Invariant probabilities for discrete time linear dynamics via thermodynamic formalism

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

June 22, 2020

Abstract: We show the existence of invariant ergodic $\sigma$-additive probability measures with full support on $X$ for a certain class of weighted shift linear operators $L : X \to X$, where $X$ is either the Banach space $c_0(\mathbb{R})$ or $l^p(\mathbb{R})$ for $1 \leq p < \infty$. In order to do so, we will adapt ideas from thermodynamic formalism. For a given bounded Hölder continuous potential $A : X \to \mathbb{R}$, we associate a transfer transfer operator $L_A$ (depending on the linear operator $L$), which acts on continuous functions on $X$, and prove that this operator satisfies a Ruelle-Perron-Frobenius theorem. Moreover, the invariant ergodic probability measure for $L$ obtained from this theorem is a Gibbs probability for the potential $A$. In order to prove our main result, we will consider the action of the dual of the Ruelle operator on the 1-Wasserstein space of probabilities on $X$.

We point out that in order to define $L_A$ one requires the use of an a priori probability measure defined on the kernel of $L$. These results are extended to a wide class of linear operators with non-trivial kernel defined on a separable Banach space.

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


1 Introduction

The main goal in here is to show existence of invariant $\sigma$-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.

*IME - UFRGS. Partially supported by CNPq
†MAT - UNESP. Partially supported by CNPq project 311018/2018-1 and by FAPESP project 2013/24541-0
‡IM - UFRJ. Partially Supported by CNPq
§IME - UFRGS. Supported by INCTMat-CAPES grant
In here, there the dynamics is obtained from a linear 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 as a tool for the construction of equilibrium 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 (see for instance [2, 22, 24, 10, 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, 14].

Let $X$ be a topological vector space and $T : X \to X$ be a linear continuous operator. We say that $(X,T)$ - or simply the map $T$ - is 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 $\liminf_{n \to \infty} \frac{1}{n} \#(N(x,V) \cap \{1, ..., n\}) > 0$. 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 $\alpha L$, $\alpha > 1$, 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 $\alpha L$, $\alpha > 1$, 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], [14], [6]).

2
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 [16]). We point out that in there, the existence of ergodic measures with full support are 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 $\lim_{n \to \infty} x_n = 0$ equipped with the sup-norm. For a fixed operator $T : X \to X$, a Hölder continuous potential $A : X \to \mathbb{R}$ 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 $A$ which will be $T$-invariant, ergodic and has full support.

Our first result is Theorem 1. From this we will obtain a $\sigma$-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 in order to get a fixed point probabilities for the dual of the Ruelle operator (see [27, 7, 18, 19, 20, 23]). However, the arguments in here are more involved due to the metric structure of $X$. For example, in contrast to shift spaces (even with respect to uncountable, non-compact alphabets as in [10]), it seems to be impossible to obtain uniform contraction rates.

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 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.

2 Preliminaries

Let $X$ be a Banach space and $T : X \to X$ be a linear continuous operator. Hereafter, with exception of some specific cases, we will denote by $\| \cdot \|_X$ the norm for $X$.

We say that the map $T$ is positively expansive if there exists a constant $e > 1$ such that, for each point $\tilde{x} \in S_X = \{x \in X, \|x\|_X = 1\}$, there exists $n \in \mathbb{N}$ with
\[ \|T^n(\bar{x})\|_X \geq c. \] Recall also that \( T \) is said to be \textit{uniformly positively expansive} if there exist \( c > 1 \) and \( m \in \mathbb{N} \) such that, \( \|T^m(x)\|_X \geq c \), for all \( x \) in \( S_X \).

Let \( c_0(\mathbb{R}) \) be the set of real sequences \((x_n)_{n \geq 1}\) such that \( \lim_{n \to \infty} x_n = 0 \). 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, considering \( l^p(\mathbb{R}) \), \( 1 \leq p < \infty \) 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 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 of the form
\[
x = \sum_{k=1}^{\infty} x_ke_k,
\]
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 \textit{weighted shift} associated to the sequence \((\alpha_n)_{n \geq 1}\) is defined as the linear map \( L : X \to X \), 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
\[
L(e_1) = 0 \quad \text{and} \quad L(e_n) = \alpha_{n-1} e_{n-1}, \quad \text{for all} \quad n \geq 2.
\]
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 \textit{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

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

and

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

It is widely known (see for instance [8]), that the spectrum of $L : X \to X$, 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

$$r(L) := \lim_{n \to \infty} \|L^n\|_{op}^{\frac{1}{n}} = \lim_{n \to \infty} \left(\sup_{k \geq 1} \beta_k^n\right)^{\frac{1}{n}},$$  \hspace{1cm} (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 $\beta_k^n$ 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 $\limsup_{n \to \infty} \beta_1^n = \infty$ ([3], [14]).
2. $L$ is topologically mixing if and only if $\lim_{n \to \infty} \beta_1^n = \infty$ ([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 $\sum_{n=1}^{\infty} (\beta_1^n)^{-p} < \infty$ ([14]).
5. $L$ is positively expansive if and only if $\sup_{n \geq 1} \beta_1^n = \infty$ ([6]).
6. $L$ is uniformly positively expansive ([6]) if and only if

$$\sup_{n \geq 1} \left(\inf_{k \geq 1} \beta_k^n\right) = \infty.$$
For the separable Banach space $X$, we will use the notation $C(X)$ for the set of continuous functions from $X$ into $\mathbb{R}$ and $C_b(X)$ for the set of bounded continuous functions from $X$ into $\mathbb{R}$. As it is well known, $C_b(X)$ 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

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

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 $\alpha$-Hölder continuous functions from $X$ into $\mathbb{R}$ and the set $\text{Lip}_b(D^\alpha, X)$ is the set of bounded $\alpha$-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

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

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 $C_b(X)$, which we will denote by $C_b(X)^*$, 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 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, 21, 24, 1]). Even in the case of a non-compact alphabet $A$, one can get in specific cases similar results for $A^\mathbb{N}$ (see [22, 12, 10]).

Adapting some ideas from the classical setting (as for instance in [22]) 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 operator (see [15] and [11]) with a finite dimensional kernel.

We now return to the weighted shift $L : X \to X$ 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 $\mathcal{C}_0(X)$ is the defined in the following way. We fix an a priori probability measure $\mu$ on the Borel sets of $\mathbb{R}$ equipped with the usual topology and assume that the support of the probability measure $\mu$ is equal to the set $\mathbb{R}$. Since $\text{Ker}(L)$ is isometrically isomorphic with $\mathbb{R}$ it follows that $\mu$ induces an a priori probability measure on $\text{Ker}(L)$, which we will also denote by the same letter $\mu$.

As an example, one could take $\mu$ as the Gaussian distribution on $\mathbb{R}$ with mean zero and variance 1. That is $\mu = f \, dr$, 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

$$V(A) = \sum_{n=1}^{\infty} V_n(A) < \infty,$$  \tag{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 $\mathcal{S}(X)$ 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 [10]), there are bounded Lipschitz functions which are not in $\mathcal{S}(X)$. 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, $\text{Lip}_b(D_X^0, X) \cap \mathcal{S}(X)$ is non-empty which follows from the following construction. For $x, y \in X$, set

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

Note that $D_{\text{shift}}(x, y) \leq D_X(x, y)$ for any pair $x, y$ with $\sup|x_n - y_n| \leq 1$ and that, as $D_{\text{shift}}(x, y) \leq 1$ for any $x, y \in X$, each function in $\text{Lip}(D_{\text{shift}}^0, X)$ is bounded. Moreover, it follows from $\sum_n 2^{-n} < \infty$ that any function in $\text{Lip}(D_{\text{shift}}^0, X)$ has summable variations. Hence,

$$\text{Lip}(D_{\text{shift}}^0, X) \subset \text{Lip}_b(D_X^0, X) \cap \mathcal{S}(X).$$

Finally, as any function in $\text{Lip}_b(D_X^0, X)$ which only depends on a finite number of coordinates of $X$ is an element of $\text{Lip}(D_{\text{shift}}^0, X)$, the set of $D_X^0$-Lipschitz function with summable variations is non-trivial. These observations also provide an explanation of the phenomena that shift spaces with a standard Borel alphabet as in [10] show an exponential decay of correlations whereas this is not the case in the setting of linear dynamics. Namely, we obtain decay of correlations, but without a specific rate as a consequence of Theorem 2.
a) For any \( \phi \)

It is easy to verify that \( L_A \)

We say that a potential \( \phi \) to each function \( L \)

the next Lemma we will show that for any \( x,y \)

for any \( \kappa > 0 \), such that,

Consider \( L \) Lip

b) For any \( \phi \)

We now define the Ruelle operator. Given the \( a \) priori probability \( m \) and a potential \( A \in \text{Lip}_b(D^\alpha_X, X) \), the Ruelle operator is defined as the map assigning to each function \( \phi \in C_b(X) \) the function

\[
L_A(\phi)(x) := \int_{Lx=x} e^{A(x)} \varphi(\psi) \, dm(\psi) .
\]

(4)

Using the coordinates induced by the Schauder basis \( \{ e_k \}_k \geq 1 \) it is possible to write the equation (4) in the following way

\[
L_A(\phi)(x_1, x_2, \ldots) := \int_{\mathbb{R}} e^{A(r, \frac{r_1}{\alpha_1}, \frac{r_2}{\alpha_2}, \ldots)} \varphi(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots) \, dm(r) .
\]

It is easy to verify that \( L_A(\phi) \in C_b(X) \) (see the proof of Lemma 1) and that the \( n \)-th iterate \( (n \geq 2) \) of the Ruelle operator is given by

\[
L^n_A(\phi)(x) = \int_{\mathbb{R}^n} e^{\sum_{i=1}^n A(r, \frac{r_{i-1}}{\beta_{i-1}}, \frac{r_i}{\beta_i}, \ldots)} \varphi(r_1, \frac{r_{n-1}}{\beta_{n-1}}, \frac{r_n}{\beta_n}, \ldots) \, dm(r_1) \cdots dm(r_n).
\]

We say that a potential \( A : l^p(\mathbb{R}) \to \mathbb{R} \) is normalized in the case \( L_A(1) = 1 \). In the next Lemma we will show that for any \( A \in \text{Lip}_b(D^\alpha_X, X) \) the Ruelle operator \( L_A \) preserves the set of \( \alpha \)-Hölder continuous functions.

Note that in fact \( A \in \text{Lip}_b(D^\alpha_X, X) \) if and only if \( e^A \in \text{Lip}_b(D^\alpha_X, X) \), because for any \( x,y \in X \) is satisfied \( |A(x) - A(y)| \leq \sup(A) - \inf(A) < \infty \), which implies that there is \( \kappa > 0 \), such that,

\[
|e^{A(x)} - e^{A(y)}| \leq \kappa |A(x) - A(y)| .
\]

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

a) For any \( \phi \in \text{Lip}_b(D^\alpha_X, X) \), we have

\[
\text{Lip}_{L^n_A(\phi), D^\alpha_X} \leq \text{Lip}_{\phi, D^\alpha_X} e^{n} \| A \|_{\infty} (d_n)^{-\alpha} + \text{Lip}_{e^{A}, D^\alpha_X} \| \varphi \|_{\infty} \sum_{k=1}^{n} (d_k)^{-\alpha} .
\]

(5)

b) For any \( \phi \in \text{Lip}_b^{loc(\delta)}(D^\alpha_X, X) \), we have

\[
\text{Lip}_{L^n_A(\phi), D^\alpha_X}^{loc(\delta)} \leq \text{Lip}_{\phi, D^\alpha_X}^{loc(\delta)} e^{n} \| A \|_{\infty} (d_n)^{-\alpha} + \text{Lip}_{e^{A}, D^\alpha_X}^{loc(\delta)} \| \varphi \|_{\infty} \sum_{k=1}^{n} (d_k)^{-\alpha} .
\]

(6)
Proof. We will only prove a) as the proof of b) is similar. By definition, for any \( \varphi \in \text{Lip}_b(D_X^N , X) \) and each pair \( x,y \in X \),

\[
|\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)|
\]

\[
= \left| \int_{L_v=x} e^{A(v)} \varphi(v) dm(v) - \int_{L_v=y} e^{A(v)} \varphi(v) dm(v) \right|
\]

\[
= \left| \int_{\mathbb{R}} e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots)} \varphi(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots) dm(r) \right|
\]

\[
- \left| \int_{\mathbb{R}} e^{A(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \ldots)} \varphi(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \ldots) dm(r) \right|
\]

\[
\leq \int_{\mathbb{R}} e^{\|A\|_\infty} \left| \varphi(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots) - \varphi(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \ldots) \right| dm(r)
\]

\[
+ \|\varphi\|_{\infty} \left| e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots)} - e^{A(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \ldots)} \right| dm(r).
\]

Hence,

\[
|\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)|
\]

\[
\leq \left( e^{\|A\|_\infty} \text{Lip}_e, D_X^N + \|\varphi\|_{\infty} \text{Lip}_{e^A, D_X^N} \right) \int_{\mathbb{R}} \left( \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots \right) - \left( \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \ldots \right) \|x\|_X^D dm(r)
\]

\[
\leq \left( e^{\|A\|_\infty} \text{Lip}_e, D_X^N + \|\varphi\|_{\infty} \text{Lip}_{e^A, D_X^N} \right) \int_{\mathbb{R}} (d_1)^{-\alpha} \|x - y\|_X^\alpha dm(r)
\]

\[
= \left( e^{\|A\|_\infty} \text{Lip}_e, D_X^N + \|\varphi\|_{\infty} \text{Lip}_{e^A, D_X^N} \right) (d_1)^{-\alpha} \|x - y\|_X^\alpha.
\]

The above implies that \( \mathcal{L}_A(\varphi) \in \text{Lip}(D_X^N , X) \), with

\[
\text{Lip}_{\mathcal{L}_A(\varphi), D_X^N} \leq (d_1)^{-\alpha} \left( e^{\|A\|_\infty} \text{Lip}_e, D_X^N + \|\varphi\|_{\infty} \text{Lip}_{e^A, D_X^N} \right).
\]

That is, \( (5) \) holds for \( n = 1 \). The proof is by induction. Assume \( (5) \) is valid for \( n \). Then, we have

\[
|\mathcal{L}_A^{n+1}(\varphi)(x) - \mathcal{L}_A^{n+1}(\varphi)(y)|
\]

\[
= |\mathcal{L}_A^n(\mathcal{L}_A(\varphi))(x) - \mathcal{L}_A^n(\mathcal{L}_A(\varphi))(y)|
\]

\[
\leq \left( \text{Lip}_{\mathcal{L}_A(\varphi), D_X^N} e^{\|A\|_\infty} (d_n)^{-\alpha} + \|\varphi\|_{\infty} \sum_{k=1}^{n} (d_k)^{-\alpha} \right) \|x - y\|_X^\alpha
\]

\[
\leq \left( \text{Lip}_{\mathcal{L}_A(\varphi), D_X^N} e^{(n+1)\|A\|_\infty} (d_{n+1})^{-\alpha} + \|\varphi\|_{\infty} \sum_{k=1}^{n+1} (d_k)^{-\alpha} \right) \|x - y\|_X^\alpha,
\]

where the last estimate is a consequence of \( (7) \). The foregoing implies \( \mathcal{L}_A^n(\varphi) \in \text{Lip}(D_X^N , X) \) for each \( n \in \mathbb{N} \) and proves \( (5) \). Finally, we have for any \( x \in X \) that

\[
|\mathcal{L}_A^n(\varphi)(x)| \leq e^{\|A\|_\infty} \|\varphi\|_{\infty},
\]

9
that is, $L^n_A(\varphi) \in \text{Lip}(D^n_X, X)$, which concludes the proof of item a) of this Lemma.

The proof of item b) follows the same procedure under the assumption that $D_X(x, y)^\alpha < \delta$.

\[ \square \]

**Remark 3.** In the case when $A$ is a normalized potential, that is, $L^n_A(1) = 1$, we claim that for any $\varphi \in \text{Lip}^{loc}(\delta)_{b}(D^\alpha_X, X)$ the inequality (6) can be written as

\[ \text{Lip}^{loc}(\delta)_{\varphi, D_X} \sup_{k \in \mathbb{N}} (d_k)^{-\alpha} \leq \text{Lip}^{loc}(\delta)_{\varphi, D_X} \sup_{k \in \mathbb{N}} (d_k)^{-\alpha}. \]

This is so because in the above reasoning for estimating the value $|L^n_A(\varphi)(x) - L^n_A(\varphi)(y)|$,

we can alternatively write

\[ \left| \int_{\mathbb{R}} e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots)} \left( \varphi(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots) - \varphi(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \ldots) \right) dm(r) \right| \]

\[ \leq \text{Lip}^{loc}(\delta)_{\varphi, D_X} \left\| (\frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots) - (\frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, \ldots) \right\| \int_{\mathbb{R}} e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots)} dm(r) \]

\[ \leq \text{Lip}^{loc}(\delta)_{\varphi, D_X} (d_1)^{-\alpha} \left\| x - y \right\|_X^n. \]

The claim follows by induction on $n$.

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

**Remark 5.** We would like to remark now that the reason to consider the action of $L_A$ on functions which are locally Lipschitz continuous only 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 $L^*_A$ 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 $\tilde{D}$ will coincide with the space of Lipschitz functions with respect to $\tilde{D}$.

Before the statement of our main result we will present 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
\[
\sum_{n=1}^{\infty} (d_n)^{-\alpha} = d < \infty,
\] (8)
where \(d_n\) is given in (1). Observe that if \(c > 1\), then (8) is always satisfied. Moreover, since (8) implies that
\[
\sum_{n=1}^{\infty} (\beta_n^{p})^{-p} < \infty,
\]
we deduce as a consequence of Remark 1 that the weighted shift \(L\) associated to the sequence \((\alpha_n)_{n \geq 1}\) is uniformly positively expansive and topologically mixing, Devaney chaotic and frequently hypercyclic (see for instance [5], [13], and [14]). In order to guarantee the existence of invariant measures, we need the following property.

**Definition 1.** We say that \(m\) has adapted tails if, for any \(\epsilon > 0\) there exists a sequence of positive numbers \((\kappa_n)_{n \geq 1}\), such that,
1. \(\sum_{n=1}^{\infty} m([-\beta_n^{p} \kappa_n, \beta_n^{p} \kappa_n]) < \epsilon\),
2. the sequence \((\kappa_n)_{n \geq 1}\) is in \(X\).

For examples of measures with adapted tails we refer to Propositions 5 and 6. Recall that the support of a Borelian measure \(\nu\) 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 \(L : X \to X\) is a weighted shift satisfying (8) and that \(A \in \text{Lip}_b(D^\alpha_X, X)\).

a) If \(A \in SV(X)\), then there exist \(\lambda_A > 0\) and a strictly positive function \(\psi_A \in \text{Lip}_b(D^\alpha_X, X)\) such that \(L_A(\psi_A)(x) = \lambda_A \psi_A(x)\), for each \(x \in X\).

b) If \(m\) has adapted tails and \(L_A(1) = 1\), then there exists a unique \(\sigma\)-additive Borel probability measure \(\mu_A\) which is a fixed point for the operator \(L_A^*\). Moreover, \(\lim_{n \to \infty} (L_A^n)^*(\nu) = \nu\) in the weak*- topology for any Borel probability measure \(\nu\). In particular, \(\mu_A\) is an invariant and mixing probability measure (and therefore conservative and ergodic).

c) If \(A \in SV(X)\) and \(m\) has adapted tails, then there exists a unique fixed point \(\mu_A\) for the operator \(L_A^*\), where \(A = A + \log(\psi_A) - \log(\psi_A \circ L) - \log(\lambda_A)\). Moreover, \(\mu_A\) is an invariant and mixing \(\sigma\)-additive probability measure.

d) Assume that \(L_A(1) = 1\). Then, if \(\pi_1\) is the projection on the first coordinate, we get \(\text{supp}(\mu_A) \supset \pi_1^{-1}(\text{supp}(m))\). From this follows that \(\text{supp}(\mu_A) = X\).
Proof. We will adapt results that appear in [2] and [21] for the present setting.

Part a) For each $s \in (0, 1)$ define the operator $T_{s,A} : C_b(X) \to C_b(X)$ by

$$T_{s,A}(u)(x) = \log \left( \int_{L_v=x} e^{A(v)+su(v)} \, dm(v) \right). \tag{9}$$

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 $\psi_A$ for the Ruelle operator $L_A$ it will be required later to take $s \to 1$.

For fixed $s$, and using the coordinates notation obtained from the Schauder basis $\{e_k\}_{k \geq 1}$, it follows that the equation (9) can be written as

$$T_{s,A}(u)(x_1, x_2, \ldots) = \log \left( \int_{\mathbb{R}} e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots) + su(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, \ldots)} dm(r) \right). \tag{10}$$

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

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

$$|T_{s,A}(u_1)(x) - T_{s,A}(u_2)(x)| = \left| \log \left( \frac{\int_{L_v=x} e^{A(v)+su_1(v)} \, dm(v)}{\int_{L_v=x} e^{A(v)+su_2(v)} \, dm(v)} \right) \right| \leq \log \left( \frac{\int_{L_v=x} e^{A(v)+su_1(v)} \, dm(v)}{\int_{L_v=x} e^{A(v)+su_2(v)} \, dm(v)} \right) = s \|u_1 - u_2\|_\infty.$$ 

This shows that for all $0 < s < 1$, the operator $T_{s,A}$ is a uniform contraction.

Therefore, since $C_b(X)$ is a complete metric space with the distance induced by the uniform norm $\| \cdot \|_\infty$, it follows from the Banach contraction principle that $T_{s,A}$ has a unique fixed point $u_s \in C_b(X)$. That is, for all $x$ and $0 < s < 1$,

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

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

i) $\lim_{n \to \infty} s_n = 1$;

ii) $\lim_{n \to \infty} u^*_s(x) = u(x)$ for each $x \in X$;

iii) $\lim_{n \to \infty} u^*_s = u$ as uniform limit on any compact subset of $X$. 

12
Moreover, we will show that the function \( u \) such that \( e^u = \psi_A \) is the main eigenfunction for \( L_A \). That is, \( \psi_A \) satisfies \( a \), with \( \lambda_A \) maximal on the set of eigenvalues of \( L_A \).

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

\[
S_1(x, y) := \sup_{r \in \mathbb{R}} \left\{ A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, ...) - A(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, ...)
+ s\left(u_s(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, ...)
- u_s(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, ...)\right)\right\}.
\]

Using the expression in coordinates for the operator \( T_{s,A} \) (which appears in \((10)\)) and the fact that \( u_s \) is a fixed point for \( T_{s,A} \), it follows that

\[
e^{u_s(x)} = \int_{\mathbb{R}} e^{A(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, ...)} + s\left(u_s(r, \frac{x_1}{\alpha_1}, \frac{x_2}{\alpha_2}, ...)
- u_s(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, ...)\right) dm(r)
\]

\[
\leq e^{S_1(x,y)} \int_{\mathbb{R}} e^{A(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, ...)} + s\left(u_s(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, ...)
- u_s(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2}, ...)\right) dm(r)
\]

\[
= e^{S_1(x,y)} e^{u_s(y)}.
\]

That is,

\[
|u_s(x) - u_s(y)| \leq S_1(x,y).
\]

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

\[
|u_s(x) - u_s(y)| \leq S_n(x,y),
\]

is satisfied, where

\[
S_n(x, y) := \sup_{(r_1, ..., r_n)} \left\{ \left( \sum_{j=1}^{n} s^{j-1} \left( A(r_j, \frac{r_{j-1}}{\beta_1^1}, \frac{r_{j-1}}{\beta_1^2}, ..., \frac{r_{j-1}}{\beta_1^n}, \frac{x_1}{\beta_2^1}, \frac{x_2}{\beta_2^2}, ...)ight.ight.
\]

\[
\left. - A(r_j, \frac{r_{j-1}}{\beta_1^1}, \frac{r_{j-1}}{\beta_1^2}, ..., \frac{r_{j-1}}{\beta_1^n}, \frac{y_1}{\beta_2^1}, \frac{y_2}{\beta_2^2}, ...))\right)\right\}
\]

\[
+ s^n \left(u_s(r_n, \frac{r_{n-1}}{\beta_1^1}, ..., \frac{r_1}{\beta_1^n}, \frac{x_1}{\beta_2^1}, \frac{x_2}{\beta_2^2}, ...
- u_s(r_n, \frac{r_{n-1}}{\beta_1^1}, ..., \frac{r_1}{\beta_1^n}, \frac{y_1}{\beta_2^1}, \frac{y_2}{\beta_2^2}, ...))\right)\right\}.
\]
Moreover, for any pair $x, y \in X$, the sequence $(S_n(x, y))_{n \geq 1}$ satisfies

$$S_n(x, y) \leq \sup_{(r_1, \ldots, r_n)} \left\{ \left( \sum_{j=1}^{n} s_j^{-1} \text{Lip}_{A,D} \| \left( \frac{r_j}{\beta_1^j}, \ldots, \frac{r_{j-1}}{\beta_1^{j-1}}, \frac{x_1}{\beta_1^j}, \frac{x_2}{\beta_2^j}, \ldots \right) \|_X \right)^\alpha \right\} + 2s^n \|u_s\|_\infty$$

$$= \left( \sum_{j=1}^{n} s_j^{-1} \text{Lip}_{A,D} \| \left( \frac{x_1}{\beta_1^j}, \frac{x_2}{\beta_2^j}, \ldots \right) - \left( \frac{y_1}{\beta_1^j}, \frac{y_2}{\beta_2^j}, \ldots \right) \|_X \right)^\alpha + 2s^n \|u_s\|_\infty$$

$$\leq \left( \sum_{j=1}^{n} s_j^{-1} \text{Lip}_{A,D} (d_j)^{-\alpha} \|x - y\|_X \right)^\alpha + 2s^n \|u_s\|_\infty$$

$$\leq \left( \sum_{j=1}^{n} \text{Lip}_{A,D} (d_j)^{-\alpha} \|x - y\|_X \right)^\alpha + 2s^n \|u_s\|_\infty .$$

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

$$|u_s(x) - u_s(y)| \leq d \text{Lip}_{A,D} \|x - y\|_X \cdot$$

where $d$ is as in (8). 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,

$$S_n(x, y) \leq \sup_{(r_1, \ldots, r_n)} \left\{ \left( \sum_{j=1}^{n} s_j^{-1} V_j(A) + 2s^n \|u_s\|_\infty \right)^\alpha \right\}$$

$$\leq \sum_{j=1}^{n} V_j(A) + 2s^n \|u_s\|_\infty .$$

Thus, taking the limit as $n \to \infty$, it follows from (3) and (11) 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

$$|u^*_s(x)| \leq |u_s(x) - u_s(0)| \leq V(A) ,$$

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:

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

14
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, $\kappa := \lim_{n \to \infty} (1 - s_n) u_{s_n}(0)$.

Considering the sequences $u_{s_n}$ and $u_{s_n}^* = u_{s_n} - u_{s_n}(0)$, we use Arzelà-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 $n$, $n \in \mathbb{N}$.

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

$$e^{u(x)} = e^{-\kappa} \int_{L_v=x} e^{A(v) + u(v)} dm(v), \quad (14)$$

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

Indeed, since $T_{s_n} A(u_{s_n}) = u_{s_n}$, we have

$$e^{u_{s_n}^*(x)} = e^{-(1-s_n)u_{s_n}(0)} \int_{L_v=x} e^{A(v) + s_n u_{s_n}^*(v)} dm(v). \quad (15)$$

It follows that

$$\int_{L_v=x} e^{A(v) + s_n u_{s_n}^*(v)} dm(v) = \int_{\mathbb{R}} e^{A(r, \frac{e_1}{e_2}, \frac{e_2}{e_3}, \ldots) + s_n u_{s_n}^*(r, \frac{e_1}{e_2}, \frac{e_2}{e_3}, \ldots)} dm(r),$$

and by (13) we have that

$$\int_{\mathbb{R}} \left| e^{A(r, \frac{e_1}{e_2}, \frac{e_2}{e_3}, \ldots) + s_n u_{s_n}^*(r, \frac{e_1}{e_2}, \frac{e_2}{e_3}, \ldots)} \right| dm(r) \leq C \int_{\mathbb{R}} dm(r) < \infty,$$

where $C = e^{\|A\|_{\infty} + V(A)}$. Therefore, (14) is a consequence of (15) and the Dominated Convergence Theorem.

The foregoing implies that the positive function $\psi_A(x) = e^{u(x)}$, satisfies the equation $L_A(\psi_A)(x) = \lambda_A \psi_A(x)$ for all $x \in X$, with $\lambda_A = e^{-\kappa}$. Furthermore, by (12), it follows that for each $x, y \in X$ is satisfied

$$|u(x) - u(y)| = \lim_{n \to \infty} |u_{s_n}^*(x) - u_{s_n}^*(y)| \leq d \text{Lip}_A (D_X^\alpha) \|x - y\|_X^{\alpha}, \quad (16)$$

and by (13), we obtain that for any $x \in X$

$$|u(x)| = \lim_{n \to \infty} |u_{s_n}^*(x)| \leq V(A), \quad (17)$$

By (16) and (17), it follows that $u \in \text{Lip}_b (D_X^\alpha, X)$, and thus $\psi_A = e^{u} \in \text{Lip}_b (D_X^\alpha, X)$, which concludes the proof of a).

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

$$\int_X (\varphi \circ L) d\mu_A = \int_X (\varphi \circ L) d(L_A^\alpha (\mu_A)) = \int_X L_A (\varphi \circ L) d\mu_A = \int_X \varphi d\mu_A.$$

15
Hence, $\mu_A$ is $L$-invariant and, as a consequence of Poincaré’s Recurrence Theorem, conservative. We now showing that $\mu_A$ is mixing. As $(L^n_A)^*(\nu) \to \mu_A$ for any $\nu$ in the weak-* topology, it follows that, for any $x \in X$ and any bounded continuous function $f : X \to \mathbb{R}$, that

$$L^n_A(f)(x) = \int_X f d(L^n_A)^*(\delta_x) \to \int_X f d\mu_A.$$ 

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

$$\lim_{n \to \infty} \int_X f g \circ L^n d\mu_A = \lim_{n \to \infty} \int_X f (g \circ L^n) d(L^n_A)^*(\mu_A) = \int_X L^n_A(f) g d\mu_A = \int_X f d\mu_A \int_X g d\mu_A.$$ 

This implies that $\mu_A$ is mixing and part b) is proven.

Part c) This an immediate consequence of parts a) and b).

Part d) In order to show the remaining part, for each $x \in X$ and $\epsilon > 0$, we define

$$F(x; \epsilon) = \left\{ y \in X, |y_n - x_n| \leq \frac{\epsilon}{2^n} \forall n \in \mathbb{N} \right\}.$$ 

Note that $F(x; \epsilon)$ is a closed set as

$$F(x; \epsilon) = \bigcap_{n=1}^{\infty} \pi_n^{-1}([x_n - \epsilon/2^n, x_n + \epsilon/2^n]),$$

where $\pi_i : X \to \mathbb{R}$ is the bounded linear operator defined as $\pi_i(y) = y_i$, for each $y \in X$ and any $i \in \mathbb{N}$. Moreover, by definition of the norms $|| \cdot ||_X$, it follows that

$$F(x; \epsilon) \subset \frac{\epsilon}{(2^p - 1)^{\frac{1}{p}}} B_X + \{x\},$$

when $X = l^p(\mathbb{R})$, $1 \leq p < \infty$ and

$$F(x; \epsilon) \subset \frac{\epsilon}{2} B_X + \{x\},$$

when $X = c_0(\mathbb{R})$, with $B_X$ referring to the unit ball in $X$ with respect to the norm $|| \cdot ||_X$.

We now claim that the set $F(x; \epsilon)$ has positive measure with respect to $\mu_A$. Indeed, since $X$ is a separable Banach space, it follows that $\mu_A \in B(\mathcal{L}(X))$ is a regular measure. Hence there exist a closed set $F \subset X$ and an open set $U \subset X$ satisfying

1. $F(x; \epsilon/2) \subset F \subset F(x; \epsilon) \subset U$,
2. $\mu_A(U \setminus F) < \mathcal{E}$, with $\mathcal{E} = \frac{1}{2} e^{\inf(A)} m([x_1 - \epsilon/4, x_1 + \epsilon/4]).$
By Urysohn’s Lemma, there exists \( \varphi \in C_b(X) \), such that, \( \chi_F \leq \varphi \leq \chi_U \). This implies
\[
\int_X |\chi_F(x; \epsilon)(y) - \varphi(y)| d\mu_A(y) < \mathcal{E}.
\]
For \( U(x; \epsilon) = \{ y \in X, |y_1 - x_1| \leq \epsilon/2 \} \), and \( x_n = y_n : \forall n > 1 \), it follows that \( U(x; \epsilon/2) \subset F(x; \epsilon) \), implying that
\[
\int_X |\chi_F(x; \epsilon/2)(y; \varphi(y)) - \chi_U(x; \epsilon/4)(r, \frac{y_1}{\alpha_1}, \frac{y_2}{\alpha_2},...) dm(r) - d\mu_A(y) - \mathcal{E} = 0.
\]
This implies that \( \mu_A(V) > 0 \), for any open neighborhood \( V \subset X \) of \( x \). Hence, the support of \( \mu_A \) contains the inverse image of the support of the a priori measure \( m \) by the map \( \pi_1 \), which implies that the support of \( \mu_A \) is the set \( X \). 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 \( A \in \text{Lip}_b(D^p_X, X) \) and \( L : l^p(\mathbb{R}) \rightarrow l^p(\mathbb{R}) \) is a weighted shift satisfying (8). 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 (8). Using Cauchy’s criterion, we deduce that this class contains all the weighted shifts satisfying
\[
\lim_{n \to \infty} (d_n)^{-\frac{1}{p}} = \lim_{n \to \infty} \left( \frac{1}{\beta_k^n} \right)^{-\frac{1}{p}} < 1.
\]
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. $\lim_{n \to \infty} (d_n)^{-\frac{1}{p}} < 1$.

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 $S_k = \sum_{n=1}^{\infty} (\beta_k^n)$. 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

$$(\beta_k^n)^{-1} = \frac{S_k}{1 + S_{k+1}} \cdot \frac{S_{k+1}}{1 + S_{k+1}} \cdot \ldots \frac{S_{k+n-1}}{1 + S_{k+n-1}} \leq M \cdot \left( \frac{M}{1+M} \right)^{n-1}.$$ 

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

On the other hand, we have

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

Moreover, by (2) we have $\lim_{n \to \infty} \left( \sup_{k \geq 1} (1/\beta_k^n) \right)^{\frac{1}{p}} = \lim_{n \to \infty} \| L' \|^x_{\text{op}}$, 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 $\lim_{n \to \infty} (d_n)^{-\frac{1}{p}} = \limsup_{n \to \infty} (d_n)^{-\frac{1}{p}} < 1$.

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

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 $L : X \to X$ is a weighted shift satisfying

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

Then, Theorem 1 holds.

18
Proof. By Lemma 2, \( \lim_{n \to \infty} (d_n)^{-\frac{1}{n}} < 1 \) is equivalent to (8).

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 \( A \in \text{Lip}(D_X^\alpha, X) \) and \( L : X \to X \) is a weighted shift satisfying \( \lim_{n \to \infty} \left( \inf_{k \geq 1} |\beta_k^n|^{-\frac{1}{n}} \right) < 1 \).

Example 1. There are many cases which can be considered on the class of weighted shifts operators satisfying (18). For instance:

1. Take \( 1 < c < \alpha_n \leq c', \) 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,

\[
\beta_k^n = \alpha_k...\alpha_{k+n-1} \geq (a^e b)^{\left[ \frac{n}{e+1} \right]} a^e, \text{ for all } k \geq 1,
\]

where \( \left[ \frac{n}{e+1} \right] \) 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 \( T : X \to X \) 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 \( \lim_{n \to \infty} (m(T^n))^{\frac{1}{n}} \) exists (see [8]).

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

$$X = \text{Ker}(T) \oplus E.$$  \hfill (19)

We define

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

Note that

$$p(T)\|x\|_X \leq \|T(x)\|_X, \text{ for all } x \in E. \hfill (20)$$

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

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

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

$$p(L) = \inf_{k \geq 1} \alpha_k.$$  \hfill (21)

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,

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

Hence $p(L) \geq \inf_{k \geq 1} \alpha_k$ and we obtain (21).

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

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

Observe that $\lim_{n \to \infty} (p(L^n))^\frac{1}{n}$ 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 \( T : X \to X \) is onto, not one to one and satisfying

\[
X = \text{Ker}(T^n) \oplus E_n \quad \text{and} \quad E_{n+1} \subset E_n, \quad \text{for all } n \geq 1, \tag{22}
\]

where \( E_n \) are closed subspaces of \( X \).

Note that the assumption (22) is satisfied by the weighted shift \( L \) in any space \( X \in \{c_0(\mathbb{R}), X = l^p(\mathbb{R}), 1 \leq p < \infty \} \). Furthermore, (22) is also satisfied by any continuous operator \( T : X \to X \) 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 \( \lim_{n \to \infty} (p(T^n))^{-\frac{2}{n}} \) exists. Indeed, we have

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

Hence, the sequence \( (-\log(p(T^n)))_{n \geq 1} \) is sub-additive. Thus, \( \lim_{n \to \infty} -\frac{\log(p(T^n))}{n} \) 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 \textit{a priori} probability \( m \) on \( \text{Ker}(T) \). We will also denote by \( m \) the corresponding \textit{a priori} measure on \( T^{-1}(x) = \{x'\} + \text{Ker}(T) \). This then gives rise to a Ruelle operator associated to the potential \( A \in \text{Lip}_b(D^\alpha_X, X) \) and the \textit{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 done before and sometimes we just outline some of the proofs.

**Lemma 3.** Let \( X \) be a Banach space and \( T : X \to X \) a bounded linear operator, such that, \( T \) is onto, not one to one, satisfying (19), \( p(T) > 0 \) and \( A \in \text{Lip}_b(D^2_X, X) \). Then, the Ruelle operator \( \mathcal{L}_A \) preserves the spaces \( \text{Lip}_b(D^2_X, X) \) and \( \text{Lip}_b^{\text{loc}(\delta)}(D^2_X, X) \), \( \delta > 0 \).

**Remark 7.** It’s classical and not difficult to prove that \( p(T) > 0 \) is equivalent to the fact that \( T : E \to X \) is injective and \( T(E) \) is a closed subspace of \( X \).
Proof. For any $\varphi \in \text{Lip}_b(D^\alpha_X, X)$ and each $x, y \in X$, we have

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

where

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

Hence,

$$\left| \mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y) \right| \leq \int_{\text{Ker}(T)} e^{\|A\|_\infty \|\varphi\|_{\infty} |x' + z - y' + z|} \left| e^{A(x' + z)} - e^{A(y' + z)} \right| dm(z),$$

Using (20) we deduce that

$$\left| \mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y) \right| \leq e^{\|A\|_\infty \|\varphi\|_{\infty} \|\text{Lip}_e^\alpha, D^\alpha_X\|} \int_{\text{Ker}(T)} \|x' - y'\|^\alpha_X dm(z).$$

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

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

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

where $V_{T,n}(A) = \sup \{|A(x) - A(y)| : x, y \in X, T^n(x) = T^n(y), 1 \leq i \leq n\}$. We will denote by $\mathcal{SV}_T(X)$ the set of potentials $A \in \mathcal{C}(X)$ satisfying (23).

**Proposition 2.** Let $X$ be a separable Banach space and $T : X \to X$ a bounded linear operator, such that, $T$ is onto, not bijective, satisfying (22), $p(T^n) > 0$ and $\dim(\text{Ker}(T^n)) < \infty$, for each $n \in \mathbb{N}$. Consider $A \in \mathcal{SV}_T(X) \cap \text{Lip}_b(D^\alpha_X, X)$ and assume that

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

Then, Theorem 1 holds.

Proof. The proof is the same as the proof of Theorem 1. For each $s \in (0, 1)$, we consider the operator $T_{s,A} : \mathcal{C}_b(X) \to \mathcal{C}_b(X)$ given by

$$T_{s,A}(u)(x) = \log \left( \int_{z \in T^{-1}(x)} e^{A(x') + su(v)} dm(v) \right).$$
In the same way as before, for all $s \in (0, 1)$, the map $T_{s,A}$ has a unique fixed point $u_s \in C_b(X)$. That is, for all $x$ and $0 < s < 1$

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

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

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

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

$$\mathcal{S}_n(x, y) = \sup_{z_j \in \ker(T_j)} \left\{ \left( \sum_{j=1}^{n} s^{j-1} (A(x_j + z_j) - A(y_j + z_j)) \right) + s^n (u_s(x_n + z_n) - u_s(y_n + z_n)) \right\},$$

where $x_j$ and $y_j$, $1 \leq j \leq n$, are fixed elements of $X$ satisfying $T_j(x_j) = x$, $T_j(y_j) = y$ for all $1 \leq j \leq n$.

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

$$\mathcal{S}_n(x, y) \leq \sup_{z_j \in \ker(T')} \left\{ \left( \sum_{j=1}^{n} s^{j-1} \text{Lip}_{A,D,X} \| x_j - y_j \|_X^\alpha + 2s^n \| u_s \|_\infty \right) \right\}$$

$$\leq \left( \sum_{j=1}^{n} s^{j-1} \text{Lip}_{A,D,X} \frac{1}{p(T')^\alpha} \| x - y \|_X^\alpha + 2s^n \| u_s \|_\infty \right)$$

$$\leq \left( \sum_{j=1}^{n} \text{Lip}_{A,D,X} \frac{1}{p(T')^\alpha} \| x - y \|_X^\alpha + 2s^n \| u_s \|_\infty \right).$$

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

$$|u_s(x) - u_s(y)| \leq \left( \text{Lip}_{A,D,X} \sum_{n=1}^{\infty} \frac{1}{p(T')^\alpha} \right) \| x - y \|_X^\alpha.$$

Besides that, by (23), it follows that

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

This is the end of the proof.
It follows from the above that:

**Corollary 2.** Consider $X$ a separable Banach space and $T : X \to X$ a bounded linear operator such that $T$ is onto, not injective, satisfying (22), $p(T^n) > 0$ and $\dim(\ker(T^n)) < \infty$, for each $n \in \mathbb{N}$. Consider $A \in SV_T(X) \cap \text{Lip}_b(D_X^\alpha, X)$. Assume that the following is true:

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

Then, Theorem 1 holds.

In the next result, we prove that a sub class of the class of operators given by Corollary 2 is formed by uniformly hypercyclic operators.

**Proposition 3.** Let $X$ be a separable Banach space and $T : X \to X$ be a bounded linear operator, such that, $T$ is onto, not one to one, satisfying (22) and $\bigcup_{n=0}^{\infty} \ker(T^n)$ is dense in $X$. Assume that

$$\lim_{n \to \infty} \frac{p(T^n)^{-\frac{1}{2}}}{n} < 1,$$

then, $T$ is uniformly hypercyclic and Devaney chaotic.

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

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

Hence, $\bigcup_{n=0}^{\infty} \ker(L^n)$ 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 $T : X \to X$ be a continuous linear operator. Assume that there exist a dense set $D \subset X$ and a map $S : D \to D$, 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 3. Set \( D = \bigcup_{n=0}^{\infty} \ker(T^n) \), then for all \( x \in D \), there exists a non-negative integer \( n_0 = n_0(x) \), such that,

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

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

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

Hence, by (20), it follows that

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

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

Throughout this appendix, we assume that \( X = c_0(\mathbb{R}) \) or \( X = l^p(\mathbb{R}) \), \( 1 \leq p < \infty \), and that \( A \in \text{Lip}_b(D_\alpha X, X) \) 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 [17], which was adapted to the dual of the Ruelle operator by several authors (see [27],[18],[20],[7],[10],[23],[28], in chronological order). 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

\[
\bar{D}(x,y) := \min \{1, aD_X(x,y)\}.
\]

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 \( \bar{D} \) in order to obtain a contraction for Dirac measures whose base points have \( \bar{D} \)-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 \( D := D_\xi \). Moreover, we refer to \( \text{Lip}^\text{loc}(\delta)(D, X) \) as the set of continuous functions \( \varphi : X \to \mathbb{R} \) such that

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

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

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

(24)

and \( \sup(\varphi) - \inf(\varphi) \leq \text{Lip}^\text{loc}(1)(\varphi, \tilde{D}) \).

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

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

\[
\frac{\left| \varphi(x) - \varphi(y) \right|}{D(x, y)} = \begin{cases} 
\frac{\left| \varphi(x) - \varphi(y) \right|}{\frac{1}{a} \text{Lip}^\text{loc}(1/a)(\varphi, D)} & : D(x, y) < 1/a \\
\left| \varphi(x) - \varphi(y) \right| / 2\|\varphi\|_\infty & : D(x, y) \geq 1/a
\end{cases}
\]

This proves that \( \varphi \in \text{Lip}(\tilde{D}, X) \) and the left half of (24). The left half of (24) follows from the trivial estimate \( \text{Lip}^\text{loc}(1)(\varphi, \tilde{D}) \leq \text{Lip}^\text{loc}(1)(\varphi, D) \) and the first observation.

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

\[
\sup(\varphi) - \inf(\varphi) \leq \text{Lip}^\text{loc}(1)(\varphi, \tilde{D}) \sup_{\tilde{D}(x, y) \leq 1} \tilde{D}(x, y) = \text{Lip}^\text{loc}(1)(\varphi, \tilde{D}) \leq \text{Lip}^\text{loc}(1)(\varphi, D)
\]

\( \Box \)

In order to introduce the Wasserstein distance (see for details [9, 29]), recall that a coupling of \( \mu, \nu \in \mathcal{P}(X) \) is a Borel probability measure \( \Pi \) on \( X \times X \) such that the first marginal of \( \Pi \) is \( \mu \) and the second marginal is \( \nu \). We refer to \( \Gamma(\mu, \nu) \) as the set of all couplings of \( \mu \) and \( \nu \). 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 \( \sigma \)-algebra. Moreover, if \( X \) has finite diameter with respect to \( D \), then \( W_D \) is compatible with the weak convergence of measures, that is, \( \mu_n \) converges weakly to \( \mu \) if and only of \( W_D(\mu_n, \mu) \to 0 \). A further fundamental tool in this setting is Kantorovich’s duality, which states that

\[
W_D(\mu, \nu) = \sup \left\{ \left| \int_X \varphi d\mu - \int_X \varphi d\nu \right| : \text{Lip}^\text{loc}(\varphi, D) \leq 1 \right\}.
\]
We are now in position to prove the basic contraction estimates for our setting. In order to do so, observe that \( L_A^*(1) = 1 \) implies that

\[
L_A^*(\mu)(X) = \int_X 1dL_A^*(\mu) = \int_X L_A(1)d\mu = 1 \text{ for any } \mu \in \mathcal{P}(X).
\]

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

\[
C_A := \sup \left\{ \frac{e^{Lip_\alpha D} \sum_{i=1}^\infty d_i^{-\alpha} t - 1}{t} : 0 < t \leq 1 \right\}
\]

**Lemma 5** (Local contraction for Dirac measures). Assume that \( n \) is such that \( (d_n)^{-\alpha} \leq \frac{3}{4} \) and that \( a = \max\{8C_A/3, 1\} \). Then, for all \( x, y \in X \) with \( \tilde{D}(x, y) < 1 \),

\[
W_{\tilde{D}}((L_A^n)^\circ(\delta_x), (L_A^n)^\circ(\delta_y)) \leq \frac{3}{4} W_{\tilde{D}}(\delta_x, \delta_y).
\]

**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 \( \varphi \in Lip(\tilde{D}, X) \) with \( \inf(\varphi) = 0 \). Then, by Lemma 4,

\[
|L_A(\varphi)(x) - L_A(\varphi)(y)| \\
\leq \left( e^{S_n(A)\phi_{\tau_a}(x)} |\varphi \circ \tau_a(x) - \varphi \circ \tau_a(y)| dm_a(a) \\
+ \int |e^{S_n(A)\phi_{\tau_a}(x)} - e^{S_n(A)\phi_{\tau_a}(y)}| |\varphi \circ \tau_a(y)| dm_a(a) \right) \\
\leq Lip_{\varphi, \tilde{D}}(\tau_a(x), \tau_a(y)) + \sup_a \left( e^{S_n(A)\phi_{\tau_a}(x)} - e^{S_n(A)\phi_{\tau_a}(y)} - 1 \right) ||\varphi||_\infty \\
\leq d_n^{-\alpha} Lip_{\varphi, \tilde{D}}(\tilde{D}(x, y) + \left( e^{Lip_\alpha D} \sum_{i=1}^\infty d_i^{-\alpha} D(x,y) - 1 \right) Lip_{\varphi, \tilde{D}} \\
\leq Lip_{\varphi, \tilde{D}} \left( d_n^{-\alpha} \tilde{D}(x, y) + C_A D(x,y) \right) = Lip_{\varphi, \tilde{D}} \left( d_n^{-\alpha} + \frac{C_A}{a} \right) \tilde{D}(x, y).
\]

Observe that \( L_A(1) = 1 \) implies that

\[
L_A(\varphi + c)(x) - L_A(\varphi + c)(y) = L_A(\varphi)(x) - L_A(\varphi)(y).
\]

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

\[
W_{\tilde{D}}((L_A^n)^\circ(\delta_x), (L_A^n)^\circ(\delta_y)) \\
= \sup \left\{ |L_A(\varphi)(x) - L_A(\varphi)(y)| : Lip_{\varphi, \tilde{D}} \leq 1 \right\} \\
\leq \left( d_n^{-\alpha} + \frac{C_A}{a} \right) \tilde{D}(x, y) = \left( d_n^{-\alpha} + \frac{C_A}{a} \right) W_{\tilde{D}}(x, y).
\]

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

We now analyze the action of \( L_A^* \) on \( \delta_x \) and \( \delta_y \) 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 $d^{-\alpha} D(x, y) < n$, then

$$W_{\tilde{D}}((\mathcal{L}_h^n)^n(\delta_x), (\mathcal{L}_h^n)^n(\delta_y)) \leq 1 - e^{-\sum_{i=1}^{\infty} d^{-\alpha} D(x, y)} (1 - ad^{-\alpha} D(x, y)).$$

Proof. Define

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

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

$$R_n(\pi_i^{-1}(A)) = \int 1_A \circ \tau_n(\pi_i(x, y)) e^{\min\{S_n(A) \circ \tau_n(x), S_n(A) \circ \tau_n(y)\}} dm_n \leq L_n(1_A(\pi_i(x, y)) = (L_h^n)^n(\delta_{\pi_i(x, y)})(A).$$

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

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

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

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

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

$$Q_n(\Delta_{\epsilon}) \geq R_n(\Delta_{\epsilon}) \geq \int e^{\min\{S_n(A) \circ \tau_n(x), S_n(A) \circ \tau_n(y)\}} dm_n \geq \int e^{S_n(A) \circ \tau_n(x)} dm_n \epsilon^{-C D(x, y)}.$$

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

$$W_{\tilde{D}}((\mathcal{L}_h^n)^n(\delta_x), (\mathcal{L}_h^n)^n(\delta_y)) \leq \int \tilde{D}(x, y) dQ(x, y) = \int_{\Delta_{\epsilon}} \tilde{D}(x, y) dQ(x, y) + \int_{\Delta_{\epsilon}} \tilde{D}(x, y) dQ(x, y) \leq 1 - Q(\Delta_{\epsilon}) + \alpha Q(\Delta_{\epsilon}) = 1 - Q(\Delta_{\epsilon})(1 - \alpha \epsilon) \leq 1 - \epsilon^{-C D(x, y)} (1 - ad^{-\alpha} D(x, y)).$$

The assertion follows by including the case $\epsilon \geq 1/a$. \qed

28
In order to employ Lemma 5 and 6 in order to obtain estimates for arbitrary measures, we recall the following basic fact. Assume that \( n \in \mathbb{N} \) and that \( Q_{\xi, \eta} \in \Gamma((L^*_{\lambda})^n(\delta_\xi), (L^*_{\lambda})^n(\delta_\eta)) \) such that \((x, y) \mapsto \int \tilde{D}(x, y) dQ_{\xi, \eta}(x, y)\) is measurable. Then, for any \( \mu, \nu \in P(X) \) and any \( \Pi \in \Gamma(\mu, \nu) \), we get that
\[
\int dQ_{\xi, \eta}(x, y) d\Pi(\xi, \eta) \in \Gamma((L^*_{\lambda})^n(\mu), (L^*_{\lambda})^n(\nu)).
\]
And, in particular, by taking the infimum over all possible couplings \( Q_{\xi, \eta} \),
\[
W_\tilde{D}( (L^*_{\lambda})^n(\mu), (L^*_{\lambda})^n(\nu)) \leq \int W_\tilde{D}( (L^*_{\lambda})^n(\delta_x), (L^*_{\lambda})^n(\delta_y)) d\Pi(x, y). \tag{25}
\]
This allows us to summarize Lemma 5 and 6 into a single estimate. In order to do so, set
\[
r_n(t) := \begin{cases} 
1 - e^{-\sum_{i=1}^n d_i^{-\alpha} t} (1 - \min\{1, t a_n^{-\alpha}\}) & : t \geq \frac{1}{a} \\
\frac{3}{4} & : t < \frac{1}{a}
\end{cases}
\]

**Lemma 7.** Assume that \( \mu, \nu \in P(X) \) and that \( \Pi \in \Gamma(\mu, \nu) \). Then, for any \( n \) such that \( (d_n)^\alpha \geq 8/3 \),
\[
W_\tilde{D}( (L^*_{\lambda})^n(\mu), (L^*_{\lambda})^n(\nu)) \leq \int r_n(D(x, y)) \tilde{D}(x, y) d\Pi(x, y).
\]

**Proof.** By dividing the domain of integration in (25) 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 \( P(X) \) which we define now. For \( \gamma > 0 \) and \( \epsilon > 0 \), set
\[
P_{\gamma, \epsilon}(X) := \{ \mu \in P(X) : \mu(\{ x \in X : D(x, 0) > \gamma \}) < \epsilon \}.
\]

**Proposition 4.** Assume that \( \mu, \nu \in P(X) \) and that \( \gamma \geq 1/a \) is chosen such that \( \mu, \nu \in P_{\gamma, \epsilon}(X) \), for \( \epsilon := W_\tilde{D}(\mu, \nu)/4 \). Then, for any \( n \) with \((d_n)^\alpha \geq 8/3\),
\[
W_\tilde{D}( (L^*_{\lambda})^n(\mu), (L^*_{\lambda})^n(\nu)) \leq \left( 1 - \frac{1}{2} \min\{1, 2a\gamma d_n^{-\alpha}\} \right) W_\tilde{D}(\mu, \nu).
\]

**Proof.** By the triangle property, \( D(x, y) \leq 2\gamma \) whenever \( D(x, 0), D(y, 0) \leq \gamma \). Hence,
\[
\{ D(x, y) > 2\gamma \} \subset \{ D(x, 0) > 2\gamma \} \times X \cup X \times \{ y : D(0, y) > \gamma \}.
\]
Therefore, for any \( \Pi \in \Gamma(\mu, \nu) \), \( \Pi(\{ D(x, y) > 2\gamma \}) < 2\epsilon \). Now assume that \( \Pi \) 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,\[
W_D(L^*_A)^n(\mu), (L^*_A)^n(\nu))
\leq \int \left(1_{(D(x,y) \leq 2\gamma)}r_n(2\gamma) + 1_{(D(x,y) > 2\gamma)}\right) D(x,y) d\Pi
\]
= \left(\int \left(\frac{2\epsilon (1 - r_n(2\gamma))}{W_D(\mu, \nu)}\right) W_D(\mu, \nu)
\leq \left(1 - \frac{1 - r_n(2\gamma)}{2}\right) W_D(\mu, \nu).
\]
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 \((L^*_A)^n(\mu)\) and \((L^*_A)^n(\nu)\) is strictly contracted, even though the contraction rate depends on the distance as well as the tails of \(\mu\) and \(\nu\). On the other hand, as shown below in the proof of Theorem 2, the notion of adapted tails in Definition 1, allows to control the tails of \((L^*_A)^n(\mu)\) asymptotically. Recall that we say that \(m\) has adapted tails if, for any \(\epsilon > 0\) there exists a sequence of positive numbers \((\kappa_n)\) such that

1. \(\sum_{n=1}^{\infty} m([-\beta^n \kappa_n, \beta^n \kappa_n]) < \epsilon\),
2. the sequence \((\kappa_n)\) is in \(X\).

Before giving examples for this condition, we prove the main result of this appendix.

**Theorem 2.** Assume that \(X = c_0(\mathbb{R})\) or \(X = l^p(\mathbb{R})\), \(1 \leq p < \infty\) and that \(L : X \to X\) a weighted shift satisfying (8). Moreover assume that \(A \in \text{Lip}_b(D^*_A, X)\) with \(L_A(1) = 1\) and that \(m\) has adapted tails. Then, for any \(\nu \in \mathcal{P}(X)\), the limit

\[
\mu := \lim_{n \to \infty} (L^*_A)^n(\nu)
\]
exists in \(\mathcal{P}(X)\). Moreover, \(\mu\) is the unique element in \(\mathcal{P}(X)\) such that \(L^*_A(\mu) = \mu\).

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

\[
\Omega := \{(x_j) : x_j \in [-\kappa_j, \kappa_j] \text{ for } 0 \leq j < n, \|x_n, x_{n+1}, \ldots\|_X \leq K\}.
\]
Then, for \( \nu \in \mathcal{P}(X) \) and \( \beta^n_0 := 1 \),

\[
(\mathcal{L}^n_\ast)^n(\nu)(\Omega^c) = (\mathcal{L}^n_\ast)^n(\nu) \left( \bigcup_{j=0}^{n-1} \{ (x_i) : |x_j| > \kappa_j \} \cup \{(x_i) : \| (x_n, x_{n+1}, \ldots) \|x > K \} \right)
\]

\[
\leq \sum_{j=0}^{n-1} (\mathcal{L}^n_\ast)^n(\nu) ([|x_j| > \kappa_j]) + (\mathcal{L}^n_\ast)^n(\nu) ([\| (x_n, x_{n+1}, \ldots) \|x > K])
\]

\[
= \int \mathcal{L}^n_\ast \left( \sum_{j=0}^{n-1} 1_{\{|x_0/\beta^n_1 > \kappa_1\}} \circ L^j + 1_{\{|(x_n/\beta^n_1, x_{n+1}/\beta^n_2, \ldots )\|x > K\}} \circ L^n \right) d\nu
\]

\[
= \int \mathcal{L}^{n-j}_\ast \left( 1_{\{|x_0/\beta^n_1 > \kappa_1\}} \right) d\nu + \int 1_{\{|(x_n/\beta^n_1, x_{n+1}/\beta^n_2, \ldots )\|x > K\}} d\nu
\]

\[
\leq \sum_{j=0}^{n-1} \int \mathcal{L}^{n-j}_\ast \left( 1_{\{|x_0/\beta^n_1 > \kappa_1\}} \right) d\nu + \nu \left( \{(x_i) : \|x_i\| > d_n K\} \right).
\]

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 \( (\kappa_j) \) is chosen for \( \epsilon > 0 \) according to Definition 1, then, as \( A \) is normalized,

\[
(\mathcal{L}_\ast)^n(\nu)(\Omega^c) \leq C \sum_{k=0}^{n-1} m\{ (x_i) : |x| > \beta^n_1 \} + \nu \left( \{(x_i) : \|x_i\| > d_n K\} \right)
\]

\[
\leq C \epsilon + \nu \left( \{(x_i) : \|x_i\| > d_n K\} \right) \xrightarrow{n \to \infty} C \epsilon.
\] (26)

In the second step, we show that \( W_D((\mathcal{L}_\ast)^n(\mu), (\mathcal{L}_\ast)^n(\nu)) \to 0 \) as \( n \to \infty \). So assume that \( \mu, \nu \in \mathcal{P}(X) \). By Lemma 7, \( W_D((\mathcal{L}_\ast)^n(\mu), (\mathcal{L}_\ast)^n(\nu)) \) is decreasing in \( n \) and, in particular,

\[
L := \lim_{n \to \infty} W_D((\mathcal{L}_\ast)^n(\mu), (\mathcal{L}_\ast)^n(\nu))
\]

exists. Now assume that \( L > 0 \). Then, as \( m \) is adapted, there exists \( (\kappa_j) \in X \) with \( \sum_{j=0}^{n-1} m\{[-\beta^n_1 \kappa_j, \beta^n_1 \kappa_j]^c \} < L/(8C) \). For \( K := \|x\|_X \), the estimate (26) combined with \( d_n \to \infty \) implies that there exists \( n_0 \in \mathbb{N} \) such that \( (\mathcal{L}_\ast)^n(\nu)(\Omega^c) < L/4 \) for any \( n \geq n_0 \). However, as \( \Omega \) for these choices of \( n \) and \( K \) satisfies \( \Omega \subset \{ x \in X : \|x\|_X \leq 2\| (\kappa_j) \|_X \} \), it follows that

\[
(\mathcal{L}_\ast)^n(\mu), (\mathcal{L}_\ast)^n(\nu) \in \mathcal{P}_2([\| (\kappa_j) \|_X, L/4](X)) \text{ for all } n \geq n_0.
\]

We now apply Proposition 4 to the pair \( ((\mathcal{L}_\ast)^n(\mu), (\mathcal{L}_\ast)^n(\nu)) \). Namely, for \( k \) chosen such that the factor on the right hand side in Proposition 4 is strictly smaller than 1, it follows that

\[
L = \lim_{n \to \infty} W_D((\mathcal{L}_\ast)^{n+k}(\mu), (\mathcal{L}_\ast)^{n+k}(\nu)) < \lim_{n \to \infty} W_D((\mathcal{L}_\ast)^n(\mu), (\mathcal{L}_\ast)^n(\nu)) = L,
\]

which is absurd. Hence, \( L = 0 \) and the assertion of the second step is proven.
In the third step, we use (26) to obtain that \((L^*_A)^n(\delta_0)\) 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 \((\kappa_j)\) is chosen according to Definition 1, and set \(\Omega_\epsilon := \{(x_i) \in X : |x_i| \leq \kappa_i \forall i \geq 0\}\). As 0 \(\in \Omega_\epsilon\) the estimate (26) implies that \((L^*_A)^n(\delta_0)(\Omega_\epsilon) \leq C\epsilon\).

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

In the last step, we prove for any \(\nu \in \mathcal{P}(X)\) that \(\mu = \lim_{n \to \infty} (L^*_A)^n(\nu)\) and \(\mu = L^*_A(\mu)\). In order to do so, we apply Step 2 to the measures \(\delta_0\) and \(L^*_A(\delta_0)\), which implies by continuity of \(L^*_A\) that \(\mu = \lim_{k \to \infty} (L^*_A)^n(\delta_0) = \lim_{k \to \infty} (L^*_A)^{n+1}(\delta_0) = L^*_A(\mu)\).

Finally, the remaining assertion follows from a further application of Step 2 to \(\nu\) and \(\mu\).

We now provide some examples of adapted measures. Therefore recall that a measure on \(\mathbb{R}\) has polynomial tails of order \(\gamma\) 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
\[
\sum_{n=1}^\infty m(\{x : |x| > \kappa_n^\alpha\}) \leq \sum_{n=1}^\infty m(\{x : |x| > \kappa_n d_n\}) \leq C \sum_{n=1}^\infty L^{-\gamma n^{-\ell\gamma}} < \frac{C\ell\gamma}{\ell\gamma - 1} \xrightarrow{L \to \infty} 0.
\]
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 5.** Assume that \(m\) has polynomial tails of order \(\gamma\). Then the following holds.

1. If \(n/d_n \to 0\) 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 6.** Assume that $m$ has exponential tails and that $((\log n)/d_n) \in X$. Then $m$ is adapted.

**References**


