

Invariant probabilities for discrete time linear dynamics via thermodynamic formalism

Artur O. Lopes*, Ali Messaoudi†, Manuel Stadlbauer‡ and Victor Vargas§.

March 30, 2022

Abstract: We show the existence of invariant ergodic σ -additive probability measures with full support on X for a class of linear operators $L : X \rightarrow X$, where L is a weighted shift operator and X either is the Banach space $c_0(\mathbb{R})$ or $l^p(\mathbb{R})$ for $1 \leq p < \infty$. In order to do so, we adapt ideas from Thermodynamic Formalism as follows. For a given bounded Hölder continuous potential $A : X \rightarrow \mathbb{R}$, we define a transfer operator \mathcal{L}_A which acts on continuous functions on X and prove that this operator satisfies a Ruelle-Perron-Frobenius theorem. That is, we show the existence of an eigenfunction for \mathcal{L}_A which provides us with a normalized potential \bar{A} and an action of the dual operator \mathcal{L}_A^* on the 1-Wasserstein space of probabilities on X with a unique fixed point, to which we refer to as Gibbs probability. It is worth noting that the definition of \mathcal{L}_A requires an *a priori* probability on the kernel of L . These results are extended to a wide class of operators with a non-trivial kernel defined on separable Banach spaces.

Keywords: Gibbs states, linear dynamics, Ruelle operator, weighted shifts.

Mathematics Subject Classification (2010): 28D05, 37A25, 47A35

1 Introduction

The main goal in here is to show existence of invariant σ -additive probabilities for linear dynamical systems in discrete time for a certain class of linear operators for which the formalism of the Ruelle operator is applicable. That is, we obtain our main results by adapting the formalism previously considered for non-compact generalized XY models where the dynamics are given by the shift (see [20], [21] and [9]). In here, the dynamics are obtained from a linear

*IME - UFRGS. Partially supported by CNPq

†MAT - UNESP. Partially supported by CNPq, project 311018/2018-1, and FAPESP, project 2013/24541-0

‡IM - UFRJ. Partially Supported by CNPq, project 312632/2018-5

§IME - UFRGS. Supported by INCTMat.

operator L but as in the case of the XY -model, we require the choice of an *a priori* probability.

We point out that the existence of invariant probabilities for $l^p(\mathbb{R})$, $1 < p < \infty$, already is known, but our methods are quite different from the standard approach. We believe the connection with thermodynamic formalism is interesting in itself, because - among other reasons - allows an extension of the previously known results beyond the case of linear operators defined on reflexive Banach spaces.

The use of the Ruelle operator (also called transfer operator) as a tool for the construction of Gibbs measures - which are, in particular, ergodic invariant measures with full support - is a usual approach in the area of thermodynamic formalism. Moreover, there are several works in which these properties are studied for a wide variety of both compact and non-compact dynamical contexts, **even for the case of shifts defined on $l^1(\mathbb{R})$** (see for instance [2, 21, 24, 9, 26]). Thus, it is natural to choose this tool in order to solve an important question in linear dynamics: the existence of probability measures satisfying the above mentioned properties.

Linear dynamics is a relatively young branch of mathematics in which dynamical properties of linear operators defined on Fréchet spaces are studied and, in particular, some interesting phenomena even in the case of Banach spaces appear. In the finite dimensional case, it is widely known that linear dynamical systems are completely characterized by their corresponding Jordan canonical form. However, in infinite dimension, interesting properties such as the existence of dense orbits, chaos in the sense of Devaney and topological equivalence of invariant subsets with any dynamical system defined on a compact metric space might occur. For an interesting example of a linear operator defined on a Hilbert space, see for instance [3, 13].

Let X be a topological vector space and $T : X \rightarrow X$ be a linear continuous operator. We say that (X, T) - or simply the map T - is *uniformly hypercyclic*, if it has a dense orbit in X . In the particular case that X is a separable Banach space, this property is equivalent to say that T is *topologically transitive*, that is, for all non-empty open sets $U, V \subset X$, there is an integer $n \geq 0$, such that, $T^n(U) \cap V \neq \emptyset$. Moreover, we call the map T *frequently hypercyclic*, if for every non-empty open set $V \subset X$, the set $N(x, V) = \{k \in \mathbb{N}, T^k(x) \in V\}$ has positive lower density, i.e. $\liminf_{n \rightarrow \infty} \frac{1}{n} \#(N(x, V) \cap \{1, \dots, n\}) > 0$. On the other hand, we call the map T *Devaney chaotic*, if it is topologically transitive and has a dense subset of periodic points. We say that T is *topologically mixing*, if for all non-empty open subsets $U, V \subset X$, there exists an integer $N > 0$, such that, $T^n(U) \cap V$ is not empty for all $n \geq N$.

A typical example of mixing, frequently hypercyclic and Devaney chaotic operator is αL , $\alpha > 1$, where L is the shift operator acting on the Banach space $X = l^p(\mathbb{R})$, $1 \leq p < \infty$, i.e. $L((x_n)_{n \geq 1}) = (x_{n+1})_{n \geq 1}$. Furthermore, the operator αL , $\alpha > 1$, acting on $c_0(\mathbb{R})$ is mixing, frequently hypercyclic and Devaney chaotic as well. The study of the above properties is a central problem in the area of discrete time Linear Dynamical Systems (see for instance [3], [5], [13]).

In the context of ergodic theory on linear dynamics, there are some results related to the existence of invariant probability measures when the system is frequently hypercyclic and X is a reflexive Banach space (see for example [15]). We point out that in there, the existence of ergodic measures with full support is obtained through approximations of measures supported on dense orbits of the system. Here we prove a similar kind of result (even for non-reflexive Banach spaces as $l^1(\mathbb{R})$ and $c_0(\mathbb{R})$) in the context of weighted shifts using a different technique: we adapt classical tools of thermodynamic formalism to extend this result to a large class of linear operators defined on separable and non-necessarily reflexive Banach spaces.

Let X be one of the Banach spaces $c_0(\mathbb{R})$ or $l^p(\mathbb{R})$, for $1 \leq p < \infty$, where $c_0(\mathbb{R})$ refers to the Banach space of sequences $(x_n)_{n \in \mathbb{N}}$, with $\lim_{n \rightarrow \infty} x_n = 0$ equipped with the sup-norm. For a fixed operator $T : X \rightarrow X$, a Hölder continuous potential $A : X \rightarrow \mathbb{R}$ and a suitable *a priori* probability measure m defined on the kernel of T , we are able to associate a *Gibbs probability measure* for the potential A (see Definition 3 and item c) in Theorem 1) which will be T -invariant, ergodic and has full support.

Our first result is Theorem 1. From this we will obtain a σ -additive ergodic probability measure which is invariant for the weighted shift L and has full support. In Proposition 2 and Corollary 2, we then extend the previous results to a large class of linear operators defined on separable Banach spaces, among them the class of uniform hypercyclic linear operators. Furthermore, these results are satisfied for the class of linear continuous operators with non-trivial kernel defined on separable Banach spaces. An important issue here is the use of ideas of transport theory which are employed in the Appendix (Section 5) in order to prove item b) of Theorem 1, that is the existence of a fixed point of the dual of the Ruelle operator through a contraction argument (for related results, see [27, 6, 17, 18, 19, 23]). However, the arguments here are more complex due to the metric structure of X . For example, in contrast to shift spaces (even with respect to uncountable, non-compact alphabets as in [9]), it seems to be impossible to obtain uniform contraction rates (see Remark 9).

The paper is organized as follows. In Section 2, we consider some definitions and preliminaries. In Section 3, we prove the main result for weighted shifts operators. In Section 4, we consider the general case of frequently hypercyclic operators defined on separable Banach spaces. In the Appendix (Section 5) we show the remaining assertion of Theorem 1 by extending results from transport theory to the setting of linear spaces.

Our thanks to L. Cioletti for many helpful conversations during the writing of our paper. The fourth author also would to thank PNPd-CAPES, INCTMat and the Francisco José de Caldas Fund by the financial support during part of the development of this paper.

2 Preliminaries

Let X be a Banach space and $T : X \rightarrow X$ be a linear continuous operator. Hereafter, with exception of some specific cases, we will denote by $\|\cdot\|_X$ the norm for X . Let $c_0(\mathbb{R})$ be the set of real sequences $(x_n)_{n \geq 1}$ such that $\lim_{n \rightarrow \infty} x_n = 0$. It is widely known that the vector space $c_0(\mathbb{R})$ equipped with the supremum norm

$$\|x\|_{c_0(\mathbb{R})} := \sup_{n \geq 1} |x_n|$$

is a separable, non-reflexive Banach space. On other hand, we also consider $l^p(\mathbb{R})$, for $1 \leq p < \infty$, defined as the set of real sequences $(x_n)_{n \geq 1}$ satisfying $\sum_{n=1}^{\infty} |x_n|^p < \infty$, equipped with the norm

$$\|x\|_{l^p(\mathbb{R})} := \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{\frac{1}{p}} .$$

It is known that $l^p(\mathbb{R})$ is a separable Banach space and that it is reflexive when $1 < p < \infty$. Our results will include the cases of the Banach spaces $l^1(\mathbb{R})$ and $c_0(\mathbb{R})$ which are separable but not reflexive Banach spaces. In the particular case where $p = 2$, the space $l^2(\mathbb{R})$ is a separable Hilbert space when it is equipped with the following inner product

$$\langle x, y \rangle_{l^2(\mathbb{R})} := \sum_{n=1}^{\infty} x_n y_n .$$

Let $\{e_k\}_{k \geq 1}$ be such that each e_k is a vector of the form $e_k = (\delta_{ik})_{i \geq 1}$. Then, any $x \in X \in \{c_0(\mathbb{R}), l^p(\mathbb{R}), 1 \leq p < \infty\}$ can be written as

$$x = \sum_{k=1}^{\infty} x_k e_k ,$$

where the series above converges in the norm $\|\cdot\|_X$, i.e. $\{e_k\}_{k \geq 1}$ is a Schauder basis for X (in particular, $x_k = \langle x, e_k \rangle$ for any $k \in \mathbb{N}$ when $p = 2$).

Fix values $0 < c < c'$ and consider a sequence $(\alpha_n)_{n \geq 1}$ satisfying $\alpha_n \in (c, c')$ for each $n \in \mathbb{N}$. The *weighted shift* associated to the sequence $(\alpha_n)_{n \geq 1}$ is defined as the linear map $L : X \rightarrow X$, where $X \in \{c_0(\mathbb{R}), l^p(\mathbb{R}), 1 \leq p < \infty\}$, and

$$L((x_n)_{n \geq 1}) = (\alpha_n x_{n+1})_{n \geq 1} .$$

Note that $L(e_1) = 0$ and $L(e_n) = \alpha_{n-1} e_{n-1}$, for all $n \geq 2$. Besides that, for each $x \in X$ we have $\|L(x)\|_X \leq c' \|x\|_X$, thus, L is a linear continuous operator and $\text{Ker}(L) = \text{span}\{e_1\}$. Moreover, in the case $c > 1$, the linear operator L is *positively expanding* on $E = \text{span}\{e_n, n \geq 2\}$ with expanding constant equal to c , that is, $\|L(x)\|_X \geq c \|x\|_X$ for all $x \in E$.

We remark that L as above has always a one dimensional kernel but would like to point out that the results we get in the next two sections can be easily

extended to the case where the kernel is finite dimensional. In Section 4 we will consider an even more general case of linear operators with non-trivial kernels defined on a separable Banach space.

From now on, we will use the notation

$$\beta_k^n := \alpha_k \dots \alpha_{k+n-1}, \text{ for all } k, n \in \mathbb{N}$$

and

$$d_n := \inf_{k \geq 1} \beta_k^n, \text{ for all } n \in \mathbb{N}. \quad (1)$$

It is widely known (see for instance [7]), that the spectrum of $L : X \rightarrow X$, when X is $c_0(\mathbb{R})$ or $l^p(\mathbb{R})$, $1 \leq p \leq \infty$, is $\sigma(L) = D(0, r(L))$, where $D(0, r(L))$ is the closed disc of center 0 and radius $r(L)$ and $r(L)$ is the spectral radius of L given by

$$r(L) := \lim_{n \rightarrow \infty} \|L^n\|_{op}^{\frac{1}{n}} = \lim_{n \rightarrow \infty} \left(\sup_{k \geq 1} \beta_k^n \right)^{\frac{1}{n}}, \quad (2)$$

where the operator norm $\|\cdot\|_{op}$ is given by

$$\|L^n\|_{op} := \sup\{\|L^n(x)\|_X, \|x\|_X = 1\}.$$

The following characterization is classical and will be useful in the following sections in order to guarantee the existence of eigenfunctions and invariant probabilities associated to the Ruelle operator (and its corresponding dual) which is the main tool to be used in here.

Remark 1. The asymptotic behavior of β_k^n implies the following for a weighted shift L defined on $c_0(\mathbb{R})$ or $l^p(\mathbb{R})$, for $1 \leq p < \infty$.

1. L is topologically transitive if and only if $\limsup_{n \rightarrow \infty} \beta_1^n = \infty$ ([3], [13]).
2. L is topologically mixing if and only if $\lim_{n \rightarrow \infty} \beta_1^n = \infty$ ([5]).
3. L is frequently hypercyclic in $l^p(\mathbb{R})$ if and only if $\sum_{n=1}^{\infty} (\beta_1^n)^{-p} < \infty$ ([4]).
4. L is Devaney Chaotic if and only if $\sum_{n=1}^{\infty} (\beta_1^n)^{-p} < \infty$ ([13]).

Given a separable Banach space X , we will use the notation $\mathcal{C}(X)$ for the set of continuous functions from X into \mathbb{R} and $\mathcal{C}_b(X)$ for the set of bounded continuous functions from X into \mathbb{R} . As it is well known, $\mathcal{C}_b(X)$ equipped with the uniform norm $\|\cdot\|_{\infty}$ given by $\|\varphi\|_{\infty} := \sup\{|\varphi(x)| : x \in X\}$ is a Banach space.

Given a metric D on X , we will denote by $\text{Lip}(D, X)$ the set of Lipschitz continuous functions from X into \mathbb{R} . That is, the set of continuous functions satisfying

$$\text{Lip}_{\varphi, D} := \sup \left\{ \frac{|\varphi(x) - \varphi(y)|}{D(x, y)} : x \neq y \right\} < \infty.$$

We will use the notation $\text{Lip}_b(D, X)$ for the set of bounded Lipschitz continuous functions from X into \mathbb{R} . Note that for any $0 < \alpha \leq 1$, the set $\text{Lip}(D^\alpha, X)$ is the set of α -Hölder continuous functions from X into \mathbb{R} and the set $\text{Lip}_b(D^\alpha, X)$ is the set of bounded α -Hölder continuous functions from X into \mathbb{R} , where $D^\alpha(x, y) := (D(x, y))^\alpha$.

Besides that, given $\delta > 0$, we will use the notation $\text{Lip}^{loc(\delta)}(D, X)$ for the set of continuous functions from X into \mathbb{R} , such that

$$\text{Lip}_{\varphi, D}^{loc(\delta)} := \sup \left\{ \frac{|\varphi(x) - \varphi(y)|}{D(x, y)} : x \neq y, D(x, y) < \delta \right\} < \infty.$$

In addition, we will denote by $\text{Lip}_b^{loc(\delta)}(D, X)$ the set of bounded functions in $\text{Lip}^{loc(\delta)}(D, X)$ and, from now on, we will denote by D_X the metric generated by the norm $\|\cdot\|_X$, that is,

$$D_X(x, y) = \|x - y\|_X.$$

Observe that the Riesz representation theorem in this setting implies that the dual of $\mathcal{C}_b(X)$, which we will denote by $\mathcal{C}_b(X)^*$, coincides with the set of additive finite Borel signed measures on X , denoted by $\mathcal{B}_A(X)$ in here. In addition, we will use the notation $\mathcal{B}(X)$ for the set of sigma additive measures and $\mathcal{P}(X)$ for the set of Borel sigma additive probability measures on X .

3 A RPF Theorem for weighted shifts

The Ruelle operator is a powerful tool in thermodynamic formalism, especially for Hölder potentials and expanding dynamical systems (in the sense of Ruelle) defined on a compact metric space (see [25]). Results of this nature were also obtained for the case where the number of preimages of the underlying dynamics of each point is uncountable as, for example, in the case of the generalized XY model via the use of an *a priori* probability (see for instance [2, 20, 24, 1]). Even in the case of a non-compact alphabet \mathcal{A} , one can get in specific cases similar results for $\mathcal{A}^{\mathbb{N}}$ (see [21, 11, 9]).

Adapting some ideas from the classical setting (as for instance in [21]) we show in this section the existence of invariant probabilities for the above defined linear dynamical system via the Ruelle operator. After that, in section 4, we will show that our approach can be adapted to the case where the system is given by a class of frequently hypercyclic operators (see [14] and [10]) with a finite dimensional kernel.

We now return to the weighted shift $L : X \rightarrow X$ where X is $c_0(\mathbb{R})$ or $l^p(\mathbb{R})$, for $1 \leq p < \infty$. The analogue of the Ruelle operator on the set $\mathcal{C}_b(X)$ is defined in the following way. We fix an *a priori* probability measure m on the Borel sets of \mathbb{R} equipped with the usual topology and assume that the support of the probability measure m is equal to the set \mathbb{R} . Since $\text{Ker}(L)$ is isometrically isomorphic with \mathbb{R} it follows that m induces an *a priori* probability measure on $\text{Ker}(L)$ - which we will also denote by the same letter m .

As an example, one could take m as the Gaussian distribution on \mathbb{R} with mean zero and variance 1. That is $m = f dr$, with $f(r) = \frac{1}{\sqrt{2\pi}} e^{-r^2/2}$.

We say that a potential $A \in \mathcal{C}(X)$ has *summable variation* if

$$V(A) = \sum_{n=1}^{\infty} V_n(A) < \infty, \quad (3)$$

where $V_n(A) = \sup\{|A(x) - A(y)| : x, y \in X, x_i = y_i, 1 \leq i \leq n\}$. Note that (3) does not imply that A is a bounded potential. We will denote by $\mathcal{SV}(X)$ the set of potentials $A \in \mathcal{C}(X)$ satisfying (3).

Remark 2. We would like to point out that in the setting of linear dynamics in contrast to shift spaces (even with respect to uncountable, non-compact alphabets as in [9]), there are bounded Lipschitz functions which are not in $\mathcal{SV}(X)$. For example, the function $A(x) := \arctan \|x\|$ is an element of $\text{Lip}_b(D_X, X)$, but as $V_n(A) = \pi$ for any $n \in \mathbb{N}$, the function A does not have summable variations. On the other hand, $\text{Lip}_b(D_X^\alpha, X) \cap \mathcal{SV}(X)$ is non-empty which follows from the following construction. For $x, y \in X$, set

$$D_{\text{shift}}(x, y) := \begin{cases} (\sum_n \min(\{1, |x_n - y_n|^p\}) 2^{-n})^{1/p} & : X = l^p(\mathbb{R}) \\ \sup(\{\min(\{1, |x_n - y_n|\}) 2^{-n} : n \in \mathbb{N}\}) & : X = c_0(\mathbb{R}) \end{cases}$$

Note that $D_{\text{shift}}(x, y) \leq D_X(x, y)$ **for any pair** $x, y \in X$ and that, as $D_{\text{shift}}(x, y) \leq 1$ for any $x, y \in X$, each function in $\text{Lip}(D_{\text{shift}}^\alpha, X)$ is bounded. Moreover, it follows from $\sum_n 2^{-n} < \infty$ that any function in $\text{Lip}(D_{\text{shift}}^\alpha, X)$ has summable variations. Hence,

$$\text{Lip}(D_{\text{shift}}^\alpha, X) \subset \text{Lip}_b(D_X^\alpha, X) \cap \mathcal{SV}(X).$$

Finally, as any function in $\text{Lip}_b(D_X^\alpha, X)$ which only depends on a finite number of coordinates of X is an element of $\text{Lip}(D_{\text{shift}}^\alpha, X)$, the set of D_X^α -Lipschitz function with summable variations is non-trivial. These observations also provide an explanation of the phenomena that shift spaces whose alphabet is a standard Borel space show an exponential contraction under the iteration of the transfer operator (see [9]). However, this is not the case in the setting of linear dynamics (see Remark 9). Namely, we obtain contraction to zero with time, but without an explicit decay rate, as a consequence of Theorem 3.

We now define the *Ruelle operator*.

Definition 1. Given the a priori probability m and a potential $A \in \text{Lip}_b(D_X^\alpha, X)$, the Ruelle operator is defined as the map assigning to each function $\varphi \in \mathcal{C}_b(X)$ the function

$$\mathcal{L}_A(\varphi)(x) := \int_{Lv=x} e^{A(v)} \varphi(v) dm(v). \quad (4)$$

Using the coordinates induced by the Schauder basis $\{e_k\}_{k \geq 1}$ it is possible to write the equation (4) as

$$\mathcal{L}_A(\varphi)(x_1, x_2, \dots) := \int_{\mathbb{R}} e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots)} \varphi(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots) dm(r).$$

It is easy to verify that $\mathcal{L}_A(\varphi) \in \mathcal{C}_b(X)$ (see the proof of Lemma 1) and that the n -th iterate ($n \geq 2$) of the Ruelle operator is given by

$$\begin{aligned} \mathcal{L}_A^n(\varphi)(x) &= \int_{\mathbb{R}^n} e^{\sum_{j=1}^n A\left(r_j, \frac{r_{j-1}}{\beta_1^j}, \dots, \frac{r_1}{\beta_1^{j-1}}, \frac{x_1}{\beta_1^j}, \frac{x_2}{\beta_2^j}, \dots\right)} \\ &\quad \times \varphi\left(r_n, \frac{r_{n-1}}{\beta_1^n}, \dots, \frac{r_1}{\beta_1^{n-1}}, \frac{x_1}{\beta_1^n}, \frac{x_2}{\beta_2^n}, \dots\right) dm(r_1) \cdots dm(r_n). \end{aligned}$$

Definition 2. We say that a potential $A : X \rightarrow \mathbb{R}$ is normalized if $\mathcal{L}_A(1) = 1$.

Remark 3. Basic examples for normalized potentials are $A = 0$ or any potential which only depends on the first coordinate, that is $A(r_1, r_2, r_3, \dots) = A(r_1)$ with respect to the coordinates induced by the Schauder basis, and such that $\int e^{A(r_1)} dm(r_1) = 1$. A further important example is given by the following normalization process. Assume that $\mathcal{L}_A(f) = \lambda f$ for some $\lambda > 0$ and $f : X \rightarrow (0, \infty)$. Then, for $\bar{A} := A + \log f - \log f \circ L - \log \lambda$ and $x \in X$,

$$\mathcal{L}_{\bar{A}}(1)(x) = \frac{\mathcal{L}_A(f)(x)}{\lambda f(x)} = \frac{\lambda f(x)}{\lambda f(x)} = 1.$$

Hence, \bar{A} is normalized and \bar{A} is therefore known as the normalization of A . Finally, note that, if A is normalized, then the dynamics of L and the action of \mathcal{L}_A on $\mathcal{C}_b(X)$ are connected through

$$\mathcal{L}_A^n(\varphi \circ L^n) = \varphi, \quad \forall n \in \mathbb{N}.$$

In the next Lemma we will show that for any potential $A \in \text{Lip}_b(D_X^\alpha, X)$, the Ruelle operator \mathcal{L}_A preserves the set of α -Hölder continuous functions.

Lemma 1. Consider $X = c_0(\mathbb{R})$, or $X = l^p(\mathbb{R})$, $1 \leq p < \infty$, and $A \in \text{Lip}_b(D_X^\alpha, X)$. Then, for each $n \geq 1$ and any $\delta > 0$, the n -th iterate of the Ruelle operator preserves $\text{Lip}_b(D_X^\alpha, X)$ and $\text{Lip}_b^{\text{loc}(\delta)}(D_X^\alpha, X)$. Moreover, for

$$D_n := \sum_{j=1}^n (d_j)^{-\alpha}, \quad C_A := \text{Lip}_{A, D_X^\alpha}$$

we obtain the following estimates.

a) For any $\varphi \in \text{Lip}_b(D_X^\alpha, X)$ and $x, y \in X$, $x \neq y$ we have

$$\frac{|\mathcal{L}_A^n(\varphi)(x) - \mathcal{L}_A^n(\varphi)(y)|}{\|x - y\|_X \mathcal{L}_A^n(1)(x)} \leq (d_n)^{-\alpha} \text{Lip}_{\varphi, D_X^\alpha} + \|\varphi\|_\infty C_A D_n \frac{e^{2n\|A\|_\infty} - 1}{2n\|A\|_\infty}. \quad (5)$$

b) For any $\varphi \in \text{Lip}_b^{\text{loc}(\delta)}(D_X^\alpha, X)$ and $x, y \in X$, $x \neq y$ and $\|x - y\|_X < \delta$, we have

$$\frac{|\mathcal{L}_A^n(\varphi)(x) - \mathcal{L}_A^n(\varphi)(y)|}{\|x - y\|_X \mathcal{L}_A^n(1)(x)} \leq (d_n)^{-\alpha} \text{Lip}_{\varphi, D_X^\alpha} + \|\varphi\|_\infty \frac{e^{C_A D_n \delta^\alpha} - 1}{\delta^\alpha}. \quad (6)$$

Proof. First note that for $x = (x_i), y = (y_i) \in X$,

$$\begin{aligned} & D_X \left(\left(r_j, \dots, r_1/\beta_1^{j-1}, x_1/\beta_1^j, \dots \right), \left(r_j, \dots, r_1/\beta_1^{j-1}, x_1/\beta_1^j, \dots \right) \right) \\ &= \left\| \left(x_1/\beta_1^j, x_2/\beta_2^j, \dots \right) - \left(y_1/\beta_1^j, y_2/\beta_2^j, \dots \right) \right\|_X \leq (d_j)^{-1} D_X(x, y). \end{aligned}$$

It now follows from Hölder continuity of A that

$$\begin{aligned} (*) &:= \left| e^{\sum_{j=1}^n A(r_j, \dots, r_1/\beta_1^{j-1}, x_1/\beta_1^j, \dots)} - e^{\sum_{j=1}^n A(r_j, \dots, r_1/\beta_1^{j-1}, y_1/\beta_1^j, \dots)} \right| \\ &= e^{\sum_{j=1}^n A(r_j, \dots, x_1/\beta_1^j, \dots)} \left| e^{\sum_{j=1}^n A(r_j, \dots, x_1/\beta_1^j, \dots) - A(r_j, \dots, y_1/\beta_1^j, \dots)} - 1 \right| \\ &\leq e^{\sum_{j=1}^n A(r_j, \dots, x_1/\beta_1^j, \dots)} \left| e^{\text{Lip}_{\varphi, D_X^\alpha} \sum_{j=1}^n (d_j)^{-\alpha} \|x-y\|_X^\alpha} - 1 \right|. \end{aligned}$$

The above then implies that

$$(*) \leq \frac{e^{C_A D_n \|x-y\|_X^\alpha} - 1}{\|x-y\|_X^\alpha} e^{\sum_{j=1}^n A(r_j, \dots, x_1/\beta_1^j, \dots)} \|x-y\|_X^\alpha,$$

which provides an effective estimate as long as $\|x-y\|_X^\alpha$ is uniformly bounded. That is, for $\|x-y\|_X \leq \delta$, it follows that

$$\begin{aligned} & |\mathcal{L}_A^n(\varphi)(x) - \mathcal{L}_A^n(\varphi)(y)| \\ &= \left| \int_{\mathbb{R}^n} e^{\sum_{j=1}^n A\left(r_j, \dots, \frac{r_1}{\beta_1^{j-1}}, \frac{x_1}{\beta_1^j}, \dots\right)} \varphi\left(r_n, \dots, \frac{r_1}{\beta_1^{n-1}}, \frac{x_1}{\beta_1^n}, \dots\right) \right. \\ &\quad \left. - e^{\sum_{j=1}^n A\left(r_j, \dots, \frac{r_1}{\beta_1^{j-1}}, \frac{y_1}{\beta_1^j}, \dots\right)} \varphi\left(r_n, \dots, \frac{r_1}{\beta_1^{n-1}}, \frac{y_1}{\beta_1^n}, \dots\right) dm(r_1) \cdots dm(r_n) \right| \\ &\leq \int_{\mathbb{R}^n} e^{\sum_{j=1}^n A\left(r_j, \dots, \frac{x_1}{\beta_1^j}, \dots\right)} \left| \varphi\left(r_n, \dots, \frac{x_1}{\beta_1^n}, \dots\right) - \varphi\left(r_n, \dots, \frac{y_1}{\beta_1^n}, \dots\right) \right| dm(r_1) \cdots dm(r_n) \\ &\quad + \|\varphi\|_\infty \int_{\mathbb{R}^n} (*) dm(r_1) \cdots dm(r_n) \\ &\leq \mathcal{L}_A^n(1)(x) \|x-y\|_X^\alpha \left(d_n^{-\alpha} \text{Lip}_{\varphi, D_X^\alpha} + \|\varphi\|_\infty \frac{e^{C_A D_n \delta^\alpha} - 1}{\delta^\alpha} \right), \end{aligned}$$

which proves the estimate in (6). On the other hand, if $\|x-y\|_X^\alpha$ is large, then the uniform boundedness of A implies that

$$(*) \leq \frac{e^{2n\|A\|_\infty} - 1}{\|x-y\|_X^\alpha} e^{\sum_{j=1}^n A(r_j, \dots, x_1/\beta_1^j, \dots)} \|x-y\|_X^\alpha,$$

In particular, by choosing δ such that

$$\frac{e^{2n\|A\|_\infty} - 1}{\delta^\alpha} = \frac{e^{C_A D_n \delta^\alpha} - 1}{\delta^\alpha},$$

that is $\delta^\alpha = 2n\|A\|_\infty/(C_A D_n)$ one obtains that

$$\min \left\{ \frac{e^{2n\|A\|_\infty} - 1}{\delta^\alpha}, \frac{e^{C_A D_n \delta^\alpha} - 1}{\delta^\alpha} \right\} \leq C_A D_n \frac{e^{2n\|A\|_\infty} - 1}{2n\|A\|_\infty}.$$

The estimate (5) follows from this. With respect to the invariance of $\text{Lip}_b(D_X^\alpha, X)$ and $\text{Lip}_b^{\text{loc}(\delta)}(D_X^\alpha, X)$, it suffices to observe that $\|\mathcal{L}_A^n(1)\|_\infty \leq \exp(n\|A\|_\infty) < \infty$ and apply (5) and (6), respectively. \square

Remark 4. Consider $1 \leq p < q < \infty$, since $l^p(\mathbb{R}) \subset l^q(\mathbb{R})$ and $\|x\|_{l^q(\mathbb{R})} \leq \|x\|_{l^p(\mathbb{R})}$ for all $x \in l^p(\mathbb{R})$, we deduce that if $A \in \text{Lip}_b(D_X^\alpha, l^q(\mathbb{R}))$, then the Ruelle operator preserves the set $\text{Lip}_b(D_X^\alpha, l^p(\mathbb{R}))$.

Remark 5. We would like to remark now that the reason to consider the action of \mathcal{L}_A on functions which are locally Lipschitz continuous is based on the fact that the second term in (6) is uniformly bounded whenever $\sup_n D_n < \infty$. Moreover, a further advantage will become apparent in the appendix (Section 5), where we will consider a bounded metric \tilde{D} on X which is equivalent to the metric $D_X(x, y) = \|x - y\|_X$, where $\|\cdot\|_X$ stands for norm on X . The necessity for this change of metric is twofold. The first reason is based on the simple observation that the Wasserstein space with respect to a bounded metric contains all Borel probability measures (see, e.g., [30]) and therefore, provided that A is normalized, the dual \mathcal{L}_A^* acts on this space. The second reason is of technical nature as this allows to obtain contraction for probability measures whose supports are arbitrarily distant.

However, the new metric also requires a change in the underlying function spaces. Namely, as it will turn out below, the space of bounded, locally Lipschitz functions with respect to D will coincide with the space of Lipschitz functions with respect to \tilde{D} .

We now define the dual \mathcal{L}_A^* of \mathcal{L}_A based on the observation that \mathcal{L}_A acts on the space of bounded Lipschitz functions by Lemma 1. That is,

$$\varphi \rightarrow \int \varphi d\mathcal{L}_A^*(\mu) := \int \mathcal{L}_A(\varphi) d\mu$$

defines an action of on the space of finite measures. Furthermore, if A is normalized, then \mathcal{L}_A^* acts on the space of sigma-additive probability measures. In here, we put emphasis on sigma-additivity in order to avoid any ambiguity with finitely additive measures, which are canonical in the similar setting of [9].

Definition 3. *Given a normalized potential $A : X \rightarrow \mathbb{R}$, the sigma-additive probability μ_A on X , such that, $\mathcal{L}_A^*(\mu_A) = \mu_A$, will be called the Gibbs probability for the potential A .*

We emphasize here the fact that the *Gibbs probability* terminology can have several different meanings depending on the authors of each work under consideration. Here the concept appears when this invariant probability is obtained

from Ruelle's Theorem. The terminology *equilibrium probability* (not considered here) would be reserved for those derived from a principle of maximizing pressure.

Observe that Remark 3 immediately implies that a Gibbs probability μ_A is always L -invariant. Furthermore, the existence of eigenfunctions of \mathcal{L}_A and eigenprobabilities of \mathcal{L}_A^* , as given in Theorem 1 below, provides important tools for solving problems related to the construction of equilibrium probabilities, that is invariant probabilities that maximize the pressure (see [22]).

Before the statement of our main result we introduce the class of weighted shifts L for which Theorem 1 holds. Fix $0 < c < c'$, with $1 < c'$ and consider a sequence $(\alpha_n)_{n \geq 1}$ satisfying $\alpha_n \in (c, c')$, for each $n \in \mathbb{N}$. Suppose also that

$$\sum_{n=1}^{\infty} (d_n)^{-\alpha} = d < \infty, \quad (7)$$

where d_n is given in (1). Observe that if $c > 1$, then (7) is always satisfied. Moreover, since (7) implies that

$$\sum_{n=1}^{\infty} (\beta_1^n)^{-p} < \infty,$$

we deduce as a consequence of Remark 1 that the weighted shift L associated to the sequence $(\alpha_n)_{n \geq 1}$ is **topologically mixing, Devaney chaotic and frequently hypercyclic** (see for instance [5], [12], and [13]). In order to guarantee the existence of invariant measures, we need the following property.

Definition 4. *We say that a Borelian measure m defined on \mathbb{R} has adapted tails if, for any $\epsilon > 0$ there exists a sequence of positive numbers $(\kappa_n)_{n \geq 1}$, such that,*

1. $\sum_{n=1}^{\infty} m(\mathbb{R} \setminus [-\beta_1^n \kappa_n, \beta_1^n \kappa_n]) < \epsilon$;
2. the sequence $(\kappa_n)_{n \geq 1}$ is in X .

For examples of measures with adapted tails we refer to Propositions 4 and 5. Recall that the support of a Borelian measure ν on a metric space (D, X) is defined by $\text{supp}(\nu) := \{x \in X : \nu(\{y \in X : D(x, y) < \epsilon\}) > 0, \forall \epsilon > 0\}$.

Theorem 1. *Assume that $X = c_0(\mathbb{R})$ or $X = l^p(\mathbb{R})$ with $1 \leq p < \infty$, that $L : X \rightarrow X$ is a weighted shift satisfying (7) and that $A \in \text{Lip}_b(D_X^\alpha, X)$.*

- a) *If $A \in \mathcal{SV}(X)$, then there exist $\lambda_A > 0$ and $\psi_A \in \text{Lip}_b(D_X^\alpha, X)$ with $\inf_{x \in X} \psi_A(x) > 0$, such that, $\mathcal{L}_A(\psi_A)(x) = \lambda_A \psi_A(x)$, for each $x \in X$. Moreover, λ_A is equal to the spectral radius of the action of \mathcal{L}_A on $\text{Lip}_b(D_X^\alpha, X)$, and any positive eigenfunction in $\text{Lip}_b(D_X^\alpha, X)$ is colinear to ψ_A .*
- b) *If m has adapted tails and $\mathcal{L}_A(1) = 1$, then there exists a unique σ -additive Borel probability measure μ_A which is a fixed point for the operator \mathcal{L}_A^* . Moreover, $\lim_{n \rightarrow \infty} (\mathcal{L}_A^n)^*(\nu) = \mu_A$ in the weak-* topology for any Borel probability measure ν . In particular, μ_A is an invariant and mixing probability measure (and therefore conservative and ergodic).*

- c) If $A \in \mathcal{SV}(X)$ and m has adapted tails, then there exists a unique fixed point μ_A for the operator $\mathcal{L}_{\bar{A}}^*$, where $\bar{A} = A + \log(\psi_A) - \log(\psi_A \circ L) - \log(\lambda_A)$. Moreover, μ_A is an invariant and mixing sigma-additive probability measure.
- d) If $A \in \mathcal{SV}(X)$ and π_1 is the projection on the first coordinate, we get $\text{supp}(\mu_A) \supset \pi_1^{-1}(\text{supp}(m))$. From this follows that $\text{supp}(\mu_A) = X$.

Before giving the proof, we remark that item c) of Theorem 1 in combination with the normalization of Remark 3 allows us to extend Definition 3 to potentials which are not necessarily normalized. That is, for $A \in \mathcal{SV}(X)$ and m with adapted tails, we refer to μ_A as in c) as the Gibbs measure for the potential A . Moreover, we would like to remark that item b) only requires that A is a Hölder continuous and normalized potential which is not necessarily of bounded distortion (see (3)).

Proof. We postpone the discussion of the existence of μ_A in item b) to Theorem 3 in the Appendix (Section 5), but the remaining statements of item b) are proven below. For the remaining parts, we adapt the results that appear in [2] and [20].

Part a) For each $s \in (0, 1)$ define the operator $\mathcal{T}_{s,A} : \mathcal{C}_b(X) \rightarrow \mathcal{C}_b(X)$ by

$$\mathcal{T}_{s,A}(u)(x) = \log \left(\int_{Lv=x} e^{A(v)+su(v)} dm(v) \right). \quad (8)$$

The use of this kind of operator is quite common in the so called discounted method (see Section 1 in [2]). In order to show existence of the eigenfunction ψ_A for the Ruelle operator \mathcal{L}_A it will be required later to take $s \rightarrow 1$.

For fixed s , and using the coordinates notation obtained from the Schauder basis $\{e_k\}_{k \geq 1}$, it follows that the equation (8) can be written as

$$\mathcal{T}_{s,A}(u)(x_1, x_2, \dots) = \log \left(\int_{\mathbb{R}} e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots) + su(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots)} dm(r) \right). \quad (9)$$

Since the maps A and u are bounded continuous functions and m is a probability measure, we deduce that $\mathcal{T}_{s,A}(u) \in \mathcal{C}_b(X)$ for all $u \in \mathcal{C}_b(X)$.

Now, observe that for all $x, y \in X$,

$$\begin{aligned} |\mathcal{T}_{s,A}(u_1)(x) - \mathcal{T}_{s,A}(u_2)(x)| &= \left| \log \left(\frac{\int_{Lv=x} e^{A(v)+su_1(v)} dm(v)}{\int_{Lv=x} e^{A(v)+su_2(v)} dm(v)} \right) \right| \\ &\leq \log \left(\frac{\int_{Lv=x} e^{A(v)+su_2(v)+s\|u_1-u_2\|_{\infty}} dm(v)}{\int_{Lv=x} e^{A(v)+su_2(v)} dm(v)} \right) \\ &= s\|u_1 - u_2\|_{\infty}. \end{aligned}$$

This shows that for all $0 < s < 1$, the operator $\mathcal{T}_{s,A}$ is a uniform contraction. Therefore, since $\mathcal{C}_b(X)$ is a complete metric space with the distance induced by

the uniform norm $\|\cdot\|_\infty$, it follows from the Banach contraction principle that $\mathcal{T}_{s,A}$ has a unique fixed point $u_s \in \mathcal{C}_b(X)$. That is, for all x and $0 < s < 1$,

$$e^{u_s(x)} = \int_{Lv=x} e^{A(v)+su_s(v)} dm(v) .$$

Now, we will show that the collection $(u_s^*)_{0 < s < 1}$, with

$$u_s^*(x) = u_s(x) - u_s(0),$$

for each $x \in X$, is equicontinuous and uniformly bounded.

By the Arzelá-Ascoli Theorem, as a consequence of the Cantor-Tychonoff Theorem, we claim the existence of a sequence $(s_n)_{n \geq 1}$ satisfying

- i) $\lim_{n \rightarrow \infty} s_n = 1$;
- ii) $\lim_{n \rightarrow \infty} u_{s_n}^*(x) = u(x)$ for each $x \in X$;
- iii) $\lim_{n \rightarrow \infty} u_{s_n}^* = u$ as uniform limit on any compact subset of X .

We will now show that the function $e^u = \psi_A$ is the main eigenfunction for \mathcal{L}_A . That is, ψ_A satisfies a), with λ_A maximal on the set of eigenvalues of \mathcal{L}_A .

In order to prove that $(u_s^*)_{0 < s < 1}$ is equicontinuous, for any pair $x, y \in X$, set

$$\begin{aligned} \mathcal{S}_1(x, y) := \sup_{r \in \mathbb{R}} \left\{ A\left(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots\right) - A\left(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \dots\right) \right. \\ \left. + s\left(u_s\left(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots\right) - u_s\left(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \dots\right)\right) \right\} \end{aligned}$$

and note that $\sup_{x,y} \mathcal{S}_1(x, y) < \infty$ as A and u_s are bounded. Using the expression in coordinates for the operator $\mathcal{T}_{s,A}$ (which appears in (9)) and the fact that u_s is a fixed point for $\mathcal{T}_{s,A}$, it follows that

$$\begin{aligned} e^{u_s(x)} &= \int_{\mathbb{R}} e^{A\left(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots\right) + su_s\left(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots\right)} dm(r) \\ &\leq e^{\mathcal{S}_1(x,y)} \int_{\mathbb{R}} e^{A\left(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \dots\right) + su_s\left(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \dots\right)} dm(r) \\ &= e^{\mathcal{S}_1(x,y)} e^{u_s(y)} . \end{aligned}$$

That is,

$$|u_s(x) - u_s(y)| \leq \mathcal{S}_1(x, y) < +\infty .$$

Therefore, following an inductive argument on the first n coordinates of the points in which it is calculated the function $A + su_s$, it is easy to check that for any $n \in \mathbb{N}$, the inequality

$$|u_s(x) - u_s(y)| \leq \mathcal{S}_n(x, y) < +\infty , \tag{10}$$

is satisfied, where

$$\begin{aligned} \mathcal{S}_n(x, y) := & \sup_{(r_1, \dots, r_n)} \left\{ \left(\sum_{j=1}^n s^{j-1} \left(A\left(r_j, \frac{r_{j-1}}{\beta_1^1}, \dots, \frac{r_1}{\beta_1^{j-1}}, \frac{x_1}{\beta_1^j}, \frac{x_2}{\beta_2^j}, \dots\right) \right. \right. \right. \\ & \left. \left. \left. - A\left(r_j, \frac{r_{j-1}}{\beta_1^1}, \dots, \frac{r_1}{\beta_1^{j-1}}, \frac{y_1}{\beta_1^j}, \frac{y_2}{\beta_2^j}, \dots\right) \right) \right) \right. \\ & \left. + s^n \left(u_s\left(r_n, \frac{r_{n-1}}{\beta_1^1}, \dots, \frac{r_1}{\beta_1^{n-1}}, \frac{x_1}{\beta_1^n}, \frac{x_2}{\beta_2^n}, \dots\right) \right. \right. \\ & \left. \left. - u_s\left(r_n, \frac{r_{n-1}}{\beta_1^1}, \dots, \frac{r_1}{\beta_1^{n-1}}, \frac{y_1}{\beta_1^n}, \frac{y_2}{\beta_2^n}, \dots\right) \right) \right\}. \end{aligned}$$

Moreover, for any pair $x, y \in X$, the sequence $(\mathcal{S}_n(x, y))_{n \geq 1}$ satisfies

$$\begin{aligned} & \mathcal{S}_n(x, y) \\ & \leq \sup_{(r_1, \dots, r_n)} \left\{ \left(\sum_{j=1}^n s^{j-1} \text{Lip}_{A, D_X^\alpha} \left\| \left(r_j, \frac{r_{j-1}}{\beta_1^1}, \dots, \frac{r_1}{\beta_1^{j-1}}, \frac{x_1}{\beta_1^j}, \frac{x_2}{\beta_2^j}, \dots \right) \right. \right. \right. \\ & \quad \left. \left. \left. - \left(r_j, \frac{r_{j-1}}{\beta_1^1}, \dots, \frac{r_1}{\beta_1^{j-1}}, \frac{y_1}{\beta_1^j}, \frac{y_2}{\beta_2^j}, \dots \right) \right\|_X^\alpha \right) + 2s^n \|u_s\|_\infty \right\} \\ & = \left(\sum_{j=1}^n s^{j-1} \text{Lip}_{A, D_X^\alpha} \left\| \left(\frac{x_1}{\beta_1^j}, \frac{x_2}{\beta_2^j}, \dots \right) - \left(\frac{y_1}{\beta_1^j}, \frac{y_2}{\beta_2^j}, \dots \right) \right\|_X^\alpha \right) + 2s^n \|u_s\|_\infty \\ & \leq \left(\sum_{j=1}^n s^{j-1} \text{Lip}_{A, D_X^\alpha} (d_j)^{-\alpha} \|x - y\|_X^\alpha \right) + 2s^n \|u_s\|_\infty \\ & \leq \left(\sum_{j=1}^n \text{Lip}_{A, D_X^\alpha} (d_j)^{-\alpha} \|x - y\|_X^\alpha \right) + 2s^n \|u_s\|_\infty. \end{aligned}$$

Therefore, taking the limit when $n \rightarrow \infty$, it follows that for all $s \in (0, 1)$ and each pair $x, y \in X$,

$$|u_s(x) - u_s(y)| \leq d \text{Lip}_{A, D_X^\alpha} \|x - y\|_X^\alpha, \quad (11)$$

where d is as in (7). By the above, it follows that the family $(u_s)_{0 < s < 1}$ is equicontinuous. Moreover, this implies immediately that $(u_s^*)_{0 < s < 1}$ (such as defined above) is equicontinuous as well.

Besides that,

$$\begin{aligned} \mathcal{S}_n(x, y) & \leq \sup_{(r_1, \dots, r_n)} \left\{ \left(\sum_{j=1}^n s^{j-1} V_j(A) + 2s^n \|u_s\|_\infty \right) \right\} \\ & \leq \sum_{j=1}^n V_j(A) + 2s^n \|u_s\|_\infty. \end{aligned}$$

Thus, taking the limit as $n \rightarrow \infty$, it follows from (3) and (10) that for any $s \in (0, 1)$ is satisfied

$$|u_s(x) - u_s(y)| \leq V(A) .$$

Since, we have $u_s^*(0) = 0$, for all $s \in (0, 1)$, then, for any $x \in X$ is satisfied

$$|u_s^*(x)| \leq |u_s(x) - u_s(0)| \leq V(A) , \quad (12)$$

that is, the family $(u_s^*)_{0 < s < 1}$ is uniformly bounded.

Note that for all x and s we get

$$-\|A\|_\infty + s \min u_s \leq u_s(x) \leq \|A\|_\infty + s \max u_s .$$

From this, it follows:

$$-\|A\|_\infty \leq (1 - s) \min u_s \leq (1 - s) \max u_s \leq \|A\|_\infty ,$$

for any $0 < s < 1$.

The family $(1 - s) u_s(0)$, $0 < s < 1$, is bounded and determines a convergent sequence s_n , such that, $\kappa := \lim_{n \rightarrow \infty} (1 - s_n) u_{s_n}(0)$.

Considering the sequences u_{s_n} and $u_{s_n}^* = u_{s_n} - u_{s_n}(0)$, we use Arzelá-Ascoli's Theorem in order to get another subsequence (of the given sequence $u_{s_n}^*$) which converges uniformly on compact subsets of X and pointwise for all $x \in X$. We also denote this new subsequence of index by s_n , $n \in \mathbb{N}$.

Let u be the function satisfying the uniform limit $\lim_{n \rightarrow \infty} u_{s_n}^* = u$ on compact subsets of X and the pointwise limit $\lim_{n \rightarrow \infty} u_{s_n}^*(x) = u(x)$, for all $x \in X$. Then, for each $x \in X$ the value u satisfies the equation

$$e^{u(x)} = e^{-\kappa} \int_{Lv=x} e^{A(v)+u(v)} dm(v) , \quad (13)$$

with $\kappa = \lim_{n \rightarrow \infty} (1 - s_n) u_{s_n}(0)$.

Indeed, since $\mathcal{T}_{s_n, A}(u_{s_n}) = u_{s_n}$, we have

$$e^{u_{s_n}^*(x)} = e^{-(1-s_n)u_{s_n}(0)} \int_{Lv=x} e^{A(v)+s_n u_{s_n}^*(v)} dm(v) . \quad (14)$$

It follows that

$$\int_{Lv=x} e^{A(v)+s_n u_{s_n}^*(v)} dm(v) = \int_{\mathbb{R}} e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots) + s_n u_{s_n}^*(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots)} dm(r) ,$$

and by (12) we have that

$$\int_{\mathbb{R}} \left| e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots) + s_n u_{s_n}^*(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots)} \right| dm(r) \leq C \int_{\mathbb{R}} dm(r) < \infty ,$$

where $C = e^{\|A\|_\infty + V(A)}$. Therefore, (13) is a consequence of (14) and the Dominated Convergence Theorem.

The foregoing implies that the positive function $\psi_A(x) = e^{u(x)}$, satisfies the equation $\mathcal{L}_A(\psi_A)(x) = \lambda_A \psi_A(x)$ for all $x \in X$, with $\lambda_A = e^{-\kappa}$. Furthermore, by (11), it follows for each $x, y \in X$ that

$$|u(x) - u(y)| = \lim_{n \rightarrow \infty} |u_{s_n}^*(x) - u_{s_n}^*(y)| \leq d \text{Lip}_{A, D_X^\alpha} \|x - y\|_X^\alpha, \quad (15)$$

and by (12), we obtain that, for any $x \in X$,

$$|u(x)| = \lim_{n \rightarrow \infty} |u_{s_n}^*(x)| \leq V(A), \quad (16)$$

By (15) and (16), it follows that $u \in \text{Lip}_b(D_X^\alpha, X)$, and thus $\psi_A = e^u \in \text{Lip}_b(D_X^\alpha, X)$. In order to prove that λ_A is the maximal eigenvalue, we show that λ_A coincides with the spectral radius σ of the action of \mathcal{L}_A on Lip_b . In order to do so, we consider

$$R(\varphi) := \limsup_{n \rightarrow \infty} \left(\|\mathcal{L}_A^n(\varphi)\|_\infty + \text{Lip}_{\mathcal{L}_A^n(\varphi), D_X^\alpha} \right)^{1/n},$$

for $\varphi \in \text{Lip}_b(D_X^\alpha, X)$. Observe that $R(\psi_A) = \lambda_A$ and that $R(\varphi) + t\psi_A = \max\{R(\varphi), R(\psi_A)\}$ for any $t > 0$, which gives rise to the following argument. Assume that $\sigma > \lambda_A$. It then follows from the spectral radius formula that there exists $\varphi \in \text{Lip}_b(D_X^\alpha, X)$ with $R(\varphi) > \lambda_A$. As $\inf_x \psi_A(x) > 0$, the above implies that there is $t > 0$ and $c > 0$ such that $\varphi_t := \varphi + t\psi_A$ satisfies $\psi_A/c \leq \varphi_t \leq c\psi_A$ and $R(\varphi_t) = R(\varphi) > \lambda_A$. Hence, for all $n \in \mathbb{N}$,

$$0 < \inf_x \psi_A(x)/c \leq \psi_A/c \leq \mathcal{L}_A^n(\varphi_t)/\lambda_A^n \leq c\psi_A \leq c \sup_x \psi_A(x) < \infty.$$

As the same argument shows that $\mathcal{L}_A^n(1)/\lambda_A^n$ is uniformly bounded from above and below, an application of item b) of Lemma 1 shows that the local Hölder coefficients of $\mathcal{L}_A^n(\varphi_t)/\lambda_A^n$ are uniformly bounded from above. Hence, $\mathcal{L}_A^n(\varphi_t)/\lambda_A^n$ is a Hölder continuous function with uniformly bounded Lipschitz coefficients (see Lemma 4). This implies that

$$\lambda_A < R(\varphi_t) = \limsup_{n \rightarrow \infty} (\|\mathcal{L}_A^n(\varphi_t)\|_\infty)^{1/n} \leq \limsup_{n \rightarrow \infty} (\|c\mathcal{L}_A^n(\psi_A)\|_\infty)^{1/n} = \lambda_A,$$

which is absurd. Hence, $\lambda_A = \sigma$.

It remains to show that ψ_A is the unique, strictly positive eigenfunction. In order to do so, one uses ψ_A in order to define the normalization \bar{A} of A as given in Remark 3. The mixing property of item b) then implies that the unique positive eigenfunctions of \mathcal{L}_A^n in $\text{Lip}_b(D_X^\alpha, X)$ are the constant functions. This concludes the proof of a).

Part b) The proof that there exist a σ -additive probability μ_A on X which is fixed for the action of the operator \mathcal{L}_A^* and that $(\mathcal{L}_A^n)^*(\nu) \rightarrow \mu_A$ for any σ -additive probability measure ν will be proven in the Appendix (see Section 5). Moreover, note that, for any $\varphi \in \mathcal{C}_b(X)$, we have

$$\int_X (\varphi \circ L) d\mu_A = \int_X (\varphi \circ L) d(\mathcal{L}_A^*(\mu_A)) = \int_X \mathcal{L}_A(\varphi \circ L) d\mu_A = \int_X \varphi d\mu_A.$$

Hence, μ_A is L -invariant and, as a consequence of Poincaré's Recurrence Theorem, it is conservative. We now show that μ_A is mixing. As $(\mathcal{L}_A^n)^*(\nu) \rightarrow \mu_A$ for any ν in the weak-* topology, it follows that, for any $x \in X$ and any bounded continuous function $f : X \rightarrow \mathbb{R}$, that

$$\mathcal{L}_A^n(f)(x) = \int_X f d(\mathcal{L}_A^n)^*(\delta_x) \xrightarrow{n \rightarrow \infty} \int_X f d\mu_A.$$

Hence, by Dominated Convergence Theorem, for any pair of bounded continuous functions $f, g : X \rightarrow \mathbb{R}$,

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_X f g \circ L^n d\mu_A &= \lim_{n \rightarrow \infty} \int_X f (g \circ L^n) d(\mathcal{L}_A^n)^*(\mu_A) = \lim_{n \rightarrow \infty} \int_X \mathcal{L}_A^n(fg \circ L^n) d\mu_A \\ &= \lim_{n \rightarrow \infty} \int_X \mathcal{L}_A^n(f)g d\mu_A = \int_X f d\mu_A \int g d\mu_A. \end{aligned}$$

This implies that μ_A is mixing showing that all claims of item b) are true.

Part c) This is an immediate consequence of parts a) and b), as any potential $A \in \text{Lip}_b(D_X^\alpha, X) \cap \mathcal{SV}(X)$ satisfies $\mathcal{L}_{\bar{A}}(1) = 1$. Indeed, by part a), for each $x \in X$ we have

$$\begin{aligned} \mathcal{L}_{\bar{A}}(1)(x) &= \int_{\mathbb{R}} e^{\bar{A}(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots)} dm(r) \\ &= \frac{1}{\lambda_A \psi_A(x)} \int_{\mathbb{R}} e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots)} \psi_A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \dots) dm(r) \\ &= \frac{\mathcal{L}_A(\psi_A)(x)}{\lambda_A \psi_A(x)} \\ &= 1. \end{aligned}$$

Part d) Consider $U(x; \epsilon) = \{y \in X, |y_1 - x_1| \leq \epsilon, \text{ and } x_n = y_n; \forall n > 1\}$. Let U be a neighborhood of x , then, there is $\epsilon > 0$ such that $U(x, \epsilon) \subset U$. By Urysohn's Lemma, let $\varphi \in \mathcal{C}_b(X)$ such that $\chi_{U(x; \epsilon)} \leq \varphi \leq U$. Then,

$$\begin{aligned} \mu_A(U) &\geq \int_X \varphi d\mu_A \\ &= \int_X \mathcal{L}_A(\varphi) d\mu_A \\ &= \int_X \left(\int_{\mathbb{R}} e^{A(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \dots)} \varphi(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \dots) dm(r) \right) d\mu_A(y) \\ &\geq e^{\inf(A)} \int_X \left(\int_{\mathbb{R}} \chi_{U(x; \epsilon)}(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \dots) dm(r) \right) d\mu_A(y) \\ &\geq e^{\inf(A)} m([x_1 - \epsilon, x_1 + \epsilon]) > 0. \end{aligned}$$

The above implies that $\mu_A(U) > 0$, for any open neighborhood $U \subset X$ of x . Furthermore, since $\mu_A(U) \geq e^{\inf(A)} m([x_1 - \epsilon, x_1 + \epsilon])$, we also have that the support of μ_A contains the inverse image of the support of the a priori measure m by the map π_1 . This ends the proof. \square

From Theorem 1 and Remark 4 we can deduce the following result.

Corollary 1. *Let $1 \leq p < q$ be real numbers and $X = l^q(\mathbb{R})$. Assume that $A \in \text{Lip}_b(D_X^\alpha, X)$ and $L : l^p(\mathbb{R}) \rightarrow l^p(\mathbb{R})$ is a weighted shift satisfying (7). Then, Theorem 1 holds.*

3.1 A class of weighted shifts satisfying Theorem 1

In this section we will consider a different presentation of some results of last section. We have seen that Theorem 1 holds for the class of weighted shifts L satisfying (7). Using Cauchy's criterion, we deduce that this class contains all the weighted shifts satisfying

$$\lim_{n \rightarrow \infty} (d_n)^{-\frac{1}{n}} = \lim_{n \rightarrow \infty} \left(\inf_{k \geq 1} \beta_k^n \right)^{-\frac{1}{n}} < 1.$$

In fact, we will prove in the next proposition that the last two classes coincide.

Lemma 2. *(see [5]) The following assertions are equivalent:*

1. $\lim_{n \rightarrow \infty} (d_n)^{-\frac{1}{n}} < 1$.
2. $\sum_{n=1}^{\infty} (d_n)^{-1} < \infty$.
3. $\sup_{k \geq 1} \sum_{n=1}^{\infty} (\beta_k^n)^{-1} < \infty$.

Proof. It is easy to prove that (i) implies (ii) and (ii) implies (iii). Let us show that (iii) implies (i). For each $k \in \mathbb{N}$, define $S_k = \sum_{n=1}^{\infty} (\beta_k^n)^{-1}$. By hypothesis, there is a constant $M > 0$, such that, $S_k \leq M$ for all $k \geq 1$.

Since $\alpha_k^{-1} S_{k+1} = S_k - \alpha_k^{-1}$ for each $k \in \mathbb{N}$, it follows that

$$\begin{aligned} (\beta_k^n)^{-1} &= \frac{S_k}{1 + S_{k+n}} \cdot \frac{S_{k+1}}{1 + S_{k+1}} \cdot \dots \cdot \frac{S_{k+n-1}}{1 + S_{k+n-1}} \\ &\leq S_k \cdot \frac{S_{k+1}}{1 + S_{k+1}} \cdot \dots \cdot \frac{S_{k+n-1}}{1 + S_{k+n-1}} \\ &\leq M \cdot \left(\frac{M}{1 + M} \right)^{n-1}. \end{aligned}$$

This implies that $\limsup_{n \rightarrow \infty} (d_n)^{-\frac{1}{n}} \leq \frac{M}{1 + M} < 1$.

On the other hand, we have

$$(d_n)^{-\frac{1}{n}} = \left(\sup_{k \geq 1} (1/\beta_k^n) \right)^{\frac{1}{n}}.$$

Moreover, by (2) we have $\lim_{n \rightarrow \infty} \left(\sup_{k \geq 1} (1/\beta_k^n) \right)^{\frac{1}{n}} = \lim_{n \rightarrow \infty} \|(L')^n\|_{op}^{\frac{1}{n}}$, which is the spectral radius $r(L')$ of the weighted shift operator L' defined by

$$L'((x_n)_{n \geq 1}) = (\alpha_n^{-1} x_{n+1})_{n \geq 1} .$$

Hence $\lim_{n \rightarrow \infty} (d_n)^{-\frac{1}{n}} = \limsup_{n \rightarrow \infty} (d_n)^{-\frac{1}{n}} < 1$. \square

We deduce from Lemma 2 that for all $\alpha > 0$, the condition $\sum_{n=1}^{\infty} (d_n)^{-\alpha} < \infty$ is equivalent to $\lim_{n \rightarrow \infty} (d_n)^{-\frac{1}{n}} < 1$.

As a consequence, we obtain the following result.

Proposition 1. *Let X be either $c_0(\mathbb{R})$ or $l^p(\mathbb{R})$, $1 \leq p < \infty$. Assume that $A \in \text{Lip}_b(D_X^\alpha, X)$ and $L : X \rightarrow X$ is a weighted shift satisfying*

$$\lim_{n \rightarrow \infty} (d_n)^{-\frac{1}{n}} < 1 . \quad (17)$$

Then, the claims of items a), b), c) and d) of Theorem 1 are satisfied.

Proof. By Lemma 2, $\lim_{n \rightarrow \infty} (d_n)^{-\frac{1}{n}} < 1$ is equivalent to (7). \square

Remark 6. Consider $X \in \{c_0(\mathbb{C}), l^p(\mathbb{C}), 1 \leq p < \infty\}$. Note that the claims of Theorem 1 and Proposition 1 hold if $A \in \text{Lip}_b(D_X^\alpha, X)$ and $L : X \rightarrow X$ is a weighted shift satisfying $\lim_{n \rightarrow \infty} \left(\inf_{k \geq 1} |\beta_k^n| \right)^{-\frac{1}{n}} < 1$.

Example 1. There are many cases which can be considered on the class of weighted shifts operators satisfying (17). For instance:

1. Take

$$1 < c < \alpha_n \leq c', \text{ for all } n \geq n_0 ,$$

where n_0 is a non-negative integer and c is a fixed real number.

2. Assume that there exists an increasing sequence of non-negative integers $(k_n)_{n \geq 0}$, such that, $k_0 = 0$ and $k_i < k_{i+1}$, for all integer $i \geq 0$. Let a and b to be two real numbers, such that, $c \leq a < 1 < b \leq c'$, and

$$a \leq \alpha_j < 1, \text{ for all } k_{2i} + 1 \leq j \leq k_{2i+1}, i \geq 0$$

and

$$1 < b \leq \alpha_j < c', \text{ for all } k_{2i+1} + 1 \leq j \leq k_{2i+2}, i \geq 0 .$$

Set

$$r_n = k_n - k_{n-1} \text{ for all } n \geq 1 .$$

Assume that there exists a real number $e \geq 1$, such that,

$$1 \leq r_n \leq e, \text{ for all } n \geq 1 .$$

Then,

$$\beta_k^n = \alpha_k \dots \alpha_{k+n-1} \geq (a^e b)^{\lfloor \frac{n}{e+1} \rfloor} a^e, \text{ for all } k \geq 1 ,$$

where $\lfloor \frac{n}{e+1} \rfloor$ is the integer part of $\frac{n}{e+1}$. Assuming that $a^e b > 1$ we are able to present an interesting example.

4 Extension to a wider class of operators

Let X be a Banach space and $T : X \rightarrow X$ be a linear continuous operator. We denote by $m(T)$ the co-norm of T ; that is, $m(T) = \inf\{\|T(x)\|_X, \|x\|_X = 1\}$.

It is known that $m(T) > 0$, if and only if, T is one to one and has closed range. Moreover, it is easy to see that for all $x \in X$, $m(T)\|x\|_X \leq \|T(x)\|_X$. Another important property is that $\lim_{n \rightarrow \infty} (m(T^n))^{\frac{1}{n}}$ exists (see [7]).

Now assume that T is onto, not necessarily one to one. Suppose also that there exists a closed subspace E of X , such that, X is the direct sum of $\text{Ker}(T)$ and E , i.e., any vector $x \in X$ can be written in a unique form as $x = z + v$, with $z \in \text{Ker}(T)$ and $v \in E$. We will use the notation

$$X = \text{Ker}(T) \oplus E. \quad (18)$$

We define

$$p(T) = \inf\{\|T(x)\|_X, \|x\|_X = 1, x \in E\}.$$

Note that

$$p(T)\|x\|_X \leq \|T(x)\|_X, \text{ for all } x \in E. \quad (19)$$

We point out that in case the Banach space we consider in this section is a Hilbert space, then the required assumptions for splittings are automatically true.

Example 2. We consider the weighted shift $L : X \rightarrow X$, where X is $c_0(\mathbb{R})$, or $l^p(\mathbb{R})$, $1 \leq p < \infty$, which was given by

$$L(x_1, x_2, \dots) = (\alpha_1 x_2, \alpha_2 x_3, \dots).$$

Then, we have

$$\text{Ker}(L) = \{(x_i)_{i \geq 1} \in X, x_i = 0 \text{ for all } i \geq 2\}$$

and

$$E = \{(x_i)_{i \geq 1} \in X, x_1 = 0\}.$$

It's not difficult to see that

$$p(L) = \inf_{k \geq 1} \alpha_k. \quad (20)$$

Indeed, for all $n \geq 2$, we have $\|L(e_n)\|_X = \|\alpha_{n-1} e_{n-1}\|_X = \alpha_{n-1}$. Hence

$$p(L) \leq \inf_{k \geq 1} \alpha_k.$$

On the other hand,

$$\|L((x_n)_{n \geq 1})\| \geq \left(\inf_{k \geq 1} \alpha_k\right) \|(x_{n+1})_{n \geq 1}\|, \text{ for all } (x_2, x_3, \dots) \in X.$$

Hence $p(L) \geq \inf_{k \geq 1} \alpha_k$ and we obtain (20).

Now fixing an integer $n \geq 1$, we get $X = \text{Ker}(L^n) \oplus E_n$, where

$$\text{Ker}(L^n) = \{(x_i)_{i \geq 1} \in X, x_i = 0 \text{ for all } i \geq n+1\}$$

and

$$E_n = \{(x_i)_{i \geq 1} \in X, x_i = 0, \forall 1 \leq i \leq n\}.$$

In a similar way as in the case $n = 1$, we obtain

$$p(L^n) = d_n = \inf_{k \geq 1} \beta_k^n.$$

Observe that $\lim_{n \rightarrow \infty} (p(L^n))^{\frac{1}{n}}$ is the inverse of the constant given in Proposition 1.

The above claim can be stated in a more general form. For instance, assume that $T : X \rightarrow X$ is onto, not one to one and satisfies

$$X = \text{Ker}(T^n) \oplus E_n \text{ and } T(E_{n+1}) = E_n, \text{ for all } n \geq 1, \quad (21)$$

where the E_n 's are closed subspaces of X . We also assume that $T(E_1) = X$.

Note that the assumption (21) is satisfied by the weighted shift L in any space $X \in \{c_0(\mathbb{R}), l^p(\mathbb{R}), 1 \leq p < \infty\}$. Furthermore, (21) is also satisfied by any continuous operator $T : X \rightarrow X$ onto, but not one to one, where X is a Hilbert space, since in this case, for all integer $n \geq 1$, the set E_n equals the orthogonal space of $\text{Ker}(T^n)$.

Now, suppose that $p(T^n) > 0$ for all $n \geq 1$, then $\lim_{n \rightarrow \infty} (p(T^n))^{-\frac{1}{n}}$ exists.

Indeed, we have

$$p(T^{n+m}) \geq p(T^n)p(T^m), \text{ for all } n, m \in \mathbb{N}.$$

Hence, the sequence $(-\log(p(T^n)))_{n \geq 1}$ is sub-additive. Thus, $\lim_{n \rightarrow \infty} -\frac{\log(p(T^n))}{n}$

exists and is equal to $\inf_{n \geq 1} -\frac{\log(p(T^n))}{n}$.

Now, for all $x \in X \setminus \{0\}$, denote

$$T^{-1}(\{x\}) := \{v \in X, T(v) = x\}.$$

Since T is onto, $T^{-1}(\{x\})$ is not empty and

$$T^{-1}(\{x\}) = \{x'\} + \text{Ker}(T) = \{x' + z, z \in \text{Ker}(T)\},$$

where x' is an arbitrary element of $T^{-1}(\{x\})$.

We now fix an *a priori* probability m on $\text{Ker}(T)$. We will also denote by m the corresponding *a priori* measure on $T^{-1}(\{x\}) = \{x'\} + \text{Ker}(T)$. Then, this gives rise to a Ruelle operator associated to the potential $A \in \text{Lip}_b(D_X^\alpha, X)$ and

the a priori probability m , defined as the map assigning to each $\varphi \in \mathcal{C}_b(X)$ the function

$$\mathcal{L}_A(\varphi)(x) := \int_{v \in T^{-1}(\{x\})} e^{A(v)} \varphi(v) dm(v) = \int_{z \in \text{Ker}(T)} e^{A(x'+z)} \varphi(x'+z) dm(z) .$$

Some of the results in this section follow the same reasoning as before, and sometimes we just outline the proof.

Lemma 3. *Let X be a Banach space and $T : X \rightarrow X$ a bounded linear operator, such that, T is onto, not one to one, satisfying (18), $p(T) > 0$ and $A \in \text{Lip}_b(D_X^\alpha, X)$. Then, the Ruelle operator \mathcal{L}_A preserves the spaces $\text{Lip}_b(D_X^\alpha, X)$ and $\text{Lip}_b^{\text{loc}(\delta)}(D_X^\alpha, X)$, $\delta > 0$.*

Remark 7. It's classical result (and not difficult to prove) that $p(T) > 0$ is equivalent to the fact that $T : E \rightarrow X$ is injective and $T(E)$ is a closed subspace of X .

Proof. For any $\varphi \in \text{Lip}_b(D_X^\alpha, X)$ and each $x, y \in X$, we have

$$\begin{aligned} & |\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| \\ &= \left| \int_{z \in \text{Ker}(T)} e^{A(x'+z)} \varphi(x'+z) dm(z) - \int_{z \in \text{Ker}(T)} e^{A(y'+z)} \varphi(y'+z) dm(z) \right|, \end{aligned}$$

where

$$T^{-1}(\{x\}) = \{x'\} + \text{Ker}(T), \quad T^{-1}(\{y\}) = \{y'\} + \text{Ker}(T) .$$

Hence,

$$\begin{aligned} & |\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| \\ &\leq \int_{\text{Ker}(T)} e^{\|A\|_\infty} |\varphi(x'+z) - \varphi(y'+z)| + \|\varphi\|_\infty \left| e^{A(x'+z)} - e^{A(y'+z)} \right| dm(z) \\ &\leq \left(e^{\|A\|_\infty} \text{Lip}_{\varphi, D_X^\alpha} + \|\varphi\|_\infty \text{Lip}_{e^A, D_X^\alpha} \right) \int_{\text{Ker}(T)} \|x' - y'\|_X^\alpha dm(z) . \end{aligned}$$

Using (19) we deduce that

$$|\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| \leq \frac{e^{\|A\|_\infty} \text{Lip}_{\varphi, D_X^\alpha} + \|\varphi\|_\infty \text{Lip}_{e^A, D_X^\alpha}}{p(T)^\alpha} \|x - y\|_X^\alpha .$$

Then, $\mathcal{L}_A(\varphi) \in \text{Lip}_b(D_X^\alpha, X)$. Under the assumption $\|x - y\|_X^\alpha < \delta$ the proof for the case $\text{Lip}_b^{\text{loc}(\delta)}(D_X^\alpha, X)$ follows the same reasoning. \square

We say that a potential $A \in \mathcal{C}(X)$ has *summable variation* with respect to the linear operator $T : X \rightarrow X$ if satisfies

$$V_T(A) = \sum_{n=1}^{\infty} V_{T,n}(A) < \infty , \quad (22)$$

where

$$V_{T,n}(A) := \sup\{|A(z_n + x_n) - A(z_n + y_n)| : z_n \in \text{Ker}(T^n), x_n, y_n \in E_n\}.$$

We will denote by $\mathcal{SV}_T(X)$ the set of potentials $A \in \mathcal{C}(X)$ satisfying (22).

When considering a more general setting (as in this section) we need an assumption that is stronger than adapted tails condition which was used before.

Definition 5. *We say that a Borelian measure m defined on a vector space V isomorphic to \mathbb{R}^k , with $k \geq 1$, has strong adapted tails if, for any $\epsilon > 0$ there exists a sequence of positive numbers $(\kappa_n)_{n \geq 1}$, such that,*

1. $\sum_{n=1}^{\infty} i_* m(\mathbb{R}^k \setminus [-p(T^n)\kappa_n, p(T^n)\kappa_n]^k) < \epsilon$, where $i : V \rightarrow \mathbb{R}^k$ is an isomorphism between V and \mathbb{R}^k and $i_* m(E) = m(i^{-1}(E))$;
2. the sequence $(\kappa_n)_{n \geq 1}$ is in $l^1(\mathbb{R})$.

Theorem 2. *Let X be a separable Banach space and $T : X \rightarrow X$ a bounded linear operator, such that, T is onto, not bijective, satisfying (21), $p(T^n) > 0$ and $\dim(\text{Ker}(T^n)) < \infty$, for each $n \in \mathbb{N}$. Consider $A \in \text{Lip}_b(D_X^\alpha, X)$, assume that the a priori measure m has strong adapted tails and*

$$\sum_{n=1}^{\infty} (p(T^n))^{-\alpha} < \infty.$$

Then:

- a) *If $A \in \mathcal{SV}_T(X)$, then there exist $\lambda_A > 0$ and a strictly positive function $\psi_A \in \text{Lip}_b(D_X^\alpha, X)$ such that $\mathcal{L}_A(\psi_A)(x) = \lambda_A \psi_A(x)$, for each $x \in X$.*
- b) *If m has adapted tails and $\mathcal{L}_A(1) = 1$, then there exists a unique σ -additive Borel probability measure μ_A which is a fixed point for the operator \mathcal{L}_A^* . Moreover, $\lim_{n \rightarrow \infty} (\mathcal{L}_A^n)^*(\nu) = \nu$ in the weak-* topology for any Borel probability measure ν . In particular, μ_A is an invariant and mixing probability measure (and therefore conservative and ergodic).*
- c) *If $A \in \mathcal{SV}_T(X)$ and m has adapted tails, then there exists a unique fixed point μ_A for the operator \mathcal{L}_A^* , where $\bar{A} = A + \log(\psi_A) - \log(\psi_A \circ L) - \log(\lambda_A)$. Moreover, μ_A is an invariant and mixing σ -additive probability measure.*
- d) *If $A \in \mathcal{SV}_T(X)$ and $\pi : X \rightarrow \text{Ker}(T)$ is the projection on the kernel of T , we get $\text{supp}(\mu_A) \supset \pi^{-1}(\text{supp}(m))$. From this follows that $\text{supp}(\mu_A) = X$.*

Proof. The proof follows basically the same reasoning of the proof of Theorem 1.

Part a) For each $s \in (0, 1)$, we consider the operator $\mathcal{T}_{s,A} : \mathcal{C}_b(X) \rightarrow \mathcal{C}_b(X)$ given by

$$\mathcal{T}_{s,A}(u)(x) = \log \left(\int_{v \in T^{-1}(\{x\})} e^{A(v) + su(v)} dm(v) \right).$$

In the same way as before, for all $s \in (0, 1)$, the map $\mathcal{T}_{s,A}$ has a unique fixed point $u_s \in \mathcal{C}_b(X)$. That is, for all x and $0 < s < 1$

$$e^{u_s(x)} = \int_{v \in T^{-1}(\{x\})} e^{A(v) + s u_s(v)} dm(v) .$$

Following reasoning of the proof of Theorem 1 we need to prove that the family of bounded continuous functions is equicontinuous $(u_s)_{0 < s < 1}$.

In the same way as was done before, we have $|u_s(x') - u_s(y)| \leq \mathcal{S}_1(x, y)$, where

$$\mathcal{S}_1(x, y) = \sup_{z \in \text{Ker}(T)} \left\{ A(x' + z) - A(y' + z) + s(u_s(x' + z) - u_s(y' + z)) \right\} .$$

For elements x', y' in X , such that, $T(x') = x$ and $T(y') = y$, take

$$\begin{aligned} \mathcal{S}_n(x, y) = \sup_{\substack{z_j \in \text{Ker}(T^j) \\ 1 \leq j \leq n}} \left\{ \left(\sum_{j=1}^n s^{j-1} (A(x_j + z_j) - A(y_j + z_j)) \right) \right. \\ \left. + s^n (u_s(x_n + z_n) - u_s(y_n + z_n)) \right\} , \end{aligned}$$

where x_j and y_j , $1 \leq j \leq n$, are fixed elements of X satisfying

$$T^j(x_j) = x, \quad T^j(y_j) = y \text{ for all } 1 \leq j \leq n .$$

The sequence $(\mathcal{S}_n(x, y))_{n \geq 1}$ satisfies

$$\begin{aligned} \mathcal{S}_n(x, y) &\leq \sup_{\substack{z_j \in \text{Ker}(T^j) \\ 1 \leq j \leq n}} \left\{ \left(\sum_{j=1}^n s^{j-1} \text{Lip}_{A, D_X^\alpha} \|x_j - y_j\|_X^\alpha + 2s^n \|u_s\|_\infty \right) \right\} \\ &\leq \left(\sum_{j=1}^n s^{j-1} \text{Lip}_{A, D_X^\alpha} \frac{1}{p(T^j)^\alpha} \|x - y\|_X^\alpha \right) + 2s^n \|u_s\|_\infty \\ &\leq \left(\sum_{j=1}^n \text{Lip}_{A, D_X^\alpha} \frac{1}{p(T^j)^\alpha} \|x - y\|_X^\alpha \right) + 2s^n \|u_s\|_\infty . \end{aligned}$$

We deduce that for all $s \in (0, 1)$ and each pair $x, y \in X$ is satisfied

$$|u_s(x) - u_s(y)| \leq \left(\text{Lip}_{A, D_X^\alpha} \sum_{n=1}^{\infty} \frac{1}{p(T^n)^\alpha} \right) \|x - y\|_X^\alpha .$$

Besides that, by (22), it follows that

$$|u_s(x) - u_s(y)| \leq V_T(A) .$$

This is the end of the proof of item a).

Part b) This part follows exactly the same reasoning followed in the proof of Theorem 1 and will not be presented.

Part c) This part is a consequence of a) and b). Indeed, for any $A \in \text{Lip}_b(D_X^\alpha, X) \cap \mathcal{SV}_T(X)$ and any $x \in X$ we have

$$\begin{aligned} \mathcal{L}_{\bar{A}}(1)(x) &= \int_{\text{Ker}(T)} e^{\bar{A}(x'+z)} dm(z), \quad x' \in T^{-1}(\{x\}) \\ &= \frac{1}{\lambda_A \psi_A(x)} \int_{\text{Ker}(T)} e^{A(x'+z)} \psi_A(x'+z) dm(z), \quad x' \in T^{-1}(\{x\}) \\ &= \frac{\mathcal{L}_A(\psi_A)(x)}{\lambda_A \psi_A(x)} \\ &= 1. \end{aligned}$$

Part d) Define $U(x; \epsilon) := \{y \in X, |\pi(y) - \pi(x)| \leq \epsilon, \text{ and } y - \pi(y) = x - \pi(x)\}$, and let U be a neighborhood of x . There exists $\epsilon > 0$ such that the closed set $U(x; \epsilon)$ is contained in U . By Urysohn's Lemma, let $\varphi \in \mathcal{C}_b(X)$ such that $\chi_{U(x; \epsilon)} \leq \varphi \leq \chi_U$. Then,

$$\begin{aligned} \mu_A(U) &\geq \int_X \varphi d\mu_A \\ &= \int_X \mathcal{L}_A(\varphi) d\mu_A \\ &= \int_X \left(\int_{\text{Ker}(T)} e^{A(y'+z)} \varphi(y'+z) dm(z) \right) d\mu_A(y), \quad y' \in T^{-1}(\{y\}) \\ &\geq e^{\inf(A)} \int_X \left(\int_{\text{Ker}(T)} \chi_{U(x; \epsilon)}(y'+z) dm(z) \right) d\mu_A(y), \quad y' \in T^{-1}(\{y\}) \\ &\geq e^{\inf(A)} m(\epsilon B_{\text{Ker}(T)} + \{\pi(x)\}) > 0. \end{aligned}$$

The above implies that $\mu_A(U) > 0$, for any open neighborhood $U \subset X$ of x . moreover, since $\mu_A(U) \geq e^{\inf(A)} m(\epsilon B_{\text{Ker}(T)} + \{\pi(x)\})$, we have $\text{supp}(\mu_A) \supset \pi^{-1}(\text{supp}(m))$. \square

It follows from the above that:

Corollary 2. *Consider X a separable Banach space and $T : X \rightarrow X$ a bounded linear operator such that T is onto, not injective, satisfying (21), $p(T^n) > 0$ and $\dim(\text{Ker}(T^n)) < \infty$, for each $n \in \mathbb{N}$. Consider $A \in \mathcal{SV}_T(X) \cap \text{Lip}_b(D_X^\alpha, X)$. Assume that the a priori probability measure m has strong adapted tails and that the following is true:*

$$\lim_{n \rightarrow \infty} (p(T^n))^{-\frac{1}{n}} < 1.$$

Then, items a), b), c) and d) of Theorem 2 are true.

In the next result, we prove that a sub class of the class of operators satisfying the claim of last Corollary is formed by frequently hypercyclic operators.

Proposition 2. Let X be a separable Banach space and $T : X \rightarrow X$ be a bounded linear operator, such that, T is onto, not one to one, satisfying (21) and $\bigcup_{n=0}^{\infty} \text{Ker}(T^n)$ is dense in X . Assume that

$$\lim_{n \rightarrow \infty} (p(T^n))^{-\frac{1}{n}} < 1 ,$$

then, T is frequently hypercyclic and Devaney chaotic.

Example 3. If $L : X \rightarrow X$ is a weighted shift and X is $c_0(\mathbb{R})$ or $l^p(\mathbb{R})$, $1 \leq p < \infty$. Then,

$$\bigcup_{n=0}^{\infty} \text{Ker}(L^n) = \{(x_n)_{n \geq 1} \in X, \exists n_0 \in \mathbb{N}, x_n = 0, \forall n \geq n_0\} .$$

Hence, $\bigcup_{n=0}^{\infty} \text{Ker}(L^n)$ is dense in X .

Remark 8. There is an efficient criterion which guarantees that T is Devaney chaotic and frequently hypercyclic (see [3]).

Let X be a separable Banach space and $T : X \rightarrow X$ be a continuous linear operator. Assume that there exist a dense set $D \subset X$ and a map $S : D \rightarrow D$, such that,

1. For any $x \in D$, the series $\sum_{n=0}^{\infty} T^n(x)$ and $\sum_{n=0}^{\infty} S^n(x)$ are unconditionally convergent (all sub-series of both series are convergent).
2. For every $x \in D$, $T \circ S(x) = x$.

Then, T is Devaney chaotic and frequently hypercyclic.

Proof of Proposition 2. Set $D = \bigcup_{n=0}^{\infty} \text{Ker}(T^n)$, then for all $x \in D$, there exists a non-negative integer $n_0 = n_0(x)$, such that,

$$T^n(x) = 0, \text{ for all } n \geq n_0 .$$

Hence, the series $\sum_{n=0}^{\infty} T^n(x)$ is unconditionally convergent.

On the other hand, since T is onto, then for any $x \in D \setminus \{0\}$, there exists $y \in X$ such that $T(y) = x$. Taking $S(x) = y$, we deduce by induction that

$$T^n \circ S^n(x) = x, \text{ for all } n \in \mathbb{N} .$$

Hence, by (19), it follows that

$$\|S^n(x)\|_X \leq \frac{\|x\|_X}{p(T^n)}, \text{ for all } n \geq 1 .$$

Since $\lim_{n \rightarrow \infty} (p(T^n))^{-\frac{1}{n}} < 1$, we deduce that the series $\sum_{n=0}^{\infty} S^n(x)$ is absolutely convergent and hence unconditionally convergent, and we are done. \square

5 Appendix

The purpose of this appendix is to prove the existence of a unique fixed point of the operator \mathcal{L}_A^* , when A is normalized. Throughout, we assume that $X = c_0(\mathbb{R})$ or $X = l^p(\mathbb{R})$, $1 \leq p < \infty$, and that $A \in \text{Lip}_b(D_X^\alpha, X)$ and $\mathcal{L}_A(1) = 1$. For this setting, our aim is to prove that the operator \mathcal{L}_A^* has a fixed point Borel probability measure on X . In order to do so, we are following the approach to spectral gaps for Markov operators by M. Hairer and J. C. Mattingly in [16], which was adapted to the dual of the Ruelle operator by several authors (see [27],[17],[19],[6],[9],[23],[28], in chronological order).

Remark 9. However, as a consequence of the metric structure of X , the adapted approach does not provide uniform contraction rates and, in particular, we only obtain convergence but not geometric convergence of the iterates of \mathcal{L}_A^* .

We now recall the strategy of the approach which relies on a change to the bounded and equivalent metric

$$\tilde{D}(x, y) := \min \{1, aD_X^\alpha(x, y)\}.$$

By choosing $a > 0$ properly, a version of the Doeblin-Fortet inequality then implies that $(\mathcal{L}_A^n)^*$ locally contracts the Wasserstein distance for some $n \in \mathbb{N}$. Moreover, a coupling construction then makes use of the boundedness of \tilde{D} in order to obtain a contraction for Dirac measures whose base points have \tilde{D} -distance bigger than 1. By joining these two results on a further coupling, one then obtains the desired contraction property.

We begin with a comparison of the spaces of Lipschitz continuous functions with respect to the change of metric. For ease of notation, set $D := D_X^\alpha$. Moreover, recall that $\text{Lip}^{\text{loc}(\delta)}(D, X)$ is the set of continuous functions $\varphi : X \rightarrow \mathbb{R}$ such that

$$\text{Lip}_{\varphi, D}^{\text{loc}(\delta)} := \sup \left\{ \frac{|\varphi(x) - \varphi(y)|}{D(x, y)} : x \neq y, D(x, y) < \delta \right\} < \infty.$$

Lemma 4. *A function φ is in $\text{Lip}(\tilde{D}, X)$ if and only if $\varphi \in \text{Lip}_b^{\text{loc}(1/a)}(D, X)$. Moreover, $a\text{Lip}_{\varphi, D}^{\text{loc}(1/a)} = \text{Lip}_{\varphi, \tilde{D}}^{\text{loc}(1)}$,*

$$a\text{Lip}_{\varphi, D}^{\text{loc}(1/a)} \leq \text{Lip}_{\varphi, \tilde{D}} \leq \max \left\{ 2\|\varphi\|_\infty, \frac{1}{a} \text{Lip}_{\varphi, D}^{\text{loc}(1/a)} \right\}, \quad (23)$$

and $\text{sup}(\varphi) - \text{inf}(\varphi) \leq \text{Lip}_{\varphi, \tilde{D}}$.

Proof. Note that $D(x, y) < 1/a$ if and only if $\tilde{D}(x, y) < 1$ which proves that $a\text{Lip}_{\varphi, D}^{\text{loc}(1/a)} = \text{Lip}_{\varphi, \tilde{D}}^{\text{loc}(1)}$.

Now assume that $\varphi \in \text{Lip}^{\text{loc}(1/a)}(D, X)$ is bounded. Then

$$\frac{|\varphi(x) - \varphi(y)|}{\tilde{D}(x, y)} = \begin{cases} \frac{|\varphi(x) - \varphi(y)|}{aD(x, y)} \leq \frac{1}{a} \text{Lip}_{\varphi, D}^{\text{loc}(1/a)} & : D(x, y) < 1/a \\ |\varphi(x) - \varphi(y)| \leq 2\|\varphi\|_{\infty} & : D(x, y) \geq 1/a \end{cases}$$

This proves that $\varphi \in \text{Lip}(\tilde{D}, X)$ and the left half of (23). The left half of (23) follows from the trivial estimate $\text{Lip}_{\varphi, \tilde{D}}^{\text{loc}(1)} \leq \text{Lip}_{\varphi, \tilde{D}}$ and the first observation.

Finally, as $\sup \tilde{D}(x, y) = 1$,

$$\sup(\varphi) - \inf(\varphi) \leq \text{Lip}_{\varphi, \tilde{D}} \sup_{x, y \in X} \tilde{D}(x, y) = \text{Lip}_{\varphi, \tilde{D}}.$$

□

In order to introduce the Wasserstein distance (see for details [8, 29]), recall that a coupling of $\mu, \nu \in \mathcal{P}(X)$ is a Borel probability measure Π on $X \times X$ such that the first marginal of Π is μ and the second marginal is ν . We refer to $\Gamma(\mu, \nu)$ as the set of all couplings of μ and ν . The D -Wasserstein distance on $\mathcal{P}(X)$ is then defined by

$$W_D(\mu, \nu) = \inf \left\{ \int_{X \times X} D(x, y) d\Pi(x, y) : \Pi \in \Gamma(\mu, \nu) \right\},$$

where D is a metric whose open balls generate the Borel σ -algebra. Moreover, if X has finite diameter with respect to D , then W_D is compatible with the weak convergence of measures, that is, μ_n converges weakly to μ if and only if $W_D(\mu_n, \mu) \rightarrow 0$. A further fundamental tool in this setting is Kantorovich's duality, which states that

$$W_D(\mu, \nu) = \sup \left\{ \left| \int_X \varphi d\mu - \int_X \varphi d\nu \right| : \text{Lip}_{\varphi, D} \leq 1 \right\}.$$

We are now in position to prove the basic contraction estimates for our setting. In order to do so, observe that $\mathcal{L}_A(1) = 1$ implies that

$$\mathcal{L}_A^*(\mu)(X) = \int_X 1 d\mathcal{L}_A^*(\mu) = \int_X \mathcal{L}_A(1) d\mu = 1 \text{ for any } \mu \in \mathcal{P}(X).$$

Hence, as A is a normalized potential, \mathcal{L}_A^* acts on $\mathcal{P}(X)$. Set

$$C_A := \sup \left\{ \frac{e^{\text{Lip}_{A, D} \sum_{i=1}^{\infty} d_i^{-\alpha} t} - 1}{t} : 0 < t \leq 1 \right\}$$

Lemma 5 (Local contraction for Dirac measures). *Assume that n is such that $(d_n)^{-\alpha} \leq \frac{3}{8}$ and that $a = \max\{8C_A/3, 1\}$. Then, for all $x, y \in X$ with $\tilde{D}(x, y) < 1$,*

$$W_{\tilde{D}}((\mathcal{L}_A^*)^n(\delta_x), (\mathcal{L}_A^*)^n(\delta_y)) \leq \frac{3}{4} W_{\tilde{D}}(\delta_x, \delta_y).$$

Proof. As A is normalized, the arguments in the proof of Lemma 1 simplify as follows. Assume that $a \geq 1$, $x, y \in X$ with $\tilde{D}(x, y) < 1$ and that $\varphi \in \text{Lip}(\tilde{D}, X)$ with $\inf(\varphi) = 0$. Then, by Lemma 4,

$$\begin{aligned}
& |\mathcal{L}_A^n(\varphi)(x) - \mathcal{L}_A^n(\varphi)(y)| \\
& \leq \int e^{S_n(A) \circ \tau_a(x)} |\varphi \circ \tau_a(x) - \varphi \circ \tau_a(y)| dm_n(a) \\
& \quad + \int \left| e^{S_n(A) \circ \tau_a(x)} - e^{S_n(A) \circ \tau_a(y)} \right| |\varphi \circ \tau_a(y)| dm_n(a) \\
& \leq \text{Lip}_{\varphi, \tilde{D}} \tilde{D}(\tau_a(x), \tau_a(y)) \\
& \quad + \sup_a \left| e^{S_n(A) \circ \tau_a(x) - S_n(A) \circ \tau_a(y)} - 1 \right| \int e^{S_n(A) \circ \tau_a(y)} |\varphi \circ \tau_a(y)| dm_n(a) \\
& \leq \text{Lip}_{\varphi, \tilde{D}} \tilde{D}(\tau_a(x), \tau_a(y)) + \sup_a \left| e^{S_n(A) \circ \tau_a(x) - S_n(A) \circ \tau_a(y)} - 1 \right| \|\varphi\|_\infty \\
& \leq d_n^{-\alpha} \text{Lip}_{\varphi, \tilde{D}} \tilde{D}(x, y) + \left(e^{\text{Lip}_{A, D} \sum_{i=1}^\infty d_i^{-\alpha} D(x, y)} - 1 \right) \text{Lip}_{\varphi, \tilde{D}} \\
& \leq \text{Lip}_{\varphi, \tilde{D}} \left(d_n^{-\alpha} \tilde{D}(x, y) + C_A D(x, y) \right) = \text{Lip}_{\varphi, \tilde{D}} \left(d_n^{-\alpha} + \frac{C_A}{a} \right) \tilde{D}(x, y).
\end{aligned}$$

Observe that $\mathcal{L}_A(1) = 1$ implies that

$$\mathcal{L}_A(\varphi + c)(x) - \mathcal{L}_A(\varphi + c)(y) = \mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y).$$

Hence, by combining Kantorovich's duality with the above estimate,

$$\begin{aligned}
& W_{\tilde{D}}((\mathcal{L}_A^*)^n(\delta_x), (\mathcal{L}_A^*)^n(\delta_y)) \\
& = \sup \left\{ |\mathcal{L}_A^n(\varphi)(x) - \mathcal{L}_A^n(\varphi)(y)| : \text{Lip}_{\varphi, \tilde{D}} \leq 1 \right\} \\
& \leq \left(d_n^{-\alpha} + \frac{C_A}{a} \right) \tilde{D}(x, y) = \left(d_n^{-\alpha} + \frac{C_A}{a} \right) W_{\tilde{D}}(x, y).
\end{aligned}$$

The result follows from this and the above choices for a and n . \square

We now analyze the action of \mathcal{L}_A^* on δ_x and δ_y for $\tilde{D}(x, y) = 1$.

Lemma 6 (Global contraction for Dirac measures). *If $\tilde{D}(x, y) = 1$ and n is chosen such that $aD(x, y) < d_n^\alpha$, then*

$$W_{\tilde{D}}((\mathcal{L}_A^*)^n(\delta_x), (\mathcal{L}_A^*)^n(\delta_y)) \leq 1 - e^{-\sum_{i=1}^\infty d_i^{-\alpha} D(x, y)} (1 - a d_n^{-\alpha} D(x, y)).$$

Proof. Define

$$R_n := \int e^{\min\{S_n(A) \circ \tau_a(x), S_n(A) \circ \tau_a(y)\}} \delta_{(\tau_a(x), \tau_a(y))} dm_n,$$

which is a measure on $X \times X$ such that, for any Borel set A and $\pi_i(x_1, x_2) := x_i$,

$$\begin{aligned}
R_n(\pi_i^{-1}(A)) &= \int 1_A \circ \tau_a(\pi_i(x, y)) e^{\min\{S_n(A) \tau_a(x), S_n(A) \tau_a(y)\}} dm_n \\
&\leq \mathcal{L}_A^n(1_A)(\pi_i(x, y)) = (\mathcal{L}_A^*)^n(\delta_{\pi_i(x, y)})(A).
\end{aligned}$$

As it is well known from the theory of couplings, this estimate implies that R_n can be extended to an element in $\Gamma((\mathcal{L}_A^*)^n(\delta_x), (\mathcal{L}_A^*)^n(\delta_y))$. Or, by a straightforward calculation, one might check that the measure

$$Q_n := R_n + \frac{((\mathcal{L}_A^*)^n(\delta_x) - R_n \circ \pi_1^{-1}) \otimes ((\mathcal{L}_A^*)^n(\delta_y) - R_n \circ \pi_2^{-1})}{1 - R_n(X \times X)}$$

in fact is such a coupling. Now observe that $|S_n(A) \circ \tau_a(x)(x) - S_n(A) \circ \tau_a(y)| \leq CD(x, y)$ for $C := \sum_{i=1}^{\infty} d_i^{-\alpha}$. Hence, for

$$\Delta_\epsilon := \{(x, y) : \tilde{D}(x, y) \leq \epsilon\}$$

and $\epsilon = d_n^{-\alpha} D(x, y)$, we have that

$$\begin{aligned} Q_n(\Delta_\epsilon) &\geq R_n(\Delta_\epsilon) \\ &\geq \int e^{\min\{S_n(A) \circ \tau_a(x), S_n(A) \circ \tau_a(y)\}} dm_n \\ &\geq \int e^{S_n(A) \circ \tau_a(x)} dm_n e^{-CD(x, y)}. \end{aligned}$$

Hence, if n is chosen such that $\epsilon = d_n^{-\alpha} D(x, y) < 1/a$,

$$\begin{aligned} &W_{\tilde{D}}((\mathcal{L}_A^*)^n(\delta_x), (\mathcal{L}_A^*)^n(\delta_y)) \\ &\leq \int \tilde{D}(x, y) dQ_n(x, y) \\ &= \int_{\Delta_\epsilon^c} \tilde{D}(x, y) dQ_n(x, y) + \int_{\Delta_\epsilon} \tilde{D}(x, y) dQ_n(x, y) \\ &\leq 1 - Q_n(\Delta_\epsilon) + a\epsilon Q_n(\Delta_\epsilon) \\ &= 1 - Q_n(\Delta_\epsilon)(1 - a\epsilon) \\ &\leq 1 - e^{-CD(x, y)}(1 - ad_n^{-\alpha} D(x, y)). \end{aligned}$$

The assertion follows when including the case $\epsilon \geq 1/a$ for completeness. \square

In order to employ Lemma 5 and 6 for getting estimates for arbitrary measures, we recall the following basic fact: assume that $n \in \mathbb{N}$ and $Q_{\xi, \eta} \in \Gamma((\mathcal{L}_A^*)^n(\delta_\xi), (\mathcal{L}_A^*)^n(\delta_\eta))$, are such that, $(\xi, \eta) \mapsto \int \tilde{D}(x, y) dQ_{\xi, \eta}(x, y)$ is measurable, then, for any $\mu, \nu \in \mathcal{P}(X)$ and any $\Pi \in \Gamma(\mu, \nu)$, we get that

$$dQ_{\xi, \eta}(x, y) d\Pi(\xi, \eta) \in \Gamma((\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu)).$$

And, in particular, by taking the infimum over all possible couplings $Q_{\xi, \eta}$,

$$W_{\tilde{D}}((\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu)) \leq \int W_{\tilde{D}}((\mathcal{L}_A^*)^n(\delta_x), (\mathcal{L}_A^*)^n(\delta_y)) d\Pi(x, y). \quad (24)$$

This allows us to summarize Lemma 5 and 6 into a single estimate. In order to do so, set

$$r_n(t) := \begin{cases} 1 - e^{-\sum_{i=1}^{\infty} d_i^{-\alpha} t} (1 - \min\{1, tad_n^{-\alpha}\}) & : t \geq \frac{1}{a} \\ \frac{3}{4} & : t < \frac{1}{a} \end{cases}$$

Lemma 7. *Assume that $\mu, \nu \in \mathcal{P}(X)$ and that $\Pi \in \Gamma(\mu, \nu)$. Then, for any n such that $(d_n)^\alpha \geq 8/3$,*

$$W_{\tilde{D}}(\mathcal{L}_A^*(\mu), (\mathcal{L}_A^*)^n(\nu)) \leq \int r_n(D(x, y)) \tilde{D}(x, y) d\Pi(x, y).$$

Proof. By dividing the domain of integration in (24) into $\{(x, y) : D(x, y) < 1/a\}$, $\{(x, y) : 1/a \leq D(x, y) < d_n^\alpha/a\}$ and $\{(x, y) : D(x, y) \geq d_n^\alpha/a\}$, the first immediately follows by application of Lemma 5 and 6. \square

As an application, we obtain uniform contraction on certain subsets of $\mathcal{P}(X)$ which we define now. For $\gamma > 0$ and $\epsilon > 0$, set

$$\mathcal{P}_{\gamma, \epsilon}(X) := \{\mu \in \mathcal{P}(X) : \mu(\{x \in X : D(x, 0) > \gamma\}) < \epsilon\}.$$

Proposition 3. *Assume that $\mu, \nu \in \mathcal{P}(X)$ and that $\gamma \geq 1/a$ is chosen such that $\mu, \nu \in \mathcal{P}_{\gamma, \epsilon}(X)$, for $\epsilon := W_{\tilde{D}}(\mu, \nu)/4$. Then, for any n with $(d_n)^\alpha \geq 8/3$,*

$$W_{\tilde{D}}((\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu)) \leq \left(1 - \frac{1 - \min\{1, 2a\gamma d_n^{-\alpha}\}}{2e^{2\gamma \sum_{i=1}^\infty d_i^{-\alpha}}}\right) W_{\tilde{D}}(\mu, \nu).$$

Proof. By the triangle property, $D(x, y) \leq 2\gamma$ whenever $D(x, 0), D(y, 0) \leq \gamma$. Hence,

$$\{D(x, y) > 2\gamma\} \subset (\{x : D(x, 0) > \gamma\} \times X) \cup (X \times \{y : D(0, y) > \gamma\}).$$

Therefore, for any $\Pi \in \Gamma(\mu, \nu)$, $\Pi(\{D(x, y) > 2\gamma\}) < 2\epsilon$. Now assume that Π is an optimal coupling, that is, $W_{\tilde{D}}(\mu, \nu) = \int \tilde{D} d\Pi$, and that $\gamma > 1/a$. Then, by Lemma 7,

$$\begin{aligned} & W_{\tilde{D}}(\mathcal{L}_A^*(\mu), (\mathcal{L}_A^*)^n(\nu)) \\ & \leq \int (1_{\{D(x, y) \leq 2\gamma\}} r_n(2\gamma) + 1_{\{D(x, y) > 2\gamma\}}) \tilde{D}(x, y) d\Pi \\ & = r_n(2\gamma) \int \tilde{D}(x, y) d\Pi + (1 - r_n(2\gamma)) \Pi(\{D(x, y) > 2\gamma\}) \\ & \leq \left(r_n(2\gamma) + \frac{2\epsilon(1 - r_n(2\gamma))}{W_{\tilde{D}}(\mu, \nu)}\right) W_{\tilde{D}}(\mu, \nu) \\ & = \left(1 - \frac{1 - r_n(2\gamma)}{W_{\tilde{D}}(\mu, \nu)} (W_{\tilde{D}}(\mu, \nu) - 2\epsilon)\right) W_{\tilde{D}}(\mu, \nu) \\ & = \left(1 - \frac{1 - r_n(2\gamma)}{2}\right) W_{\tilde{D}}(\mu, \nu). \end{aligned}$$

This proves the assertion. \square

As an immediate consequence, one obtains that for any $\mu, \nu \in \mathcal{P}(X)$, there exists n such that the distance between $(\mathcal{L}_A^*)^n(\mu)$ and $(\mathcal{L}_A^*)^n(\nu)$ is strictly contracted, even though the contraction rate depends on the distance as well as

the tails of μ and ν . In particular, $L = \lim W_{\tilde{D}}((\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu))$ exists. In order to guarantee that $L = 0$ and that $((\mathcal{L}_A^*)^n(\mu))_n$ converges, we now employ the notion of adapted tails as introduced in Definition 4. Recall that we say that m has adapted tails if, for any $\epsilon > 0$ there exists a sequence of positive numbers (κ_n) such that

1. $\sum_{n=1}^{\infty} m(\mathbb{R} \setminus [-\beta_1^n \kappa_n, \beta_1^n \kappa_n]) < \epsilon$,
2. the sequence (κ_n) is in X .

As shown below, this condition allows to control the tails of $(\mathcal{L}_A^*)^n(\mu)$ asymptotically. Before giving examples, we prove the main result of this appendix.

Theorem 3. *Assume that $X = c_0(\mathbb{R})$ or $X = l^p(\mathbb{R})$, $1 \leq p < \infty$ and that $L : X \rightarrow X$ a weighted shift satisfying (7). Moreover assume that $A \in \text{Lip}_b(D_X^\alpha, X)$ with $\mathcal{L}_A(1) = 1$ and that m has adapted tails. Then, for any $\nu \in \mathcal{P}(X)$, the limit*

$$\mu := \lim_{n \rightarrow \infty} (\mathcal{L}_A^*)^n(\nu)$$

exists in $\mathcal{P}(X)$. Moreover, μ is the unique element in $\mathcal{P}(X)$ such that

$$\mathcal{L}_A^*(\mu) = \mu.$$

Proof. As the first step in the proof, we derive a formula for the asymptotic tail of $(\mathcal{L}_A^*)^n(\nu)$. In order to do so, for $(\kappa_j) \in X$, $n \in \mathbb{N}$ and $K > 0$ given, set

$$\Omega := \{(x_j) : x_j \in [-\kappa_j, \kappa_j] \text{ for } 0 \leq j < n, \|(x_n, x_{n+1}, \dots)\|_X \leq K\}.$$

Then, for $\nu \in \mathcal{P}(X)$ and $\beta_1^0 := 1$,

$$\begin{aligned} & (\mathcal{L}_A^*)^n(\nu)(\Omega^c) \\ &= (\mathcal{L}_A^*)^n(\nu) \left(\bigcup_{j=0}^{n-1} \{(x_i) : |x_j| > \kappa_j\} \cup \{(x_i) : \|(x_n, x_{n+1}, \dots)\|_X > K\} \right) \\ &\leq \sum_{j=0}^{n-1} (\mathcal{L}_A^*)^n(\nu) (\{|x_j| > \kappa_j\}) + (\mathcal{L}_A^*)^n(\nu) (\{\|(x_n, x_{n+1}, \dots)\|_X > K\}) \\ &= \int \mathcal{L}_A^n \left(\sum_{j=0}^{n-1} 1_{\{|x_0|/\beta_1^j > \kappa_j\}} \circ L^j + 1_{\{\|(x_n/\beta_1^n, x_{n+1}/\beta_2^n, \dots)\|_X > K\}} \circ L^n \right) d\nu \\ &= \int \mathcal{L}_A^{n-j} \left(1_{\{|x_0| > \beta_1^j \kappa_j\}} \right) d\nu + \int 1_{\{\|(\frac{x_n}{\beta_1^n}, \frac{x_{n+1}}{\beta_2^n}, \dots)\|_X > K\}} d\nu \\ &\leq \sum_{j=0}^{n-1} \int \mathcal{L}_A^{n-j} \left(1_{\{|x_0| > \beta_1^j \kappa_j\}} \right) d\nu + \nu(\{(x_i) : \|(x_i)\| > d_n K\}). \end{aligned}$$

As A is bounded, it follows that $e^A = C^{\pm 1}$. In particular, $\mathcal{L}_A(1_B) = C^{\pm 1}m(B)$ for any Borel set B which only depends on the first coordinate. Hence, if (κ_j)

is chosen for $\epsilon > 0$ according to Definition 4, then, as A is normalized,

$$\begin{aligned} (\mathcal{L}_A^*)^n(\nu)(\Omega^c) &\leq C \sum_{k=0}^{n-1} m(\{x \in \mathbb{R} : |x| > \beta_1^j \kappa_j\}) + \nu(\{(x_i) : \|(x_i)\| > d_n K\}) \\ &\leq C\epsilon + \nu(\{(x_i) : \|(x_i)\| > d_n K\}) \xrightarrow{n \rightarrow \infty} C\epsilon. \end{aligned} \quad (25)$$

In the second step, we show that $W_{\bar{D}}((\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu)) \rightarrow 0$ as $n \rightarrow \infty$. So assume that $\mu, \nu \in \mathcal{P}(X)$. By Lemma 7, $W_{\bar{D}}((\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu))$ is decreasing in n and, in particular,

$$L := \lim_{n \rightarrow \infty} W_{\bar{D}}((\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu))$$

exists. Now assume that $L > 0$. Then, as m is adapted, there exists $(\kappa_j) \in X$ with $\sum_{j=1}^{\infty} m([- \beta_1^j \kappa_j, \beta_1^j \kappa_j]^c) < L/(8C)$. For $K := \|(\kappa_j)\|_X$, the estimate (25) combined with $d_n \rightarrow \infty$ implies that there exists $n_0 \in \mathbb{N}$ such that $(\mathcal{L}_A^*)^n(\nu)(\Omega^c) < L/4$ for any $n \geq n_0$. However, as Ω for these choices of n and K satisfies $\Omega \subset \{x \in X : \|x\|_X \leq 2\|(\kappa_i)\|_X\}$, it follows that

$$(\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu) \in \mathcal{P}_{2\|(\kappa_i)\|_X, L/4}(X) \text{ for all } n \geq n_0.$$

We now apply Proposition 3 to the pair $((\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu))$. Namely, for k chosen such that the factor on the right hand side in Proposition 3 is strictly smaller than 1, it follows that

$$L = \lim_{n \rightarrow \infty} W_{\bar{D}}((\mathcal{L}_A^*)^{n+k}(\mu), (\mathcal{L}_A^*)^{n+k}(\nu)) < \lim_{n \rightarrow \infty} W_{\bar{D}}((\mathcal{L}_A^*)^n(\mu), (\mathcal{L}_A^*)^n(\nu)) = L,$$

which is absurd. Hence, $L = 0$ and the assertion of the second step is proven.

In the third step, we use (25) to obtain that $(\mathcal{L}_A^*)^n(\delta_0)$ is tight, where 0 is the origin in X . We therefore recall the notion of tightness and Prokhorov's theorem. Namely, a sequence $(\mu_n : n \in \mathbb{N})$ of probability measures is tight if for any $\epsilon > 0$ there exists a compact set K such that $\mu_n(K) > 1 - \epsilon$ for any $n \in \mathbb{N}$. Prokhorov's theorem then states that tightness and sequential compactness are equivalent.

Now assume that $\epsilon > 0$ and that (κ_j) is chosen according to Definition 4, and set $\Omega_\epsilon := \{(x_i) \in X : |x_i| \leq \kappa_i \forall i \geq 0\}$. As $0 \in \Omega_\epsilon$ the estimate (25) implies that

$$(\mathcal{L}_A^*)^n(\delta_0)(\Omega_\epsilon) \leq C\epsilon.$$

Moreover, as it easily can be shown, Ω_ϵ is compact and therefore, $(\mathcal{L}_A^*)^n(\delta_0)$ is tight. Hence, by Prokhorov's theorem, there exists a subsequence (n_k) with $n_k \nearrow \infty$ such that $\mu := \lim_{k \rightarrow \infty} (\mathcal{L}_A^*)^{n_k}(\delta_0)$ exists.

In the last step, we prove for any $\nu \in \mathcal{P}(X)$ that $\mu = \lim_{n \rightarrow \infty} (\mathcal{L}_A^*)^n(\nu)$ and that $\mu = \mathcal{L}_A^*(\mu)$. In order to do so, we apply Step 2 to the measures δ_0 and $\mathcal{L}_A^*(\delta_0)$, which implies by continuity of \mathcal{L}_A^* that

$$\mu = \lim_{k \rightarrow \infty} (\mathcal{L}_A^*)^{n_k}(\delta_0) = \lim_{k \rightarrow \infty} (\mathcal{L}_A^*)^{n_k+1}(\delta_0) = \mathcal{L}_A^*(\mu).$$

Finally, the remaining assertion follows from a further application of Step 2 to ν and μ . \square

We now provide some examples of adapted measures. Therefore recall that a measure on \mathbb{R} has polynomial tails of order γ if there exist $C > 0$ and $\gamma > 1$ such that

$$m(\{x : |x| > z\}) \leq Cz^{-\gamma} \text{ for all } z > 0.$$

Hence, for class of measures and $\kappa_n := Bn^\ell/d_n$, for some $B > 0$ and $\ell > \gamma^{-1}$, it follows that

$$\begin{aligned} \sum_{n=1}^{\infty} m(\{x : |x| > \kappa_n \beta_1^n\}) &\leq \sum_{n=1}^{\infty} m(\{x : |x| > \kappa_n d_n\}) \leq C \sum_{n=1}^{\infty} L^{-\gamma} n^{-\ell\gamma} \\ &< \frac{C\ell\gamma}{\ell\gamma - 1} L^{-\ell\gamma} \xrightarrow{L \rightarrow \infty} 0. \end{aligned}$$

Hence, m is adapted whenever $(n^\ell/d_n) \in X$ for some $\gamma^{-1} < \ell \leq 1$. The following slightly more specific examples of adapted measures m with polynomial tails easily follows from this observation.

Proposition 4. *Assume that m has polynomial tails of order γ . Then the following holds.*

1. *If $n/d_n \rightarrow 0$ for some $\ell > \gamma^{-1}$ and $X = c_0(\mathbb{R})$, then m is adapted.*
2. *If $d_n > Cn^\ell$ for some $\ell > \gamma^{-1} + 1/p$, and $X = l^p(\mathbb{R})$ with $1 \leq p < \infty$, then m is adapted.*

In particular, if d_n grows exponentially, then any m with polynomial tails is adapted.

Moreover, we say that a measure on \mathbb{R} has exponential tails if there exist $C > 0$ and $0 < \gamma < 1$ such that

$$m(\{x : |x| > z\}) \leq C\gamma^z \text{ for all } z > 0.$$

By a similar argument as above, one obtains the following criterion.

Proposition 5. *Assume that m has exponential tails and that $((\log n)/d_n) \in X$. Then m is adapted.*

References

- [1] D. Aguiar, L. Cioletti, and R. Ruviano. A variational principle for the specific entropy for symbolic systems with uncountable alphabets. *Math. Nachrichten* 291 (17-18): 2506–2515, (2018).
- [2] A. T. Baraviera, L. M. Cioletti, A. O. Lopes, J. Mohr, and R. R. Souza. On the general one-dimensional XY model: positive and zero temperature, selection and non-selection. *Rev. Math. Phys.* 23 (10): 1063–1113, (2011).
- [3] F. Bayart and E. Matheron. *Dynamics of Linear Operators*. Cambridge University Press, Cambridge, (2009).

- [4] F. Bayart and I. Ruzsa, Difference sets and frequently hypercyclic weighted shifts. *Ergod. Th. Dynam. Sys.* (2015), 35, 691–709.
- [5] N. Bernardes and A. Messaoudi. Shadowing and Structural Stability in Linear Dynamical Systems. *Ergod. Th. Dynam. Sys.* (2020), published online: doi.org/10.1017/etds.2019.107.
- [6] M. Bessa and M. Stadlbauer. On the Lyapunov spectrum of relative transfer operators. *Stoch. Dyn.*, v. 16, n. 6, 1650024 (25 pages), 2016
- [7] A. Bourhim. *Bounded Point Evaluations and Local Spectral Theory*. Dissertation, Trieste, (2000) (arXiv:math/0008197v1).
- [8] F. Bolley. Separability and completeness for the Wasserstein distance. In *Seminaire de probabilités XLI*, volume 1934 of *Lecture Notes in Math.*, pages 371–377. Springer, Berlin, 2008
- [9] L. Cioletti, E. A. Silva and M. Stadlbauer. Thermodynamic formalism for topological Markov chains on standard Borel spaces. *Discrete Contin. Dyn. Syst.* 39(11), 6277–6298, 2019
- [10] P. Cirilo, B. Gollobit and E. Pujals. Generalized hyperbolicity for linear operators. *Adv. Math.* 387 (2021), 107830
- [11] R. Freire and V. Vargas. Equilibrium states and zero temperature limit on topologically transitive countable Markov shifts. *Trans. Amer. Math. Soc.*, 370 (2018), 8451–8465.
- [12] K.-G. Grosse-Erdmann. Hypercyclic and chaotic weighted shifts. *Studia Math.* 139, (2000), no. 1, 47–68.
- [13] K.-G. Grosse-Erdmann and A. Peris Manguillot. *Linear Chaos*. Springer-Verlag, London, (2011).
- [14] S. Grivaux. Frequently hypercyclic operators with irregularly visiting orbits. *J. Math. Anal. Appl.* 462, no. 1, 542–553 (2018).
- [15] S. Grivaux and E. Matheron. Invariant measures for frequently hypercyclic operators. *Adv. Math.* 265, 371–427 (2014).
- [16] M. Hairer and J. C. Mattingly. Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations. *Ann. Probab.* v. 36, n. 6, 20502091, 2008
- [17] B. Kloeckner, A. O. Lopes, and M. Stadlbauer. Contraction in the Wasserstein metric for some Markov chains, and applications to the dynamics of expanding maps. *Nonlinearity* v.28, N. 11, 4117-4137, 2015.
- [18] B. Kloeckner. An optimal transportation approach to the decay of correlations for non-uniformly expanding maps. *Ergodic Theory Dynam. Systems*, 1-37, 2018.

- [19] A. O. Lopes, An introduction to Coupling, in “Modeling, Dynamics, Optimization and Bioeconomics II”, Editors: Alberto Pinto and David Zilberman, Springer Proceedings in Mathematics and Statistics, pp 307-335, (2017) Springer Verlag
- [20] A. O. Lopes, J. K. Mengue, J. Mohr, and R. R. Souza. Entropy and variational principle for one-dimensional lattice systems with a general *a priori* probability: positive and zero temperature. *Ergodic Theory Dynam. Systems*, 35 (6): 1925–1961, (2015).
- [21] A. O. Lopes and V. Vargas. Gibbs States and Gibbsian Specifications on the space $\mathbb{R}^{\mathbb{N}}$. *Dyn. Systems: an Int. Jour.*, Volume 35, Issue 2, 216–241 (2020)
- [22] A. O. Lopes and V. Vargas, Entropy, pressure, ground states and calibrated sub-actions for linear dynamics, arXiv (2021)
- [23] E. G. C. Mesquita, Formalismo termodinamico em espaços não-compactos via Transporte Ótimo, PhD thesis, UNB, Brasilia (2019)
- [24] S. Muir and M. Urbanski. Thermodynamic formalism for a modified shift map, *Stoch. and Dyn.*, Vol. 14, No. 02, 1350020 (2014)
- [25] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Asterisque*. (187-188): 1–268 (1990).
- [26] O. Sarig, Thermodynamic formalism for countable Markov shifts, in Hyperbolic Dynamics, Fluctuations and Large Deviations, Proc. Sympos. Pure Math., 89, Amer. Math. Soc., Providence, RI, 2015, 81–117.
- [27] M. Stadlbauer, Coupling methods for random topological Markov chains, *Ergodic Theory Dynam. Systems*, Vol. 37, Issue 3, 971–994 (2017)
- [28] M. Stadlbauer, P. Varandas and X. Zhang Quenched and annealed equilibrium states for random Ruelle expanding maps and applications. arXiv:2004.04763, 2020.
- [29] C. Villani, Topics in optimal transportation, AMS, Providence (2003)
- [30] C. Villani, Optimal transport, Grundlehren der mathematischen Wissenschaften Series 338, Springer-Verlag, 2009.