (n+1)^{-1}(\log(n+1))^{-2}$ satisfy these assumptions (notice that the first part of (R2) holds if $x \mapsto \log \alpha_x$ is convex). These sequences make it easy to solve some recursive equations (cf. Appendix F).

Theorem 7. *Let (A0) be satisfied and T be a $|\!|\!|.|\!|\!|$- and $\|.\|$-continuous operator on E satisfying (A1), Equations (7) to (11) and (18). Let α_n be a sequence satisfying*

(R1) *$n \mapsto \alpha_n$ is decreasing*

(R2) *$n \mapsto \dfrac{\alpha_{n+1}}{\alpha_n}$ is increasing and converges to 1*

(R3) $\displaystyle\sum_{n\ge 1} \frac{\alpha_n^2}{\alpha_{2n}} < \infty.$

We assume that T can be rewritten as $T = K + V$ with

$$\|V^k\|_{E0} \le C_1\alpha_k, \quad k > 0, \tag{64}$$
$$|\!|\!|KV^k|\!|\!| \le C_2\alpha_k, \quad k > 0, \tag{65}$$
$$|\!|\!|KQ^k|\!|\!| \to 0 \quad as \quad k \to \infty \tag{66}$$

(Equations (65) and (66) are clearly satisfied if (64) and (26) hold true). Then one has for some $C > 0$ and all $n > 0$

$$\|Q^n\|_{E0} \le C\sum_{k\ge n} \alpha_k.$$

If in addition $\sup_n \|T^n\| < \infty$, then

$$\|Q^n\|_{E0} \le C\alpha_n. \tag{67}$$

The proof is based on (27) and on the key result of Proposition 12. It is postponed to Appendix E.

REMARKS. 1/ If Theorem 1 is used for checking the assumptions, there is no need to check (25), which is automatically satisfied thanks to (24), (26) and the summability of α_n (consequence of (R1) and (R3)). 2/ Condition (R2) excludes geometric rates. The theorem is indeed wrong in this case: For example in the finite dimensional case, Theorem 1 holds with $V = 0$, and (67) only holds with some geometric rate.

Application to Markov chains. We consider here the Markov chain of Theorem 2 but (30) is strengthened as

$$Tv(x) \le v(x) - \theta(v(x)), \quad x \notin K_0 \tag{68}$$

for some function θ, e.g. $\theta(u) = u^q$, $0 < q < 1$. Our goal here is to use this information for bounding the sequence ρ_n in (38).

We need a preparatory lemma which will be essential for working with (68); the point of this lemma is to bring out a function ζ which satisfies (71), is significantly larger than $\zeta(x) = x$ and that can be easily iterated (Equation (70) implies $\zeta^{(n)}(x) = \psi^{(-1)}(\psi(x) + n)$):

Lemma 8. *Let θ be a non-decreasing non-negative concave differentiable function on $[0,+\infty)$ with a derivative which tends to zero at infinity, and define for $x \geq 0$*

$$\psi(x) = \int_0^x \frac{1}{\theta(y)} dy \tag{69}$$

$$\zeta(x) = \psi^{(-1)}(\psi(x)+1). \tag{70}$$

We assume that ψ is finite and tends to infinity. Let $\theta_1 \geq 0$ be such that for $x \geq \theta_1$

$$\theta(x) \leq x.$$

Then ζ is concave and for $x \geq \theta_1$

$$\zeta(x - \theta(x)) \leq x. \tag{71}$$

The proof is postponed to Appendix G.

Theorem 9. *Let all the assumptions and notations of Theorem 2 hold and assume that (30) is strengthened as*

$$Tv(x) \leq v(x) - \theta(v(x)), \qquad x \notin K_0 \tag{72}$$

for some concave function θ satisfying the assumptions of Lemma 8 with $\theta_1 = \min_x(v(x))$. In addition we assume that $\min_{x \geq \theta_1} x - \theta(x) > 0$ and that the sequence

$$\alpha_n = \frac{1}{\psi^{(-1)}(n)}$$

(ψ is given by (69)) satisfies the conditions (R1) to (R3) of Theorem 7. Then for some $c > 0$ and any bounded measurable function f

$$\sup_x \Big| \frac{Q^n f(x)}{v(x)} \Big| \leq c\rho_n \|f\|_\infty, \qquad \rho_n = \sum_{k \geq n} \alpha_k \tag{73}$$

$$\pi(|Q^n f|) \leq c\alpha_n \|f\|_\infty. \tag{74}$$

The proof is postponed to Appendix H.
We find the following matchings between drift function and rates

$\theta(t)$	α_n	ρ_n
$\log(t+1)^2$	$\sim n^{-1}(\log n)^{-2}$	$\sim (\log n)^{-1}$
t^q, $0 < q < 1$	$\sim n^{-1/(1-q)}$,	$\sim n^{-q/(1-q)}$
$\frac{ct}{\log(t+1)}$	$\sim e^{-\sqrt{2cn}}$	$\sim e^{-\sqrt{2cn}}\sqrt{n}$.

Applications of this kind of result in the field of Markov chains are not uncommon. For example in [11] (Example

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

where W_n is an i.i.d. sequence with $E[W_1] < 0$. Under the assumption that there exists an integer $m \geq 2$ such that

$$E[W_1^m] < \infty$$

they prove that the drift condition (72) is satisfied with

$$v(x) = (x+1)^m$$
$$\theta(x) = x^\alpha, \quad \alpha = \frac{m-1}{m}.$$

In Theorem 3.6 they state that for any x, $\sup_{\|f\|_\infty \leq 1} |Q^n f(x)| = o(n^{-\alpha/(1-\alpha)})$, which is somewhat intermediary between (73) and (74). On this example, we clearly see the interpretation of the difference of rates between (73) and (74): if the initial state $X_0 = x_0$ is very large, it takes a long time to come back to the invariant measure (this time is certainly proportional to x_0), but if the initial state is drawn from π, it won't be large and the convergence rate is increased.

They also give an application to Monte Carlo Markov Chains. We would like now to convince the reader that Theorem 7 can also be applied to dynamical systems.

Application to dynamical systems. Consider the following application defined on $[0,1]$

$$v(x) = \begin{cases} x(1+2^\gamma x^\gamma) & 0 \leq x < 1/2 \\ 2x-1 & 1/2 \leq x \leq 1 \end{cases} \tag{75}$$

where $0 < \gamma < 1$ is fixed, and the corresponding operator

$$Tf(x) = f(v(x)). \tag{76}$$

We are interested in the asymptotics of T^n. There exists an extensive literature on the subject [12, 6] and the result we are going to present here, Equation (80), is already known [19]; our point is to give a new and direct proof of this estimate which plays a key role in the obtainment of central limit theorems (through the Gordin-Liverani Theorem), and which is known to be optimal [16]. Notice that this proof does not require any explicit assumption on the invariant measure (see Equation (5.2) in [12]). We detail only here the example (75) but it will appear clearly that the following development extends to many other cases. Nevertheless, we feel that such extensions fall beyond the scope of this paper.

For any integrable function f on $[0,1]$, we set

$$F(x) = \int_0^x f(t)dt - x\bar{f}, \quad \bar{f} = \int_0^1 f(t)dt.$$

We start with the following identity which we prove below:

$$\begin{aligned} Tf(x) &= \left(v'(x)^{-1}F(v(x))\right)' + \left(\bar{f} - \left(v'(x)^{-1}\right)'F(v(x))\right) \\ &= Vf(x) + Kf(x) \end{aligned} \tag{77}$$

where the prime denotes **in the whole present section** the a.c. part of the distributional derivative. We shall take $E = L_\infty([0,1])$:

$$|\!|\!| f |\!|\!| = \|f\|_\infty.$$

In order to prove (77), note that $v'(x)^{-1}F(v(x))$ is clearly Lipschitz because $F(v(x))$ cancels at the discontinuity point of v', implying that this function as well as its distributional derivative belongs to E with

$$\left(v'(x)^{-1}F(v(x))\right)' = f(v(x)) - \bar{f} + \left(v'(x)^{-1}\right)'F(v(x))$$

which proves (77). We obtain also by induction on n that

$$V^n f(x) = \left(v_n'(x)^{-1}F(v_n(x))\right)' \tag{78}$$

where v_n is the n-th iterate of v. In order to prove this, notice that $v_n'(x)^{-1}F(v_n(x))$ being Lipschitz, it is the integral of its derivative and (78) leads to

$$V^{n+1}f(x) = \Big(v'(x)^{-1}\Big(v_n'(.)^{-1}F(v_n(.))\Big)(v(x))\Big)' = \Big(v_{n+1}'(x)^{-1}F(v_{n+1}(x))\Big)'.$$

On the other hand, it is proved by induction in appendix I that

$$v_n'(x) \geq c_1 n^{1/\gamma} v_n(x) \tag{79}$$

with $c_1 = (2^\gamma - 1)^{1/\gamma}$. Hence, if we consider the norm $\|f\| = \|\int_0^\cdot f(t)dt\|_\infty$, we are led to

$$\begin{aligned}
\|V^n f\| &= \|v_n'(x)^{-1}F(v_n(x))\|_\infty \\
&\leq c_1^{-1} n^{-1/\gamma} \|x^{-1}F(x)\|_\infty \\
&\leq c_1^{-1} n^{-1/\gamma} \sup_{0\leq x\leq 1} x^{-1} \int_0^x |f(y)|dy \\
&\leq c_1^{-1} n^{-1/\gamma} \left|\!\left| f \right|\!\right|.
\end{aligned}$$

Because $B = \{f \in E : \left|\!\left| f \right|\!\right| \leq 1\}$ is $\|.\|$-compact (F is 1-Lipschitz if $f \in B$), the assumptions of Theorem 1 and of Theorem 7 are all satisfied (but here $\|T^n\|$ is not bounded). Thanks to classical distortion arguments (see for instance [19] Theorem 1), one knows that T admits a unique absolutely continuous invariant probability measure, which is ergodic and mixing. In particular, there is no nontrivial eigenfunction for any eigenvalue of modulus 1 and we can conclude that

$$\|T^n f - \nu(f)\| \leq C n^{1-1/\gamma} \left|\!\left| f \right|\!\right|. \tag{80}$$

A Proof of Theorem 1

The proof of Theorem 1 requires two preliminary results which are the subject of the forthcoming section.

A.1 Asymptotically almost periodic powers of an operator

Theorem 10 below gives conditions under which, in some sense, the powers of an operator T can be rewritten

$$T^n = \sum_{i\geq 1} \lambda_i^n P_i + T^n P_0$$

where each P_i is a projection, $P_i P_j = 0$, $i \neq j$, and $T^n P_0$ tends to zero in some sense. However, if each term of the series will be well defined (eigenspace and eigenvalue), the series may fail to converge, as in the case of almost periodic sequences; but since the set of points x for which $P_i x = 0$ except for a finite number of indices i will appear to be dense, the series $\sum_{i\geq 1} \lambda_i^n P_i x$ will converge at least on a dense subspace of E. Lemma 11 will give a condition under which there is only a fimite number of non zero λ_i's.

Let us say a few words concerning Assumptions (B1) and (B2) below, since they are the key assumptions and may appear somehow complicated; it is easily shown that under these assumptions, for any $x \in E$ the sequence $T^n x$ has $\|.\|$-compact closure. These assumptions are essentially used to prove the total boundedness of the sequence $(T^n)_{n>0}$ for a certain norm (Step 1 of the proof of Theorem 10). These assumptions are reminiscent of that of the De Leeuw-Glicksberg theorem [3], but here we consider $\|.\|$-total boundedness rather than $\left|\!\left| . \right|\!\right|$-weak total boundedness (which is actually not a weaker assumption).

For the statement of this theorem, we refer to the equations (19) to (21)

Theorem 10. *Let T be a continuous operator on the Banach space $(E, \left|\!\left| . \right|\!\right|)$ satisfying assumptions (A0), (A1) and*

(B1) *The sequence T^n is uniformly $\|.\|$-equicontinuous on $|.|$-bounded sets in the following sense:*

$$\lim_{x\in B,\|x\|\to 0}\sup_n \|T^n x\| = 0 \tag{81}$$

(B2) *$T^n B$ is asymptotically $\|.\|$-totally bounded in the following sense: There exist a sequence of finite sets $K_n \subset E$, and a sequence $\varepsilon_n \to 0$ such that for any $n \geq 0$*

$$T^n B \subset K_n + \varepsilon_n B_0. \tag{82}$$

Then the following facts hold true: The space E is the direct sum of two $|.|$-closed spaces

$$E = \{x : \|T^n x\| \to 0\} \oplus \{x : \liminf_n \|x - T^n x\| = 0\} = E_0 \oplus E_c. \tag{83}$$

The projection P_c on E_c parallel to E_0 satisfies $|P_c| \leq C_T$. There exist a non-negative sequence ρ_n converging to 0 such that

$$\|T^n x\| \leq \rho_n |x|, \quad x \in E_0, \quad n \geq 0. \tag{84}$$

The space E_u of the finite linear combinations of eigenvectors with eigenvalue of modulus one is $\|.\|$-dense in E_c.

The set Λ of these eigenvalues is at most countable, and for each $\lambda \in \Lambda$ there exists a continuous projection P_λ on the corresponding eigenspace parallel to the others and to E_0. It satisfies $|P_\lambda| \leq C_T$ and

$$\lim_{n\to\infty} \|P_\lambda x - \frac{1}{n}\sum_{i=0}^{n-1} \lambda^{-i} T^i x\| = 0, \quad x \in E. \tag{85}$$

There exists a sequence k_i such that the projection P_c on E_c satisfies

$$\lim_{i\to\infty} \sup_{x\in B} \|P_c x - T^{k_i} x\| = 0. \tag{86}$$

The unit ball of E_c, $B \cap E_c$, is $\|.\|$-totally bounded.

If the integer powers of T extend to a $\|.\|$-C_0-semi-group $(T^t)_{t\geq 0}$, i.e.

$$\forall x \in E, \quad \lim_{t\to 0} \|T^t x - x\| = 0 \tag{87}$$

the space E_c is generated by the vectors x such that for some ω_x, $T^t x = e^{i\omega_x t} x$ for any $t \geq 0$.

Proof. Step 1: The non-negative powers of T form a totally bounded set for the distance

$$d(f,g) = \sup_{|x|\leq 1} \|f(x) - g(x)\|$$

on bounded functions on B. Any limit point of its closure is a continuous operator on $(E, |.|)$, with norm $\leq C_T$.

We start with a simple modification of K_n in order to imbed it in $C_T B$. Fix $n > 0$, denote by $y_k, 1 \leq k \leq N_n$ the points of K_n, choose arbitrary N_n points $x_k \in T^n B$ such that $\|x_k - y_k\| \leq \varepsilon_n$ and define $\tilde{K}_n = \{x_k, 1 \leq k \leq N_n\}$. Assumption (B2) is still satisfied with $\tilde{K}_n$ but ε_n is now two times larger; in addition $\tilde{K}_n \subset C_T B$.

Hence there exist two functions u_n and v_n such that for $|x| \leq 1$

$$T^n x = u_n(x) + v_n(x), \quad u_n(x) \in \tilde{K}_n,$$

and

$$\|v_n(x)\| \leq 2\varepsilon_n, \quad |v_n(x)| \leq 2C_T. \tag{88}$$

Fix n large; for any p:

$$\begin{aligned} T^{2n+p}x &= C_T T^n (C_T^{-1} T^p u_n(x)) + T^{n+p} v_n(x) \\ &= C_T u_n(C_T^{-1} T^p u_n(x)) + C_T v_n(C_T^{-1} T^p u_n(x)) + T^{n+p} v_n(x) \\ &= \alpha_p(x) + \beta_p(x) + \gamma_p(x). \end{aligned}$$

The set of functions $\{\alpha_p(.), p \geq 0\}$ has at most $N_n^{N_n}$ elements; clearly $\|\beta_p(x)\| \leq 2C_T\varepsilon_n$; and Assumption (B1) with Equation (88) implies that $\|\gamma_p(x)\| \leq \eta_n$, for all $p \geq 0$ and some sequence $\eta_n \to 0$. We have just proved that the set $\{T^k, k \geq 2n\}$ can be covered with $N_n^{N_n}$ d-balls of radius $2C_T\varepsilon_n + \eta_n$; hence $\{T^k, k \geq 0\}$ is totally bounded for the distance d.

For any $x \in B$, the sequence $T^n x$ belongs to $C_T B$, hence any $\|.\|$-cluster point of this sequence belongs to $C_T B$ (because of (A0)), and the continuity follows.

Step 2: For any limits $d(T^{u_k}, U) \to 0$ and $d(T^{v_k}, V) \to 0$, one has $d(T^{u_j+v_k}, UV) \to 0$ if $\min(j,k) \to +\infty$. In particular $UV = VU$ and for any third similar limit operator W, $d(WU, WV) \leq C_T d(U,V)$.

One has indeed:

$$\begin{aligned} d(T^{u_j+v_k}, UV) &\leq d(T^{u_j+v_k}, T^{u_j}V) + d(T^{u_j}V, UV) \\ &\leq \sup\{\|T^{u_j}x\| : \|x\| \leq d(T^{v_k}, V), |x| \leq 2C_T\} + d(T^{u_j}, U)|V|. \end{aligned}$$

The second term obviously converges to zero, and the first one also because of Assumption (B1). For the last assertion

$$d(WU, WV) = d(UW, VW) \leq d(U,V)|W|.$$

Step 3: Proof of Equations (83) and (86).

Let n_k be a sequence such that T^{n_k} d-converges to some limit S. We can assume that $n_k - n_{k-1} \to \infty$. From the sequence $n_k - n_{k-1}$ one can extract a sequence $p_i = n_{k_i+1} - n_{k_i}$ such that T^{p_i} and T^{p_i-1} d-converge to some limit P_c and R. Set $m_i = n_{k_i}$.

$$S = d\text{-}\lim T^{m_i+p_i} = SP_c.$$

Since $p_i \to \infty$, there exists $q_i \to \infty$ such that $P_c = d\text{-}\lim T^{m_i+q_i}$ and we get

$$P_c = d\text{-}\lim T^{m_i} T^{q_i} = d\text{-}\lim ST^{q_i} = d\text{-}\lim P_c S T^{q_i} = P_c^2.$$

P_c is a projection on $P_c E$ and Equation (86) holds. We shall prove now that $P_c E$ is indeed E_c and that (83) holds true.

Clearly $P_c E \subset E_c$. On the other hand, for any $x \in E_c$ there exists a sequence r_k such that $\|x - T^{r_k}x\|$. converges to 0. We can assume that $r_k > p_k$ and that $d(T^{r_k-p_k}, U) \to 0$ for some U; in particular $d(T^{r_k}, P_c U) \to 0$. Hence $x = P_c U x \in P_c E$. Finally $P_c E = E_c$. The null space of P_c clearly contains E_0. On the other hand for any point $x \notin E_0$, there exists a sequence r_k such that $\|T^{r_k}x\| \geq \varepsilon$ and $T^{r_k-p_k}$ d-converges to some limit V; the bound $\|VP_c x\| \geq \varepsilon$ leads to $P_c x \neq 0$. This implies by contradiction that any point of the null space of P_c belongs to E_0; hence the null space of P_c is E_0 and $E = E_0 \oplus E_c$.

The bound on the norm of P_c is a consequence of the last point of Step 1.

Step 4: T is one-to-one on E_c. The powers of T on E_c generate a compact G group of operators on E_c with the distance

$$d_c(f,g) = \sup_{|x| \leq 1, x \in E_c} \|f(x) - g(x)\|.$$

Since $TP_c = P_cT$ and $P_c = TR = RT$ (R is defined in Step 3), E_c is T-stable and R is its inverse on E_c. The monoid generated by the powers of T

$$G = \overline{\{T^n, n \geq 0\}}.$$

is a group since we have seen that $R \in G$. The continuity of the multiplication on G comes from Step 2, and the compactness from Step 1.

Step 5: E_u is $\|.\|$-dense in E_c. Properties of P_λ.

Each character χ on G is uniquely determined by the value of $\chi(T)$, because of the definition of G and $\chi(T^n) = \chi(T)^n$.

For any eigenvalue λ of T with modulus 1, there exists a unique character χ such that $\chi(T) = \lambda$ which can be defined as follows: pick an eigenvector x, a $\|.\|$-continuous linear form u such that $u(x) = 1$ and set $\chi(S) = u(Sx)$; χ is indeed a character since it is d_c-continuous with $\chi(T^n) = \chi(T)^n$; in particular since the set of characters of a compact group is at most countable, there is at most a countable number of eigenvalues of modulus one.

In order to show now that for any character χ, $\chi(T)$ is an eigenvalue we proceed as follows. Let μ be the Haar measure on G, consider a character χ on G and define

$$Q_\chi = \int_G \chi(S)^{-1} S \mu(dS). \tag{89}$$

(as a continuous function on G, $f(S) = S$ is the uniform limit of simple functions (by compactness) and this integral is well defined with the usual properties, cf [5] III.2). If x is a $\chi(T)$-eigenvector then the relation $T^n x = \chi(T^n)x$ extends to G as $Sx = \chi(S)x$, and clearly $Q_\chi x = x$.

The invariance of μ implies that for $U \in G$:

$$Q_\chi = \int_G \chi(SU)^{-1} SU \mu(dS) = \chi(U)^{-1} U Q_\chi. \tag{90}$$

In particular, taking $U = T$, for any $x \in E$, $Q_\chi x$ is 0 or an eigenvector with eigenvalue $\chi(T)$. In addition integrating this expression w.r.t. $\mu(dU)$ we get that Q_χ is a projector. If Q_χ is non-zero, Q_χ is thus a projector on the $\chi(T)$-eigenspace. If $Q_\chi = 0$, for any $\|.\|$-continuous linear form u on E and $y \in E$, one has

$$\int_G \chi(S)^{-1} u(Sy) \mu(dS) = 0.$$

The Fourier transform of $S \mapsto u(Sy)$ being 0, this d_c-continuous function is itself 0. Hence $u(Sy) = 0$ for any such u and y and any $S \in G$, which is impossible. Hence Q_χ is non-zero, $\chi(T)$ is an eigenvalue, an Q_χ is a projection whose range is the $\chi(T)$-eigenspace.

In summary, there is a one-to-one correspondence between characters and eigenvalues with modulus one, defined by $\lambda = \chi(T)$, and Q_χ is a projector whose range is exactly the eigenspace.

Since $|S| \leq C_T$ we have $|Q_\chi| \leq C_T$, and since by (89) they commute, Q_χ is a projector parallel to the other eigenspaces.

In order to show that E_u is $\|.\|$-dense in E_c, consider a $\|.\|$-continuous linear form u such that $u(x) = 0$ for any eigenvector x, then for any $y \in E_c$, $S \mapsto u(Sy)$ is d_c-continuous and for any character χ one has

$$\int_G \chi(S)^{-1} u(Sy) \mu(dS) = u(Q_\chi(y)) = 0.$$

The Fourier transform of $S \mapsto u(Sy)$ being 0, this continuous function is itself 0. Hence $u(y) = 0$. E_u is $\|.\|$-dense in E_c. The projection P_λ is finally well defined on E by setting $P_\lambda x = Q_{\chi(T)} x$ if $x \in E_c$ and $P_\lambda x = 0$ if $x \in E_0$.

We now prove (85). This equation holds on E_0 and E_u. Set

$$P_{\lambda,n} = \frac{1}{n}\sum_{i=0}^{n-1} \lambda^{-i} T^i.$$

For any $x \in E_c$ we can pick out $y \in E_u$ such that $\|x-y\| \le \varepsilon$ and get

$$\begin{aligned}\|P_{\lambda,q}x - P_{\lambda,n}x\| &\le \|P_{\lambda,q}(x-y)\| + \|P_{\lambda,n}(x-y)\| + \|P_{\lambda,q}y - P_{\lambda,n}y\| \\ &\le 2\sup_k \|T^k(x-y)\| + \|P_{\lambda,q}y - P_{\lambda,n}y\|.\end{aligned}$$

Since this quantity can be made smaller than 3ε by taking n and q large, this proves that $P_{\lambda,q}x$ is a $\|.\|$-Cauchy sequence, and its limit $P_\lambda x$ satisfies (85). Since for all $x \in E$, $|P_{\lambda,n}x| \le C_T|x|$ and $\|P_{\lambda,n}x - P_\lambda x\| \to 0$, Assumption (A0) implies that $|P_\lambda x| \le C_T|x|$.

Step 6: Equation (84).

Using a sequence p_k such that $d(T^{p_k}, P_c) = \alpha_k \to 0$, we obtain $\|T^{p_k}x\| \le \alpha_k$ for $x \in B \cap E_0$. For $n \ge p_k$ large, one can write $\|T^n x\| \le \|T^{p_k}(T^{n-p_k}x)\| \le C_T\alpha_k$. This implies (84).

Step 7: $B_c = E_c \cap B$ is $\|.\|$-totally bounded.

Using the same sequence p_k we get with (82)

$$B_c \subset (P_c - T^{p_k})B_c + T^{p_k}B_c \subset \alpha_k B_0 + K_{p_k} + \varepsilon_{p_k}B_0.$$

This means that B_c is $\|.\|$-totally bounded.

Step 8: Case of semi-group T^t.

We can carry on Steps 1 to 4 with $t \in \mathbb{R}_+$ instead of $n \in \mathbb{N}$. The group G is now $G = \overline{\{T^s, s \ge 0\}}$. In equation (90) we take $U = T^t$ and we obtain that $y = P_\chi x$ is a vector such that $T^t y = \chi(T^t)y$. In particular if $y \ne 0$ we have $\chi(T^{s+t}) = \chi(T^s)\chi(T^t)$, and on the other hand assumption (87) implies that the function $t \to \|T^t y\|$ is continuous, and so is $t \to \chi(T^t)$; hence $\chi(T^t) = e^{i\omega t}$ for some $\omega \in \mathbb{R}$. □

The following lemma gives a condition for checking that E_c is finite dimensional. This could be checked specifically on examples but we shall see in Theorem 1 that this holds naturally in general situations; in addition, this finite dimensionality assumption is very important in Theorem 7.

Lemma 11. *If in addition to (A0) and (A1), T is $\|.\|$-continuous and satisfies the following assumption*

(B1') *There exists two sequences $\eta_n \to 0$ and $\eta'_{n,p} \to 0$ (as $\min(n,p) \to \infty$), such that for any $n, p > 0$*

$$T^n(B \cap p^{-1}B_0) \subset \eta_n B_0 + \eta'_{n,p}B,$$

then (B1) is also satisfied. If (B2) is also satisfied, then (7) to (11) hold true and

$$\|Q^n x\| \le \rho_n |x|, \quad \rho_n \to 0. \tag{91}$$

Proof. We start with (B1). We have to prove that any sequence x_p of B such that $\|x_p\| \to 0$ satisfies $\sup_{n>0}\|T^n x_p\| \to 0$. Without loss of generality, we can assume that $\|x_p\| \le 1/p$. One has

$$\|T^n x_p\| \le \eta_n + \eta'_{n,p}C_0.$$

Since on the other hand

$$\|T^n x_p\| \le \|T\|^n \|x_p\|$$

we have for any n_0

$$\sup_{n>0} \|T^n x_p\| \leq \max_{n \geq n_0} (\eta_n + \eta'_{n,p} C_0) + \frac{1}{p} \max_{n<n_0} \|T\|^n$$

which can be made arbitrarily small by taking n_0 large first and then by increasing p.

Let us prove now that E_c is finite-dimensional. It suffices to prove that $B_c = E_c \cap B$ is $|.|$-totally bounded; since we already know that $E_c \cap B$ is $\|.\|$-totally bounded, it suffices to prove that $|.|$ and $\|.\|$ induce the same topology on B_c. Notice first that if $x \in E$ and $\|x - x_n\| \to 0$ then

$$|x| \leq \overline{\lim_n} |x_n|$$

because of (A0) (the inequality is obviously true if $|x_n|$ is not bounded). Let $x \in B \cap E_c$. We want to prove that $|x|$ can be made arbitrarily small by taking $\|x\|$ small enough. Consider an integer p such that $\|x\| \leq p^{-1}$. There exists a sequence n_k such that $\|x - T^{n_k} x\|$ tends to zero. Thanks to (B1'), there exist $u_k \in B_0$ and $v_k \in B$ such that

$$T^{n_k} x = \eta_{n_k} u_k + \eta'_{n_k,p} v_k.$$

Since $\|x - T^{n_k} x + \eta_{n_k} u_k\|$ tends to zero, using the previous remark:

$$|x| \leq \overline{\lim_k} |T^{n_k} x - \eta_{n_k} u_k| = \overline{\lim_k} |\eta'_{n_k,p} v_k| \leq \overline{\lim_k} \eta'_{n_k,p}$$

which can be made arbitrarily small by taking p large. Hence $|.|$ and $\|.\|$ are topologically equivalent on E_c and the compactness holds.

Now that E_c is finite dimensional, Equations (7) to (11) and (91) are an immediate rewording of the conclusion of Theorem 10 (notice that ρ_n has changed from equation (84) by a factor $|P_0|$). □

A.2 Proof of Theorem 1

Let us recall the identity (27)

$$T^n = \sum_{i=1}^{n} T^{n-i}(T - V)V^{i-1} + V^n = \sum_{i=1}^{n} T^{n-i} K V^{i-1} + V^n.$$

In particular, Assumption (A1) together with (25) implies that the sequence $|V^n|$ is bounded by a constant C_V, and $KV^n KB$ is $\|.\|$-totally bounded. We set $\alpha_n = |KV^n|$ and $\bar{\alpha}_k = \sum_{i=k}^{\infty} \alpha_i$. Let $x \in E$, for any $0 \leq k \leq n$:

$$\begin{aligned}
|(T^n - V^n)x| &\leq \sum_{i=1}^{n} |T^{n-i} K V^{i-1} x| \\
&\leq C_T \sum_{i=1}^{k} |KV^{i-1}x| + C_T \sum_{i=k+1}^{n} |KV^{i-1}x| \\
&\leq C_T C_K \sum_{i=1}^{k} \|V^{i-1}x\| + C_T \bar{\alpha}_k |x| \\
&\leq c_k \|x\| + C_T \bar{\alpha}_k |x|.
\end{aligned}$$

In particular if $x \in B \cap p^{-1} B_0$ one has

$$|(T^n - V^n)x| \leq \min_{k \leq n} \Big(\frac{c_k}{p} + C_T \bar{\alpha}_k \Big).$$

This implies (B1') where $\eta'_{n,p}$ is the right hand side of the previous equation and $\eta_n = \varepsilon'_n$. We proceed now with (82):

$$\begin{aligned} T^n &= \sum_{i=1}^{n} T^{i-1} K V^{n-i} + V^n \\ &= \sum_{i=1}^{n} (\sum_{j=1}^{i-1} T^{j-1} K V^{i-j-1} + V^{i-1}) K V^{n-i} + V^n \\ &= \sum_{1 \le j < i \le n} T^{j-1} K V^{i-j-1} K V^{n-i} + \sum_{i=1}^{n} V^{i-1} K V^{n-i} + V^n \\ &= A_n + B_n + C_n. \end{aligned}$$

The set $A_n B$ is $\|.\|$-totally bounded; on the other hand

$$C_n B + B_n B \subset (\varepsilon'_n + \sum_{i=1}^{n} \alpha_{n-i} \varepsilon'_{i-1}) B_0.$$

The sum tends to zero as n tends to infinity and this leads finally to (82).

We turn now to the last assertion. If T^k satisfies (B1) and (B2) and T is $|.|$-continuous and $\|.\|$-continuous, clearly T also satisfies (B1) and (B2). Theorem 10 applies to T. Since any eigenvector of T associated with an eigenvalue of modulus one is an eigenvector of T^k associated with an eigenvalue of modulus one, E_c is finite dimensional, and (7) to (11) and (18) hold.

B Proof of Theorem 2

(A0) is clearly satisfied. In addition T is a $|.|$-contraction, and (A1) holds true. Up to a replacement of v with v/c_v, we can assume that $c_v = 1$. Since $T1 = 1$, Equations (30), (31) and (32) imply

$$Vv \le Tv \le v - 1 + c1_{K_0} \tag{92}$$

$$V1 \le 1 - \varepsilon 1_{K_0} \tag{93}$$

for some $c > 0$. Combining these equations, we obtain that the function $\bar{v} = v + c/\varepsilon$ satisfies

$$V\bar{v} \le \bar{v} - 1. \tag{94}$$

Multiplying (94) by V^k and summing, up we obtain

$$V^n \bar{v} + \sum_{k=0}^{n-1} V^k 1 \le \bar{v}. \tag{95}$$

Equation (26) is obvious from (34) and (35) and Equation (25) is a consequence of (95) and (35) since $|KV^n| = \|KV^n 1\|_\infty \le \nu(V^n 1)$. For (24) notice that $V^n 1$ is a decreasing sequence of functions, because $V1 \le 1$, hence:

$$V^n 1 \le \frac{1}{n} \sum_{k=0}^{n-1} V^k 1 \le \frac{\bar{v}}{n}.$$

and (24) holds. It remains to prove the compactness of KT^pK. Notice first that in the assumptions we can replace ν with $\bar{\nu}$ defined as

$$\bar{\nu}(f) = \sum_{i \ge 0} 2^{-i} \nu(T^i f),$$

which makes T continuous on $L_1(\bar{\nu})$ with norm ≤ 2. Second, notice that $\bar{\nu}(\bar{v}) < \infty$, that is (35) still holds. Equation (34) implies that the measure $\mu(g) = \int g(x,y)K(x,dy)\bar{\nu}(dx)$ is absolutely continuous w.r.t. $\bar{\nu}(dx) \otimes \bar{\nu}(dy)$, and let $p(x,y)$ be its density; if g has the form $g(x,y) = h(x)f(y)$, one has

$$\int h(x)f(y)p(x,y)\bar{\nu}(dx)\bar{\nu}(dy) = \int h(x)(Kf)(x)\bar{\nu}(dx)$$

hence one has for any bounded measurable function f and for $\bar{\nu}$-a.e. x,

$$Kf(x) = \int f(y)p(x,y)\bar{\nu}(dy).$$

The function p can be approximated in $L_1(\bar{\nu} \otimes \bar{\nu})$ as

$$p(x,y) = \sum_{i=1}^{n} q_i(x)r_i(y) + \rho(x,y), \qquad \int |\rho(x,y)|\bar{\nu}(dx)\bar{\nu}(dy) < \varepsilon.$$

This finite rank approximation implies that KB is totally bounded in $L_1(\bar{\nu})$. By continuity, the same property holds for T^pKB. Next, Equation (34) implies that KT^pKB is totally bounded in $(E, |.|)$. The assumptions of Theorem 1 are thus satisfied.

To obtain (38), it remains to prove that the space E_c is one dimensional. For this, let n_k be a sequence such that $\lambda_i^{n_k} \to \lambda_i$ for each eigenvalue λ_i with modulus 1, and denote by P_c the projector on E_c parallel to E_0. Then $\|T^{n_k}f - TP_cf\|$ converges to 0 for any $f \in E$. Hence TP_c is a Markov transition operator with the same one-modulus eigenvectors as T. It is compact on E and if there exists more than one eigenvector, one can find two non trivial measurable sets A and B such that $TP_c1_A = 1_B$ ([15] Ch.6 § 3, Th.3.7). Notice now that the function $f = P_c1_A$ satisfies $0 \leq f \leq 1$ and by Jensen's inequality

$$T(f^n) \geq (Tf)^n = 1_B.$$

On the other hand, since $f^n \leq f$, we have $T(f^n) \leq Tf = 1_B$, and we obtain that $T(f^n) = 1_B$ for all $n > 0$; letting n tend to infinity, we get

$$T(1_{f=1}) = 1_B.$$

C Proof of Theorem 3

We begin with the case $q = 1$. Elementary inductions lead to

$$\begin{aligned} |T^nx| &\leq \gamma^n|x| + c\|T^{n-1}x\| + \gamma c\|T^{n-2}x\| + ... + \gamma^{n-1}c\|x\| \\ &\leq \gamma^n|x| + c_n\|x\|. \end{aligned} \tag{96}$$

This may be improved as

$$|T^nx| \leq C_T \min_{k\leq n} |T^kx| \leq C_T \min_{k\leq n} \left(\gamma^k|x| + c_k\|x\|\right).$$

This implies (B1') of Lemma 11 with $\eta_n = 0$ and

$$\eta'_{n,p} = C_T \min_{k\leq n} \left(\gamma^k + \frac{c_k}{p}\right). \tag{97}$$

We have similarly

$$T^nB \subset \gamma^nB + \gamma^{n-1}K_B + \gamma^{n-2}TK_B + \cdots + T^{n-1}K_B$$

and this implies now (B2) in Theorem 10.

It remains to prove that Q (from Eq.(7)) has a spectral radius < 1. Notice that for any $n > 0$, $Q^n = T^{n-1}Q$, this proves that $C_Q = \sup_n \left|Q^n\right|$ is finite. For any $x \in B$ we have from (7), (96) and (18)

$$\left|Q^{n+k}x\right| = \left|T^n Q^k x\right| \leq (C_Q+1)\left|T^n \frac{Q^k x}{\left|Q^k x\right|+1}\right| \leq (C_Q+1)\eta'_{n,p}$$

with

$$p^{-1} = \frac{\|Q^k x\|}{\left|Q^k x\right|+1} \leq \rho_k.$$

By choosing n and k large enough, this ensures that some power of Q is a $|.|$-contraction.

If now $q > 1$, the operator T^q satisfies the assumptions for the case $q = 1$, thus T^q satisfies (A1), (B1') and (B2). Since T is $|.|$ and $\|.\|$-continuous, this clearly implies that T also satisfies these assumptions, by writing $T^n = T^{r+kq}$ with $0 \leq r < q$.

D Proof of Theorem 4

For the proof, we shall change $\|.\|$ into

$$\|f\|' = \sup_x \frac{|f(x)|}{v'(x)}, \quad v'(x) = \frac{v(x)+A}{1+A}$$

for some constant $A \geq 1$ which will be chosen later, and $\left|f\right|$ as

$$\left|f\right|' = \|f\|' + q[f]$$

for a small constant q, and prove that the assumptions of Theorem 3 are fulfilled. Notice that $\|f\| \leq \|f\|'$. For any $f \in E$, by the positivity of T and Equation (44),

$$|Tf(x)| \leq \|f\|' T v'(x) \leq \|f\|' \frac{\gamma_v v(x) + A + c_v}{1+A} \leq \|f\|' \Big(v'(x) + \frac{c_v}{1+A} \Big) \tag{98}$$

hence

$$\|Tf\|' \leq \Big(1 + \frac{c_v}{A}\Big)\|f\|'. \tag{99}$$

T is $\|.\|'$-continuous. In addition, Equation (44) implies that for any $n > 0$

$$T^n v(x) \leq \gamma_v^n v(x) + \frac{c_v}{1-\gamma_v} \tag{100}$$

hence $\|T^n\|'$ is bounded. Equation (43) with (99) implies (40) with $\gamma = \gamma_b$, and $c = 1 + c_d/A$. With (100), it implies also that $\left|T^n\right|'$ is bounded. Thus (A0), (A1) and (40) are satisfied.

In order to prove that Theorem 3 applies, it remains to prove that (39) holds true. Each S_x contains an open neighbourhood O_x of x. Consider $A_0 > 0$ which will be chosen large enough later; if $v(x) \leq A_0$ the set $O'_x = O_x \cap \{v < 2A_0\}$ is still an open neighbourhood of x because v is continuous. Consider a finite sequence $(x_i)_{1 \leq i \leq I}$ such that $v(x_i) \leq A_0$ and $\{v \leq A_0\} \subset \cup_{i=1}^I O'_{x_i}$. This is possible thanks to the compactness of $\{v \leq A_0\}$. There exist $\theta_1(x), \theta_2(x), \ldots, \theta_{I+1}(x)$ a locally Lipschitz partition of the unity of S such that the support of each θ_i, $i \leq I$, is contained in O'_{x_i}, and the support of θ_{I+1} is contained in $\{x : v(x) > A_0\}$ (see [1] Th.2 p.10). We define $\varphi = 1 - \theta_{I+1}$ which is 0 on $\{v \geq 2A_0\}$ and 1 on $\{v \leq A_0\}$. We split Tf as

$$\begin{aligned} Tf(x) &= \left(\sum_{i=1}^I \{Tf(x) - \varepsilon\varphi(x) K_{x_i} f(x_i)\}\, \theta_i(x) + Tf(x)\theta_{I+1}(x) \right) + \varepsilon\varphi(x) \sum_{i=1}^I K_{x_i} f(x_i)\theta_i(x) \\ &= Vf(x) + Sf(x) \end{aligned}$$

Clearly, for $\mathbf{|}f\mathbf{|} \le 1$, Sf belongs to a fixed $\|.\|$-compact set because the sum is finite. We are going to show that

$$\mathbf{|}Vf\mathbf{|} \le \gamma_2 \mathbf{|}f\mathbf{|} \tag{101}$$

for some $\gamma_2 < 1$; this will imply (39). Notice that $f \mapsto Vf$ is not linear, but this does not matter. One has

$$\begin{aligned}
[Vf] &\le [Tf] + \varepsilon \sum_i |K_{x_i} f(x_i)|\, [\varphi\theta_i] \\
&\le \gamma_b[f] + \varepsilon\|f\| \sum_i (\gamma_v v(x_i) + c_v)\, [\varphi\theta_i] \\
&\le \gamma_b[f] + \varepsilon c_0 \|f\|', \qquad c_0 = (A_0 + c_v) \sum_i [\varphi\theta_i].
\end{aligned} \tag{102}$$

It is more complicated to bound $\|Vf\|'$. For $i \le I$ and $\theta_i(x) > 0$ then $x \in S_{x_i}$ and Equation (98) and (47) imply that

$$\begin{aligned}
\Big|Tf(x) - \varepsilon\varphi(x) K_{x_i} f(x_i)\Big| &= \Big|(1 - \varepsilon\varphi(x)) Tf(x)\Big| + \Big|\varepsilon\varphi(x)(Tf(x) - K_{x_i} f(x))\Big| + \Big|\varepsilon\varphi(x)(K_{x_i} f(x) - K_{x_i} f(x_i))\Big| \\
&\le \big(1 - \varepsilon\varphi(x)\big)\big(\gamma_v v(x) + c_v + A\big) \frac{\|f\|'}{1 + A} + \varepsilon\varphi(x)(T - K_{x_i}) v'(x) \|f\|' + \varepsilon c_1 [f]
\end{aligned}$$

where c_1 is the maximum of c_d on $\{\varphi > 0\}$. Since $\varphi(x) > 0$ implies $v(x) \le 2A_0$, if we denote by γ_0 the maximum of $1 - \varepsilon_d$ on $\{v \le 2A_0\}$ the second term can be bounded as

$$(T - K_{x_i}) v'(x) \le \frac{Tv(x) + A - K_{x_i}(v + A)(x)}{1 + A} \le \frac{\gamma_v v(x) + c_v + \gamma_0 A}{1 + A} \le \gamma_d v'(x)$$

with $\gamma_d = \max(\gamma_v, \gamma_0 + c_v/A)$. Notice that $\gamma_d < 1$ as soon as $A > c_v/(1 - \gamma_0)$. Our bound becomes

$$\Big|Tf(x) - \varepsilon\varphi(x) K_{x_i} f(x_i)\Big| \le (1 - \varepsilon\varphi(x))(\gamma_v v(x) + c_v + A) \frac{\|f\|'}{1 + A} + \varepsilon\gamma_d v'(x) \|f\|' + \varepsilon c_1 [f]$$

If in this expression, $\varphi(x) < 1$, then $v(x) \ge A_0$ and

$$\begin{aligned}
(1 - \varepsilon\varphi(x))(\gamma_v v(x) + c_v + A) &\le \gamma_v v(x) + c_v + A \\
&\le \Big(\sup_{u \ge A_0} \frac{\gamma_v u + c_v + A}{u + A}\Big)\, (v(x) + A) \\
&= \frac{\gamma_v A_0 + c_v + A}{A_0 + A} (v(x) + A)
\end{aligned}$$

and if $\varphi(x) = 1$:

$$\begin{aligned}
(1 - \varepsilon\varphi(x))(\gamma_v v(x) + c_v + A) &\le (1 - \varepsilon)\Big(\sup_{u \ge 0} \frac{\gamma_v u + c_v + A}{u + A}\Big)\, (v(x) + A) \\
&= (1 - \varepsilon)\Big(\frac{c_v}{A} + 1\Big)(v(x) + A).
\end{aligned}$$

In any case, we get

$$|Tf(x) - \varepsilon\varphi(x) K_{x_i} f(x_i)| \le \gamma_1 \|f\|' v'(x) + \varepsilon\gamma_d \|f\|' v'(x) + \varepsilon c_1 [f] \tag{103}$$

with

$$\gamma_1 = \max\left(\frac{\gamma_v A_0 + c_v + A}{A_0 + A}, (1 - \varepsilon)\Big(1 + \frac{c_v}{A}\Big)\right).$$

In order to bound the factor of θ_{I+1} in the expression of Vf, we notice that in the case where $\theta_{I+1}(x) > 0$, we have $v(x) \geq A_0$ and

$$\begin{aligned}
|Tf(x)| &\leq \frac{\gamma_v v(x) + c_v + A}{1 + A} \|f\|' \\
&= \Big(\gamma_v v'(x) + \frac{c_v + (1-\gamma_v)A}{1+A}\Big)\|f\|' \\
&\leq \Big(\gamma_v + \frac{c_v + (1-\gamma_v)A}{A_0 + A}\Big)\|f\|' v'(x) \\
&\leq \frac{c_v + A + \gamma_v A_0}{A_0 + A} \|f\|' v'(x) \\
&\leq \gamma_1 \|f\|' v'(x). \qquad (104)
\end{aligned}$$

Since (103) is true if $\theta_i(x) > 0$, and (104) holds if $\theta_{I+1}(x) > 0$, we obtain for all x

$$|Vf(x)| \leq \gamma_1 v'(x)\|f\|' + \varepsilon\gamma_d \|f\|' v'(x) + \varepsilon c_1[f]$$

thus

$$\|Vf\|' \leq (\gamma_1 + \varepsilon\gamma_d)\|f\|' + \varepsilon c_1[f] \qquad (105)$$

and combining (105) and (102) leads to

$$\|Vf\|' + q[Vf] \leq \Big(\gamma_1 + \varepsilon\gamma_d + \varepsilon q c_0\Big)\|f\|' + \Big(q\gamma_b + \varepsilon c_1\Big)[f]. \qquad (106)$$

In order to get (101) for some $\gamma_2 < 1$, we need simultaneously:

$$\begin{aligned}
&\frac{\gamma_v A_0 + c_v + A}{A_0 + A} + \varepsilon\gamma_d + \varepsilon q c_0 < 1 \\
&1 + \frac{c_v}{A} - \varepsilon\Big(1 + \frac{c_v}{A} - \gamma_d - c_0 q\Big) < 1 \\
&\gamma_b + \varepsilon\frac{c_1}{q} < 1.
\end{aligned}$$

In other words, it suffices that

$$\begin{aligned}
&\varepsilon\Big(\gamma_d + c_0 q\Big) < \frac{A_0 - \gamma_v A_0 - c_v}{A_0 + A} \\
&c_v < \varepsilon A\Big(1 - \gamma_d - c_0 q\Big) \\
&\varepsilon\frac{c_1}{q} < 1 - \gamma_b.
\end{aligned}$$

Remember that γ_d is a function of A and A_0, actually

$$1 - \gamma_d = \min\left(1 - \gamma_v, \min_{v \leq 2A_0} \varepsilon_d(x) - c_v/A\right).$$

For some large value of A_0, we set $A = A_0^2, \varepsilon = (1-\gamma_v)/(2A_0), q = (1-\gamma_d)/2c_0$. The above conditions become

$$\begin{aligned}
&\frac{(1-\gamma_v)(1+\gamma_d)}{4} < \frac{A_0 - \gamma_v A_0 - c_v}{1 + A_0} \\
&c_v < \tfrac{1}{4}(1-\gamma_v)A_0(1-\gamma_d) \\
&\frac{(1-\gamma_v)c_1 c_0}{A_0(1-\gamma_d)} < 1 - \gamma_b.
\end{aligned}$$

Since by assumption (48) and (49), $(1-\gamma_d)A_0$ tends to infinity with A_0 these equation are satisfied for A_0 large enough, as well as the condition $A \geq c_v/(1-\gamma_0)$ that has been required before.

It remains to prove the last assertion. Since 1 is the only eigenvalue of modulus one and since its multiplicity is one, there exists a linear form π on E such that (50) holds. This equation implies that for $f \in E$

$$\|\pi(f)1 - T^n f\| \leq C\rho^n |\!|\!| f |\!|\!|$$

hence

$$|\pi(f)| \ \|1\| \leq \sup_k \|T^k\| \ \|f\| + C\rho^n |\!|\!| f |\!|\!|.$$

Now we can let n tend to infinity and conclude that π is $\|.\|$-continuous. This $\|.\|$-continuous linear functional defined on the set of compactly supported Lipschitz functions extends to a positive functional on $C_c(S)$, the set of all compactly supported functions on S. By the Riesz Theorem, there exists a Borel measure μ such that $\pi(f) = \mu(f)$ for any $f \in C_c(S)$; since v is the increasing limit of a sequence of functions of $C_c(S)$, we have $\pi(v) = \mu(v) < \infty$. Any f in E being the $\|.\|$-limit of compactly supported Lipschitz functions, by $\|.\|$-continuity of π we obtain that $\pi(f) = \mu(f)$, $f \in E$.

E Proof of Theorem 7

Multiplying both sides of (27) by P_0 on the left and by Q^q on the right we get

$$Q^{n+q} = \sum_{i=1}^{n-1} Q^{n-i} K V^{i-1} Q^q + P_0 K V^{n-1} Q^q + P_0 V^n Q^q. \tag{107}$$

We consider first the simpler case when $\|T^n\|$ is bounded, say $\|T^n\| \leq c$. In this case, considering a sequence n_k such that $\lambda_i^{n_k}$ converges to 1, for $i = 1, \dots p$ (this can be done by considering a converging subsequence λ^{m_k} of $\lambda^m = (\lambda_i^m, \dots \lambda_p^m)$ and taking $n_k = m_{2k} - m_k$) Equation (13) implies that for any $x \in E$

$$\Big\| \sum_{i=1}^{p} \lambda_i^{n_k} P_i x \Big\| \leq \|T^{n_k} x\| + \|Q^{n_k} x\|$$

and letting k tend to infinity, thanks to (18):

$$\Big\| \sum_{i=1}^{p} P_i x \Big\| \leq c\|x\|.$$

Hence $\|P_0\| \leq 1 + c$ is finite, and Equation (107) leads directly to

$$\|Q^{n+q}\|_{E0} \leq \sum_{i=1}^{n-1} \|Q^{n-i}\|_{E0} |\!|\!| K V^{i-1} Q^q |\!|\!| + \|P_0\| |\!|\!| Q^q |\!|\!| \|K V^{n-1}\|_{E0} + \|P_0\| |\!|\!| Q^q |\!|\!| \|V^n\|_{E0}$$

We plan to apply Proposition 12 of the Appendix with $u_n = \|Q^n\|_{E0}$ and $\beta_i = |\!|\!| K V^{i-1} Q^q |\!|\!|$ for some q large enough. We remark that (112) is satisfied since

$$|\!|\!| K V^i Q^q |\!|\!| = |\!|\!| K V^i T^q P_0 |\!|\!| \leq \alpha_i C_2 C_T |\!|\!| P_0 |\!|\!| \tag{108}$$

Because of the summability of α_i (a consequence of (R1) and (R3)), and with the help of the Lebesgue Dominated Convergence Theorem, Equation (113) will be satisfied for q large enough if we can prove that for any $i \geq 0$

$$\lim_q |\!|\!| K V^i Q^q |\!|\!| = 0. \tag{109}$$

But this is easily obtained by induction on i since it is true for $i = 0$ and for any $i, q > 0$

$$\begin{aligned}
\left|KV^iQ^q\right| &= \left|KV^{i-1}(T-K)Q^q\right| \\
&\leq \left|KV^{i-1}Q^{q+1}\right| + \left|KV^{i-1}KQ^q\right| \\
&\leq \left|KV^{i-1}Q^{q+1}\right| + \left|KV^{i-1}\right|\left|KQ^q\right|.
\end{aligned}$$

Hence Proposition 12 applies and (67) holds.

If now $\|T^n\|$ is not bounded we have to work slightly more on Equation (107). Consider

$$f(z) = \prod_{i=1}^{p}(1 - z\bar{\lambda}_i).$$

Since Equations (7) to (11) imply that $T^n = \sum \lambda_i^n P_i + P_0 Q^n$, $n \geq 0$ (this differs from (13) because we have to take into account the case $n = 0$) we have $f(T) = P_0 f(Q)$. Hence after multiplication on the left by $f(Q)$ Equation (107) becomes

$$f(Q)Q^{n+q} = \sum_{i=1}^{n-1} f(Q)Q^{n-i}KV^{i-1}Q^q + f(T)KV^{n-1}Q^q + f(T)V^nQ^q$$

thus

$$\|f(Q)Q^{n+q}\|_{E0} \leq \sum_{i=1}^{n-1} \|f(Q)Q^{n-i}\|_{E0}\left|KV^{i-1}Q^q\right| + \|f(T)KV^{n-1}Q^q\|_{E0} + \|f(T)V^nQ^q\|_{E0}. \tag{110}$$

Since $\|f(T)\| < \infty$, (108) implies that there exists a constant C such that

$$\|f(T)KV^{n-1}Q^q\|_{E0} + \|f(T)V^nQ^q\|_{E0} \leq C\alpha_n$$

and we obtain, as before (because (108) and (109) still hold true) that

$$\|f(Q)Q^n\|_{E0} \leq C'\alpha_n.$$

Set $g(z) = 1/f(z) = \sum_{i\geq 0} g_i z^i$. The partial fraction decomposition of g implies that $\sup_i |g_i| < \infty$. For any $n \geq 0$

$$\|Q^n\|_{E0} \leq \|Q^n g(Q)f(Q)\|_{E0} \leq \sum_k \|Q^{n+k}g_k f(Q)\|_{E0} \leq \sup_i |g_i| \sum_k \|Q^{n+k}f(Q)\|_{E0}$$

hence

$$\|Q^n\|_{E0} \leq C\sum_{k\geq n} \alpha_k.$$

F Convolution of sequences

Proposition 12. *Let $(\alpha_n)_{n\geq 1}$ be a positive sequence satisfying Assumptions (R1) to (R3) of Theorem 7, and $(\beta_i)_{i\geq 1}$ be a non-negative sequence. Let q be a non-negative integer and $(u_n)_{n\geq 1}$ be a non-negative sequence such that*

$$u_{n+q} \leq C_0\alpha_n + \sum_{i=1}^{n-1} u_{n-i}\beta_i, \quad n \geq 1 \tag{111}$$

for some $C_0 > 0$. If

$$\sup_k \frac{\beta_k}{\alpha_k} < \infty, \tag{112}$$

$$\sum_{i=1}^{\infty} \beta_i < 1 \tag{113}$$

then

$$\sup_n \frac{u_n}{\alpha_n} < \infty. \tag{114}$$

Proof. Set

$$v_n = \frac{u_n}{\alpha_n}$$
$$v_n^* = \sup_{k \le n} v_k$$
$$\theta_n = \frac{\alpha_n}{\alpha_{n+q}}$$
$$C_\beta = \sup_k \frac{\beta_k}{\alpha_k}$$

then, for any i_0 and $n > i_0$

$$\begin{aligned} v_{n+q} &\le C_0\theta_n + \theta_n \sum_{i=1}^{n-1} v_{n-i} \frac{\alpha_{n-i}\beta_i}{\alpha_n} \\ &\le C_0\theta_n + \theta_n v_{i_0}^* \sum_{i=n-i_0}^{n-1} \frac{\alpha_{n-i}\beta_i}{\alpha_n} + \theta_n v_n^* \sum_{i=i_0}^{n-i_0} \frac{\alpha_{n-i}\beta_i}{\alpha_n} + \theta_n v_n^* \sum_{i=1}^{i_0} \frac{\alpha_{n-i}\beta_i}{\alpha_n} \\ &\le C_0\theta_n + \theta_n v_{i_0}^* C_\beta \sum_{i=n-i_0}^{n-1} \frac{\alpha_{n-i}\alpha_i}{\alpha_n} + \theta_n v_n^* C_\beta \sum_{i=i_0}^{n-i_0} \frac{\alpha_{n-i}\alpha_i}{\alpha_n} + \theta_n v_n^* \frac{\alpha_{n-i_0}}{\alpha_n} \sum_{i=1}^{i_0} \beta_i \\ &\le C_0\theta_n + \theta_n v_{i_0}^* C_\beta i_0 \frac{\alpha_1\alpha_{n-i_0}}{\alpha_n} + \theta'_n v_n^* \Big(C_\beta \sum_{i=i_0}^{n-i_0} \frac{\alpha_{n-i}\alpha_i}{\alpha_n} + \sum_{i=1}^{i_0} \beta_i \Big) \end{aligned}$$

where θ'_n tends to 1 (Assumption (R2)). By assumption (R2), for any i, the sequence $j \mapsto \alpha_{j-i}/\alpha_j$, $j \ge i$ is decreasing, hence for $i \le n/2$ one has

$$\frac{\alpha_{n-i}}{\alpha_n} \le \frac{\alpha_i}{\alpha_{2i}}$$

thus for $1 \le i_0 < n$

$$\sum_{i=i_0}^{n-i_0} \frac{\alpha_{n-i}\alpha_i}{\alpha_n} \le 2 \sum_{i=i_0}^{[n/2]} \frac{\alpha_{n-i}\alpha_i}{\alpha_n} \le 2 \sum_{i=i_0}^{[n/2]} \frac{\alpha_i^2}{\alpha_{2i}} \le 2 \sum_{i=i_0}^{\infty} \frac{\alpha_i^2}{\alpha_{2i}}$$

and we get, for $n > i_0$

$$\begin{aligned} v_{n+q} &\le C' + \theta'_n \rho v_n^* \\ \rho &= 2\Big(\sum_{i=i_0}^{\infty} \frac{\alpha_i^2}{\alpha_{2i}} \Big) \sup_k \frac{\beta_k}{\alpha_k} + \sum_{i=1}^{i_0} \beta_i \end{aligned}$$

where C' depends on everything except on n. Since $\theta'_n \to 1$ and i_0 can be chosen large enough to have $\rho < 1$, this proves that for some $n_0 > 0$ and $0 < \rho' < 1$

$$v_{n+q} \leq C' + \rho' v^*_n, \quad n \geq n_0$$

In particular

$$v_{n+q} \leq C' + \rho' v^*_{n+q}, \quad n \geq n_0.$$

By increasing C' we even get

$$v_n \leq C'' + \rho' v^*_n, \quad n \geq 1$$

and since the r.h.s. is also an upper bound for v_k, $k \leq n$ (because $v^*_k \leq v^*_n$), we get

$$v^*_n \leq C'' + \rho' v^*_n, \quad n \geq 1$$

which proves that v_n is bounded.

□

G Proof of Lemma 8

The equation

$$\psi(\zeta(x)) = \psi(x) + 1$$

implies that $\zeta(x) > x$. By differentiating this equation, we get

$$\zeta'(x) = \frac{\theta(\zeta(x))}{\theta(x)} \tag{115}$$

and

$$\zeta''(x) = \frac{\theta'(\zeta(x))\zeta'(x)\theta(x) - \theta(\zeta(x))\theta'(x)}{\theta(x)^2} = \frac{\theta(\zeta(x))}{\theta(x)^2}(\theta'(\zeta(x)) - \theta'(x)) \leq 0.$$

We turn now Equation (71); since ψ is strictly increasing, (71) is equivalent to

$$\psi(x - \theta(x)) + 1 \leq \psi(x)$$

but since θ is non-decreasing

$$\psi(x) - \psi(x - \theta(x)) = \int_{x-\theta(x)}^{x} \frac{1}{\theta(y)} dy \geq \theta(x)\frac{1}{\theta(x)} = 1.$$

H Proof of Theorem 9

We plan to apply Theorem 7 with

$$|\!|\!| f |\!|\!| = \|f\|_\infty$$
$$\|f\| = \|f\|_v.$$

Clearly, since Theorem 2 applies, Equations (7) to (11) and (18) are satisfied. As in the proof of Theorem 2 we set $V = T - K$; we recall that $K(x, S) = 0$ if $x \notin K_0$ (cf. the statement of Theorem 2). We have to estimate $\|V^n\|_{E0}$. Recall that we have

$$Tv \leq v - \theta(v) + c1_{K_0}$$

for some $c > 0$, and Equation (32) with the fact that $K(x, S) = 0$ for $x \notin K_0$ imply that

$$V1 \leq 1 - \varepsilon 1_{K_0}.$$

Combining these equations we get

$$Tv \leq v - \theta(v) + \lambda(1 - V1), \quad \lambda = \frac{c}{\varepsilon}.$$

We define the functions ζ and ψ from θ as in Lemma 8 and we set for $x \geq 0$

$$\zeta_n(x) = \psi^{(-1)}(\psi(x) + n) = \zeta(\zeta_{n-1}(x)). \tag{116}$$

The function ζ_n is concave, as a composition of increasing concave functions. Using the Jensen inequality and the concavity of ζ_k we obtain

$$\begin{aligned} T(\zeta_k(v)) &\leq \zeta_k(Tv) \\ &\leq \zeta_k(v - \theta(v) + \lambda - \lambda V1) \\ &\leq \zeta_k(v - \theta(v)) + \lambda \zeta_k'(v - \theta(v))(1 - V1) \\ &\leq \zeta_{k-1}(v) + \lambda \zeta_k'(\theta_0)(1 - V1), \quad \theta_0 = \inf_{x \geq \theta_1} x - \theta(x) \end{aligned}$$

the last inequality coming from the fact that ζ_k' is decreasing. Thus, since $v > \theta_0$,

$$\begin{aligned} V\zeta_k(v) &\leq T\zeta_k(v) - \zeta_k(\theta_0)K1 \\ &\leq \zeta_{k-1}(v) - (\zeta_k(\theta_0) - \lambda \zeta_k'(\theta_0))(1 - V1). \end{aligned} \tag{117}$$

Differentiating (116) and using (115), we obtain

$$\zeta_n'(x) = \zeta'(\zeta_{n-1}(x))\, \zeta_{n-1}'(x) = \frac{\theta(\zeta_n(x))}{\theta(\zeta_{n-1}(x))} \zeta_{n-1}'(x)$$

hence

$$\zeta_n'(\theta_0) = \frac{\theta(\zeta_n(\theta_0))}{\theta(\theta_0)}.$$

Since in addition $\zeta_n(\theta_0)$ tends to infinity and $\theta(x)/x$ tends to zero (θ is concave with a derivative which tends to zero), the sequence $\zeta_n'(\theta_0)/\zeta_n(\theta_0)$ tends to 0. As a consequence, there exist n_0 such that $\lambda \zeta_k'(\theta_0) - \zeta_k(\theta_0) \leq 0$ for $k > n_0$, hence multiplying both sides of (117) by V^{k-1} and summing up from 1 to $n > n_0$, we get

$$V^n \zeta_n(v) \leq v + c$$

for some constant c. Since $\zeta_n(x) \geq \psi^{(-1)}(n)$ we get

$$V^n 1 \leq \frac{v + c}{\psi^{(-1)}(n)}.$$

Theorem 7 applies and, in particular, we obtain (73).

For (74), we consider

$$\|f\| = \pi(|f|).$$

Since $|.|$ is unchanged, Equations (7) to (11), (65) and (66) are still satisfied, as well as (18) because $\pi(|f|) \leq \|f\|_v \pi(v)$. In addition $\|T^n\| = 1$, and

$$\|V^n\|_{E0} = \pi(V^n 1) \leq \frac{c'}{\psi^{(-1)}(n)}.$$

Theorem 7 still applies and we obtain (65).

I Proof of Equation (79)

We shall prove that for $0 \le x < 1$

$$v_n'(x)^\gamma \ge 1 + anv_n(x)^\gamma, \quad a = 2^\gamma - 1. \tag{118}$$

We recall that

$$v(x) = \begin{cases} x(1+2^\gamma x^\gamma) & 0 \le x < 1/2 \\ 2x - 1 & 1/2 \le x \le 1 \end{cases} \tag{119}$$

and that the prime sign stands for the right derivative. In the case $n = 0$ the inequality is obvious. In the case $n \ge 1$, we assume by induction that (118) is satisfied and since $v_{n+1}'(x) = v_n'(x)v'(v_n(x)))$, valid for $n \ge 0$, Equation (118) with $n+1$ will be implied by

$$(1 + anv_n(x)^\gamma)v'(v_n(x))^\gamma \ge 1 + a(n+1)v_{n+1}(x)^\gamma.$$

This has to be proved for $n \ge 0$. It suffices to show that for any $0 \le y \le 1$

$$(1 + any^\gamma)v'(y)^\gamma \ge 1 + a(n+1)v(y)^\gamma \tag{120}$$

(i.e. $y = v_n(x)$). By linearity of both sides of (120) w.r.t. n, we only have to check this for $n = 0$, and $n \to \infty$, that is

$$\begin{cases} v'(y)^\gamma \ge 1 + av(y)^\gamma \\ yv'(y) \ge v(y) \end{cases} \tag{121}$$

(the first equation is (118) with $n = 1$). In the case $y < 1/2$ this is rewritten as

$$\begin{cases} (1 + (\gamma+1)2^\gamma y^\gamma)^\gamma \ge 1 + ay^\gamma(1 + 2^\gamma y^\gamma)^\gamma \\ 1 + (\gamma+1)2^\gamma y^\gamma \ge 1 + 2^\gamma y^\gamma. \end{cases}$$

The second inequality is obvious. For the first one, since $2y < 1$, setting $z = 2^\gamma y^\gamma$, this holds if

$$(1 + (\gamma+1)z)^\gamma \ge 1 + az$$

for $0 \le z \le 1$. Since the difference of both sides is a concave function of z which vanishes at $z = 0$, and is non negative at $z = 1$ (we recall that $a = 2^\gamma - 1$), the inequality is satisfied. In the case $y \ge 1/2$, (121) is

$$\begin{cases} 2^\gamma \ge 1 + a(2y-1)^\gamma \\ 2y \ge 2y - 1 \end{cases}$$

which is obviously satisfied.

References

[1] Jean-Pierre Aubin and Arrigo Cellina, *Differential inclusions*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 264, Springer-Verlag, Set-valued maps and viability theory.

[2] B. Cloez and M. Hairer, *Exponential ergodicity for markov processes with random switching*, arXiv:1303.6999v1 (2013).

[3] K. de Leeuw and I. Glicksberg, *Applications of almost periodic compactifications*, Acta Math. **105** (1961), 63–97.

[4] Wolfgang Doeblin and Robert Fortet, *Sur des chaînes à liaisons complètes*, Bull. Soc. Math. France **65** (1937), 132–148.

[5] Nelson Dunford and Jacob T. Schwartz, *Linear operators. Part I*, Wiley Classics Library, John Wiley & Sons Inc., 1988, General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[6] Sébastien Gouëzel, *Central limit theorem and stable laws for intermittent maps*, Probab. Theory Related Fields **128** (2004), no. 1, 82–122.

[7] Hubert Hennion and Loïc Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Lecture Notes in Mathematics, vol. 1766, Springer-Verlag, Berlin, 2001.

[8] C. T. Ionescu Tulcea and G. Marinescu, *Théorie ergodique pour des classes d'opérations non complètement continues*, Ann. of Math. (2) **52** (1950), 140–147. MR 0037469 (12,266g)

[9] M. Iosifescu and R. Theodorescu, *Random processes and learning*, Springer-Verlag, New York, 1969, Die Grundlehren der mathematischen Wissenschaften, Band 150.

[10] Marius Iosifescu, *A basic tool in mathematical chaos theory: Doeblin and Fortet's ergodic theorem and Ionescu Tulcea and Marinescu's generalization*, Doeblin and modern probability (Blaubeuren, 1991), Contemp. Math., vol. 149, Amer. Math. Soc., Providence, RI, 1993, pp. 111–124.

[11] Søren F. Jarner and Gareth O. Roberts, *Polynomial convergence rates of Markov chains*, Ann. Appl. Probab. **12** (2002), no. 1, 224–247.

[12] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory Dynam. Systems **19** (1999), no. 3, 671–685.

[13] Sean Meyn and Richard L. Tweedie, *Markov chains and stochastic stability*, Cambridge University Press, 1994, With a prologue by Peter W. Glynn.

[14] Esa Nummelin, *General irreducible Markov chains and nonnegative operators*, Cambridge Tracts in Mathematics, vol. 83, Cambridge University Press, 1984.

[15] D. Revuz, *Markov chains*, North-Holland Publishing Co., Amsterdam, 1975, North-Holland Mathematical Library, Vol. 11.

[16] Omri Sarig, *Subexponential decay of correlations*, Invent. Math. **150** (2002), no. 3, 629–653.

[17] Cédric Villani, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, 2009, Old and new.

[18] Kôsaku Yosida and Shizuo Kakutani, *Operator-theoretical treatment of Markoff's process and mean ergodic theorem*, Ann. of Math. (2) **42** (1941), 188–228.

[19] Lai-Sang Young, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999).