THE SPACE OF INVARIANT MEASURES FOR COUNTABLE MARKOV SHIFTS

By

GODOFREDO IOMMI* AND ANIBAL VELOZO

Abstract. It is well known that the space of invariant probability measures for transitive sub-shifts of finite type is a Poulsen simplex. In this article we prove that in the non-compact setting, for a large family of transitive countable Markov shifts, the space of invariant sub-probability measures is a Poulsen simplex and that its extreme points are the ergodic invariant probability measures together with the zero measure. In particular, we obtain that the space of invariant probability measures is a Poulsen simplex minus a vertex and the corresponding convex combinations. Our results apply to finite entropy non-locally compact transitive countable Markov shifts and to every locally compact transitive countable Markov shift. In order to prove these results we introduce a topology on the space of measures that generalizes the vague topology to a class of non-locally compact spaces, the topology of convergence on cylinders. We also prove analogous results for suspension flows defined over countable Markov shifts.

1 Introduction

Ever since the work of Parthasarathy [Par1] and Oxtoby [O] in the early 1960s a great deal of attention has been paid to the problem of describing the space of invariant probability measures of a dynamical system. Remarkable results have been obtained relating the geometry of the space with the dynamical properties of the system. A result by Downarowicz [D] states that for every Choquet simplex K there exists a minimal sub-shift (X, T) for which the space of invariant probability measures $\mathcal{M}(X, T)$ is affinely homemorphic to K. In this article we will be interested in a very special Choquet Simplex.

Definition 1.1. A metrizable convex compact Choquet simplex with at least two points K is a **Poulsen Simplex** if its extreme points are dense in K.

The first example of such a simplex was constructed by Poulsen [Pou] in 1961. It was later shown by Lindenstrauss, Olsen and Sternfeld [LOS, Theorem 2.3]

that the Poulsen simplex is unique up to affine homemorphism. This simplex enjoys remarkable properties. For example, as proved in [LOS, Section 3], the set of extreme points in the Poulsen simplex is path connected. The relation of this simplex with dynamical systems directly follows from the seminal work of Sigmund [Si1]; see also [Si2, Si3]. Indeed, if (Σ, σ) is a transitive sub-shift of finite type with infinitely many elements then $\mathcal{M}(\Sigma, \sigma)$ is affinely homeomorphic to the Poulsen simplex. Note that the extreme points in this setting correspond to the ergodic measures.

This article describes the space of invariant measures for transitive one-sided countable state Markov shifts (see Section 2 for precise definitions). The major difference with previous work on the subject is that the phase space is no longer compact and therefore the escape of mass phenomenon has to be taken into account. Notions of convergence in the space of measures are required to describe loss of mass. Indeed, the weak* topology preserves the total mass of the space, thus it can not capture the escape of mass. For locally compact spaces the space of invariant measures can be endowed with the vague topology; in this context it is possible for mass to be lost. We introduce a new notion of convergence in the space of measures, the so-called **topology of convergence on cylinders**, that generalizes the vague topology. This notion of convergence does not require the underlying space to be locally compact.

In the non-compact setting the space of invariant probability measures is not necessarily compact. The lack of compactness of the space of invariant probability measures is a major difficulty in the development of the corresponding ergodic theory: in many arguments it is natural to take limits of invariant measures and it is important to know that the limiting object is indeed a measure. We stress that this is a very subtle phenomenon, it could happen that for topologies that naturally generalize the weak* topology the limit of a sequence of invariant probability measures is not a countably additive measure. In this paper we will compactify the space of invariant probability measures for a large family of countable Markov shifts, including a wide range of non-locally compact shifts. As we will see, our compactification is strongly related to the escape of mass phenomenon.

For completeness we will briefly describe the topology on the space of invariant sub-probability measures we will focus on in this work. Let (Σ, σ) be a transitive countable Markov shift and $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ the space of σ-invariant sub-probability measures on Σ (for precise definitions we refer the reader to Sections 2 and 3). We say that $(\mu_n)_n \subset \mathcal{M}_{\leq 1}(\Sigma, \sigma)$ converges on cylinders to μ if

$$\lim_{n \to \infty} \mu_n(C) = \mu(C),$$

for every cylinder $C \subset \Sigma$. This notion of convergence induces a topology, the topology of convergence on cylinders. We prove that this topology is metrizable (see Proposition 3.15). For general facts about the topology of convergence on cylinders we refer the reader to Section 3.3.

We consider a large class of countable Markov shifts that satisfy the so-called $\mathcal{F}$-property (see Definition 4.9). This includes locally compact and finite entropy non-locally compact countable Markov shifts. The $\mathcal{F}$-property essentially rules out the possibility of having infinitely many periodic orbits of a given length that intersect a fixed cylinder. One of the main results of this work is

Theorem 1.2. *Let (Σ, σ) be a transitive countable Markov shift satisfying the $\mathcal{F}$-property. Then the space of invariant sub-probability measures $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ endowed with the topology of convergence on cylinders is affine homeomorphic to the Poulsen simplex. In particular, $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is compact with respect to the topology of convergence on cylinders.*

The topology of convergence on cylinders restricted to $\mathcal{M}(\Sigma, \sigma)$ coincides with the weak* topology (see Lemma 3.17). Theorem 1.2 has the following corollary.

Corollary 1.3. *Let (Σ, σ) be a transitive countable Markov shift satisfying the $\mathcal{F}$-property. Then the space of invariant probability measures $\mathcal{M}(\Sigma, \sigma)$ endowed with the weak* topology is affine homeomorphic to the Poulsen simplex minus a vertex and all of its convex combinations.*

In order to prove Theorem 1.2 we will need to prove three key properties: (1) there exists a sequence of ergodic measures converging on cylinders to the zero measure, (2) every sequence of periodic measures has an accumulation point (in the topology of convergence on cylinders) which is a countably additive measure, and (3) the set of periodic measures is weak* dense in $\mathcal{M}(\Sigma, \sigma)$. While point (3) is fairly standard and uses shadowing and closing properties of the shift, the other two are more subtle. Indeed, a combinatorial assumption is required on (Σ, σ) for these properties to hold (hence the $\mathcal{F}$-property assumption).

It is worth pointing out that Theorem 1.2 is optimal for the topology of convergence on cylinders. More precisely, if (Σ, σ) does not satisfy the $\mathcal{F}$-property, then $\mathcal{M}(\Sigma, \sigma)$ contains a sequence of periodic measures which converges on cylinders to a finitely additive measure which is not countably additive (see Proposition 4.19). In particular, if we want to compactify $\mathcal{M}(\Sigma, \sigma)$ we must give up the convergence on all cylinders (which does not seem reasonable) or modify the topology in a more substantial way.

We also study suspension flows defined over countable Markov shifts. These are continuous time dynamical systems defined over non-compact spaces. The suspensions we consider are constructed over arbitrary countable Markov shifts (Σ, σ) and for roof functions $\tau : \Sigma \to \mathbb{R}$ belonging to a class that we denote by $\mathcal{R}$ (see Definition 6.2). If τ is bounded away from zero there is a one-to-one correspondence between the space of invariant probability measures for the flow, which we denote by $\mathcal{M}(\Sigma, \sigma, \tau)$, and $\mathcal{M}_\tau = \{ \mu \in \mathcal{M}(\Sigma, \sigma) : \int \tau d\mu < \infty \}$. The space of sub-probability measures invariant by the suspension flow is denoted by $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$. In Section 6 we define a topology on $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$, the **topology of convergence on cylinders for the suspension flow**, that shares many properties with the topology of convergence on cylinders on $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ (see Definition 6.4).

The class of suspension flows that we study includes a wide range of symbolic models for geometric systems. For example, the symbolic model of the geodesic flow over the modular surface satisfies all of our assumptions. In this context we prove

Theorem 1.4. *Let (Σ, σ) be a transitive countable Markov shift and $\tau \in \mathcal{R}$. Then the space of invariant sub-probability measures of the suspension flow $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$, endowed with the topology of convergence on cylinders for the suspension flow, is affine homeomorphic to the Poulsen simplex. In particular, $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$ is compact with respect to the topology of convergence on cylinders.*

Some countable Markov shifts without the $\mathcal{F}$-property are particularly important (for instance, the full shift, or shifts with the BIP property [Sa2]), and we do want to have some understanding on their spaces of invariant probability measures. The work done in Section 6 and our auxiliary potential τ allows us to regain control in this setting. Indeed, in this general context we are able to describe the set of invariant probability measures for which the function τ is integrable (see Lemma 6.1 and Theorem 6.18).

We stress that the compactifications constructed in this paper have several interesting applications to the thermodynamic formalism of countable Markov shifts. For instance, in joint work with M. Todd [ITV], we consider finite entropy countable Markov shifts (Σ, σ) and study the behavior of the measure theoretic entropy of a sequence $(\mu_n)_n \subset \mathcal{M}(\Sigma, \sigma)$. By the results in this paper there exists a measure $\mu \in \mathcal{M}_{\leq 1}(\Sigma, \sigma)$ and a sub-sequence of $(\mu_n)_n$ which converges to μ in the cylinder topology. This property is used to establish stability results for the measure of maximal entropy, relate the escape of mass with the entropy of the system and to prove upper-semi continuity of the entropy map. In the compact setting there is no escape of mass and the properties of the entropy map and the

measure of maximal entropy are classical [Wa, Chapter 8], but in the non-compact case new ideas were needed. We also mention that the results in [ITV] can be pushed even further to include potentials; this is discussed in [V] by the second author.

Finally, we remark that over the last few years countable Markov shifts have been used to code relevant parts of the dynamics for a wide range of dynamical systems. For example, it was shown by Sarig [Sa2] that countable Markov partitions can be constructed for positive entropy diffeomorphisms defined on compact surfaces. The corresponding symbolic coding captures positive entropy measures. These results have recently been used to prove that C^∞ surface diffeomorphisms of positive entropy have at most finitely many measures of maximal entropy [BCS]. In a different direction, countable Markov partitions have been constructed for Sinai and Bunimovich billiards; this has allowed for the proof of lower bounds on the number of periodic orbits of a given period [LM]. Based on the work of Sarig, countable Markov partitions have been constructed for large classes of dynamical systems. Our results apply not only to all countable Markov shifts obtained as symbolic codings of these systems, but also to several symbolic models that are not locally-compact, for instance, symbolic codings of interval maps having a parabolic fixed point [H, MP, Sa1] or loop systems [BBG].

2   Countable Markov shifts

In this section we define the dynamical systems that will be studied throughout the article. Let B be a transition matrix defined on the alphabet of natural numbers. That is, the entries of the matrix $B = B(i, j)_{\mathbb{N} \times \mathbb{N}}$ are zeros and ones (with no row and no column made entirely of zeros). The one-sided countable state Markov shift (Σ, σ) defined by the matrix B is the set

$$\Sigma := \{ (x_n)_{n \in \mathbb{N}} : B(x_n, x_{n+1}) = 1 \text{ for every } n \in \mathbb{N} \},$$

together with the shift map $\sigma : \Sigma \to \Sigma$ defined by $\sigma(x_1, x_2, \dots) = (x_2, x_3, \dots)$. In what follows we will simply call (Σ, σ), **countable Markov shift**. For $(a_1, \dots, a_n) \in \mathbb{N}^n$, we define a cylinder set $[a_1 \dots a_n]$ of length n by

$$[a_1 \dots a_n] := \{x \in \Sigma : x_j = a_j \text{ for } 1 \le j \le n\}.$$

We endow Σ with the topology generated by cylinder sets. This is a metrizable non-compact space. Indeed, let $\tilde{d} : \Sigma \times \Sigma \to \mathbb{R}$ be the function defined by

$$(2.1) \qquad \tilde{d}(x, y) := \begin{cases} 1 & \text{if } x_0 \ne y_0; \\ 2^{-k} & \text{if } x_i = y_i \text{ for } i \in \{0, \dots, k-1\} \text{ and } x_k \ne y_k; \\ 0 & \text{if } x = y. \end{cases}$$

The function $\tilde{d}$ is a metric and it generates the same topology as that of the cylinder sets.

A countable Markov shift defined by the transition matrix $B = B(i, j)_{\mathbb{N} \times \mathbb{N}}$ is locally compact if and only if for every $i \in \mathbb{N}$ we have $\sum_{j=1}^{\infty} B(i, j) < \infty$ (see [Ki, Observation 7.2.3]).

An **admissible word** is a word $w = \mathbf{a_1} \cdots \mathbf{a_n}$, where $\mathbf{a_i} \in \mathbb{N}$ and $[a_1, \dots, a_n]$ is non-empty. To emphasize the difference between admissible words and points in Σ we use bold letters for admissible words.

Definition 2.1. A countable Markov shift (Σ, σ) is **transitive** if for all open sets $U, V \subset \Sigma$ there exists $n \in \mathbb{N}$ such that $U \cap \sigma^{-n}V \ne \emptyset$. Similarly, a countable Markov shift (Σ, σ) is **topologically mixing** if for all open sets $U, V \subset \Sigma$ there exists $N(U, V) \in \mathbb{N}$ such that for every $n > N(U, V)$ we have $U \cap \sigma^{-n}V \ne \emptyset$.

Let $\varphi : \Sigma \to \mathbb{R}$ be a function. We define $\text{var}_n(\varphi) = \sup_{x,y} |\varphi(x) - \varphi(y)|$, where the supremum runs over points x and y satisfying $\tilde{d}(x, y) \le 2^{-n}$. Observe that a function φ is uniformly continuous if and only if $\text{var}_n(\varphi)$ goes to zero as n goes to infinity. A potential φ has summable variations if $\sum_{k=2}^{\infty} \text{var}_k(\varphi)$ is finite.

In the late 1960's Gurevich [Gu1, Gu2] introduced a suitable notion of entropy in this setting. Note that since the space Σ is not compact the classical definition of topological entropy obtained by means of (n, ε)-separated sets (see [Wa, Chapter 7]) depends upon the metric, that is, two equivalent metrics can yield different numbers. Since the entropy of an invariant measure depends only on the Borel structure and not on the metric, this is a major problem if the entropy is to satisfy a variational principle. Gurevich introduced the following notion of entropy:

$$h(\sigma) := \limsup_{n \to \infty} \frac{1}{n} \log \sum_{x : \sigma^n x = x} 1_{[a]}(x),$$

where $a \in \mathbb{N}$ is an arbitrary symbol and $1_{[a]}$ is the characteristic function of the cylinder $[a]$. Gurevich proved that this value is independent of the symbol a if (Σ, σ) is transitive and that the limit exists if (Σ, σ) is topologically mixing. Moreover, he also proved that this notion of entropy is the correct one in the sense that it satisfies the variational principle, that is

$$h(\sigma) = \sup\{h(\mu) : \mu \in \mathcal{M}(\Sigma, \sigma)\},$$

where $h(\mu)$ is the entropy of the invariant measure μ (see [Wa, Chapter 4]) and $\mathcal{M}(\Sigma, \sigma)$ is the space of invariant probability measures.

3 Topologies in the space of measures

In this section we recall definitions and properties of the weak* and the vague topologies and define a new notion of convergence in the space of probability measures on Σ, namely the topology of convergence on cylinders. It is with respect to these three topologies that we will describe the space of invariant probability measures $\mathcal{M}(\Sigma, \sigma)$. It is worth emphasizing that the weak* topology does not allow escape of mass, but the vague topology and the topology of convergence on cylinders do allow it.

3.1 The weak* topology. Let (X, ρ) be a metric space. We denote by $C_b(X)$ the space of bounded continuous function $f : X \to \mathbb{R}$. We endow $C_b(X)$ with the C^0-topology. This is the topology induced by the norm $\|f\| = \sup_{x \in X} |f(x)|$. It is a standard fact that $C_b(X)$ is a Banach space. Denote by $\mathcal{M}(X)$ the set of Borel probability measures on the metric space (X, ρ). Our first notion of convergence in this set is the following,

Definition 3.1. A sequence of probability measures $(\mu_n)_n \subset \mathcal{M}(X)$ converges to a measure μ in the weak* topology if for every $f \in C_b(X)$ we have

$$\lim_{n \to \infty} \int f d\mu_n = \int f d\mu.$$

Remark 3.2. Note that in this notion of convergence we can replace the set of test functions (bounded and continuous) by the space of bounded uniformly continuous functions (see [B, 8.3.1 Remark]) or by the space of bounded Lipschitz functions (see [Kl, Theorem 13.16 (ii)]). That is, if for every bounded uniformly continuous (or bounded Lipschitz) function $f : X \to \mathbb{R}$ we have

$$\lim_{n \to \infty} \int f d\mu_n = \int f d\mu,$$

then the sequence $(\mu_n)_n$ converges in the weak* topology to μ.

The weak* topology is the coarsest topology such that for every $f \in C_b(X)$ the map $\mu \to \int f \, d\mu$, with $\mu \in \mathcal{M}(X)$, is continuous. The following classical result characterizes weak* convergence (see [Bi, Theorem 2.1]).

Proposition 3.3 (Portmanteau Theorem). *Let $(\mu_n)_n$, μ be probability measures on X. The following statements are equivalent:*

(a) *The sequence $(\mu_n)_n$ converges to μ in the weak* topology.*

(b) *For every open set $O \subset X$, the following holds: $\mu(O) \leq \liminf_{n \to \infty} \mu_n(O)$.*

(c) *For every closed set $C \subset X$, the following holds: $\mu(C) \geq \limsup_{n \to \infty} \mu_n(C)$.*

(d) *For every set $A \subset X$ such that $\mu(\partial A) = 0$, the following holds:*

$$\mu(A) = \lim_{n \to \infty} \mu_n(A).$$

A relevant feature of the weak* convergence is that there is no loss of mass since the constant function equal to one belongs to $C_b(X)$. That is,

Remark 3.4. If the sequence of probability measures $(\mu_n)_n$ converges in the weak* topology to μ, then μ is also a probability measure.

Also note that if the space (X, ρ) is compact, then the space of Borel probability measures, $\mathcal{M}(X)$, is also compact with respect to the weak* topology (see [Wa, Theorem 6.5]). An interesting fact is that if (X, ρ) is a separable metric space then $\mathcal{M}(X)$ can be metrized as a separable metric space (see [Par2, Theorem 6.2]). Actually, there exists an explicit metric that generates the weak* topology and for which $\mathcal{M}(X)$ is separable if X is separable. This is the so-called Prohorov metric (see [Bi, pp. 72–73]). Therefore, if X is a separable metric space then so is $\mathcal{M}(X)$, despite the fact that the space $C_b(X)$ might not be separable. Actually, if (Σ, σ) is a non-compact countable Markov shift then $C_b(\Sigma)$ is not separable, as below.

Remark 3.5. Let (Σ, σ) be a countable (non-compact) Markov shift; then $C_b(\Sigma)$ is not separable. Indeed, let $(x_n)_n$ be a fixed sequence of elements of Σ with $x_n \in [n]$, where $[n] := \{(y_1, y_2, \ldots) \in \Sigma : y_1 = n\}$. Define the set

$$\mathcal{L} := \{\varphi \in C_b(\Sigma) : \varphi(x_n) = 0 \text{ or } \varphi(x_n) = 1, \text{ for every } n \in \mathbb{N}\}.$$

Note that $\mathcal{L}$ contains uncountably many elements. For every countable subset $\mathcal{C} \subset C_b(\Sigma)$ there exists $\varphi \in \mathcal{L}$ such that for every $\psi \in \mathcal{C}$ we have

$$\|\varphi - \psi\| \geq \frac{1}{2},$$

hence $C_b(\Sigma)$ is not separable.

Let $T : (X, \rho) \rightarrow (X, \rho)$ be a continuous dynamical system defined on a metric space. We denote by $\mathcal{M}(X, T)$ the space of T-invariant probability measures. In the following lemma we collect relevant information regarding the structure of $\mathcal{M}(X, T)$.

Lemma 3.6. *Let $T : (X, \rho) \rightarrow (X, \rho)$ be a continuous dynamical system defined on a metric space. Then:*
  (a) *The space $\mathcal{M}(X, T)$, as a subset of $\mathcal{M}(X)$, is closed in the weak* topology ([Wa, Theorem 6.10]).*
  (b) *If X is compact, then so is $\mathcal{M}(X, T)$ with respect to the weak* topology (see [Wa, Theorem 6.10]).*
  (c) *The space $\mathcal{M}(X, T)$ is a convex set for which its extreme points are the ergodic measures (see [Wa, Theorem 6.10]). It is actually a Choquet simplex (each measure is represented in a unique way as a generalized convex combination of the ergodic measures [Wa, p. 153]).*

Definition 3.7. Let $T : (X, \rho) \rightarrow (X, \rho)$ be a continuous dynamical system defined on a metric space. We denote by $\mathcal{M}_e(X, T)$ the set of ergodic T-invariant probability measures. An ergodic measure $\mu \in \mathcal{M}_e(X, T)$ supported on a periodic orbit will be called a **periodic measure**. Denote by $\mathcal{M}_p(X, T)$ the set of periodic measures.

The next result was obtained by Coudène and Schapira [CS, Section 6] as a consequence of shadowing and the Anosov closing Lemma.

Theorem 3.8. *Let (Σ, σ) be a transitive countable Markov shift. Then $\mathcal{M}_p(\Sigma, \sigma)$ is dense in $\mathcal{M}(\Sigma, \sigma)$ with respect to the weak* topology.*

3.2 The vague topology. Let (X, ρ) be a locally compact metric space. Denote by $M_{\leq 1}(X)$ the set of Borel non-negative measures on X such that $\mu(X) \leq 1$. The set of continuous functions of compact support, that is continuous functions $f : X \rightarrow \mathbb{R}$ for which the closure of the set $\{x \in X : f(x) \neq 0\}$ is compact, will be denoted by $C_c(X)$. Note that $C_c(X) \subset C_b(X)$. We will consider the following notion of convergence.

Definition 3.9. A sequence $(\mu_n)_n \subset \mathcal{M}_{\leq 1}(X)$ converges to $\mu \in \mathcal{M}_{\leq 1}(X)$ in the vague topology if for every $f \in C_c(\Sigma)$ we have

$$\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu.$$

The vague topology is the coarsest topology on $\mathcal{M}_{\leq 1}(X)$ such that for every f continuous and of compact support, the map $\mu \to \int f d\mu$ is continuous. We stress that the total mass is not necessarily preserved in the vague topology. A sequence of probability measures can converge in the vague topology to a non-negative measure of total mass less than or equal to one. If X is compact, then $C_c(X) = C_b(X)$ and therefore the vague topology coincides with the weak* topology. We collect the following results:

Remark 3.10. Let (X, ρ) be a metric space. Note that the weak* topology extends to $\mathcal{M}_{\leq 1}(X)$.

(a) If X is compact, then $\mathcal{M}_{\leq 1}(X)$ is compact with respect to the weak* topology (see [Kl, Corollary 13.30]).

(b) If X is a locally compact separable metric space, then $\mathcal{M}_{\leq 1}(X)$ is compact with respect to the vague topology (see [Kl, Corollary 13.31]) and metrizable (see [Di, 13.4.2]).

(c) Let X be a locally compact separable metric space. The sequence $(\mu_n)_n$ converges vaguely to μ and $\lim_{n\to\infty} \mu_n(X) = \mu(X)$ if and only if $(\mu_n)_n$ converges in weak* topology to μ (see [Kl, Theorem 13.16]).

(d) Let X be a locally compact separable metric space. The sequence $(\mu_n)_n$ converges vaguely to μ and the sequence $(\mu_n)_n$ is tight if and only if $(\mu_n)_n$ converges in weak* topology to μ (see [Kl, Theorem 13.35]).

Let $T : (X, \rho) \to (X, \rho)$ be a continuous dynamical system defined on a metric space. If $(\mu_n)_n \subset \mathcal{M}(X, T)$ is a sequence of T-invariant probability measures that converges in the vague topology to a non-zero measure μ, then the normalized measure $\mu(\cdot)/\mu(X)$ is a T-invariant probability. We call the measure μ an invariant sub-probability and denote by $\mathcal{M}_{\leq 1}(X, T)$ the space of T-invariant sub-probability measures. Observe that the zero measure belongs to $\mathcal{M}_{\leq 1}(X, T)$.

Lemma 3.11. *Let* $T : (X, \rho) \to (X, \rho)$ *be a continuous dynamical system defined on a metric space, then*

(a) *The space* $\mathcal{M}(X, T)$ *is a closed subset of* $\mathcal{M}_{\leq 1}(X, T)$ *in the weak* topology.*

(b) *If* X *is a locally compact separable metric space, then the space* $\mathcal{M}_{\leq 1}(X, T)$ *is compact in the vague topology.*

Proof. The first claim is a consequence of Lemma 3.6, while the second follows from Remark 3.10. □

Proposition 3.12. *Let* $T : (X, \rho) \to (X, \rho)$ *be a continuous dynamical system defined on a metric space. Then the space* $\mathcal{M}_{\leq 1}(X, T)$ *is a convex set and its extreme points are the ergodic probability measures and the zero measure.*

Proof. The convexity of the space $M_{\leq 1}(X, T)$ is direct. Note that every invariant sub-probability $\mu \in \mathcal{M}_{\leq 1}(X, T)$, with $0 < \mu(X) < 1$, is the convex combination of a measure in $\mathcal{M}(X, T)$ and the zero measure. Also note that the zero measure is not the convex combination of any set of positive measures. The result then follows from Lemma 3.6. □

3.3   The topology of convergence on cylinders.

Several relevant countable Markov shifts are not locally compact. Therefore a good notion of convergence in the space of sub-probabilities is required in this setting. The vague topology is of no use in the non-locally compact setting since in this case the space $C_c(X)$ might be empty, as below.

Remark 3.13. If Σ is a non-locally compact transitive countable Markov shift then $C_c(\Sigma) = \emptyset$. Indeed, if $K \subset \Sigma$ is a compact set then it must have empty interior (see [Ki, Observation 7.2.3 (iv)]). Therefore $K = \partial K$, that is, the compact set equals its topological boundary. If $f \in C_c(\Sigma)$ then consider the open set

$$A := \{x \in \Sigma : f(x) \neq 0\} = f^{-1}(\mathbb{R} \smallsetminus \{0\}).$$

Note that the support of f is $K = \overline{A}$, that is the closure of the set A. But since $f \in C_c(\Sigma)$ the set K is compact. This means that the compact set K contains an open set A, contradicting the fact that $K = \partial K$. Therefore $C_c(\Sigma) = \emptyset$.

We now define the notion of convergence—that generalizes the vague topology to the non-locally compact setting—which is the main topic of this work.

Definition 3.14. Let (Σ, σ) be a countable Markov shift and $(\mu_n)_n$, μ invariant sub-probability measures. We say that a sequence $(\mu_n)_n$ **converges on cylinders** to μ if $\lim_{n \to \infty} \mu_n(C) = \mu(C)$, for every cylinder $C \subset \Sigma$. The topology on $\mathcal{M}_{\leq 1}(\Sigma)$ induced by this convergence is called the **topology of convergence on cylinders**.

We emphasize that this notion of convergence induces a topology because the collection of cylinders is countable and it is a basis for the topology on Σ. For brevity we will frequently say that $(\mu_n)_n \subset \mathcal{M}_{\leq 1}(\Sigma)$ **converges on cylinders** to μ if the sequence converges in the topology of convergence on cylinders.

Proposition 3.15. *The topology of convergence on cylinders on $\mathcal{M}_{\leq 1}(\Sigma)$ is metrizable.*

Proof. Consider the metric

$$(3.1) \qquad d(\mu, \nu) = \sum_{n \geq 1} \frac{1}{2^n} |\mu(C_n) - \nu(C_n)|,$$

where $(C_n)_n$ is some enumeration of the cylinders on Σ. Note that $d(\mu, \nu) = 0$, if and only if $\mu = \nu$. Indeed, if $d(\mu, \nu) = 0$ then $\mu(C) = \nu(C)$, for every cylinder C. By the outer regularity of Borel measures on a metric space we conclude that this is equivalent to saying that $\mu = \nu$. Symmetry is clear and the triangle inequality follows directly from the triangle inequality in $\mathbb{R}$. It is clear from the definition of d that it induces the desired notion of convergence on $\mathcal{M}_{\leq 1}(\Sigma)$, that is, it generates the topology of convergence on cylinders.   □

It worth pointing out that since $C_b(\Sigma)$ is not separable we can not endow the weak* topology with a metric like d.

Remark 3.16. Note that the topology of convergence on cylinders, like the vague topology, allows for mass to escape. Indeed, let Σ be the full shift on $\mathbb{N}$, which is not locally compact. Denote by $\delta_{\overline{n}}$ the atomic measure supported on the point $\overline{n} := (n, n, n, \dots)$. Then the sequence $(\delta_{\overline{n}})_n$ converges in the topology of convergence on cylinders to the zero measure.

Despite Remark 3.16 the topology of convergence on cylinders is closely related to the weak* topology. If there is no loss of mass both notions coincide.

Lemma 3.17. *Let* (Σ, σ) *be a countable Markov shift,* μ *and* $(\mu_n)_n$ *be probability measures on* Σ. *The following assertions are equivalent:*
  (a) *The sequence* $(\mu_n)_n$ *converges in the weak* topology to* μ.
  (b) *The sequence* $(\mu_n)_n$ *converges on cylinders to* μ.

Proof. First assume that $(\mu_n)_n$ converges in the weak* topology to μ. Define $f := 1_C \in C_b(\Sigma)$, where 1_C is the characteristic function of a cylinder. The weak* convergence implies that $\lim_{n \to \infty} \int f d\mu_n = \int f d\mu$, which is equivalent to say that

$$\lim_{n \to \infty} \mu_n(C) = \mu(C).$$

Since the cylinder C was chosen arbitrarily we conclude that $(\mu_n)_n$ converges on cylinders to μ. Now assume that $(\mu_n)_n$ converges on cylinders to μ. Observe that an open set $\mathcal{O}$ can be written as a countable union of disjoint cylinders, saying $\mathcal{O} = \bigcup_{k \geq 1} C_k$. Therefore

$$\liminf_{n \to \infty} \mu_n(\mathcal{O}) \geq \liminf_{n \to \infty} \mu_n \left(\bigcup_{k=1}^{M} C_k \right) = \mu \left(\bigcup_{k=1}^{M} C_k \right),$$

for every M. We conclude that

$$\liminf_{n \to \infty} \mu_n(\mathcal{O}) \geq \mu(\mathcal{O}).$$

Proposition 3.3 implies that $(\mu_n)_n$ converges in the weak* topology to μ.   □

We will now prove that the topology of convergence on cylinders generalizes the vague topology. More precisely, on locally compact countable Markov shifts both topologies coincide.

Lemma 3.18. *Let (Σ, σ) be a locally compact countable Markov shift and $\mu, (\mu_n)_n \in \mathcal{M}_{\leq 1}(\Sigma)$. The following assertions are equivalent:*

(a) *The sequence $(\mu_n)_n$ converges in the vague topology to μ.*

(b) *The sequence $(\mu_n)_n$ converges on cylinders to μ.*

Proof. First note that if Σ is locally compact then every cylinder is a compact set (see [Ki, Observation 7.2.3]). Assume that $(\mu_n)_n$ converges in the vague topology to μ. Let $C \in \Sigma$ be a cylinder set; then the characteristic function of C, denoted by 1_C, belongs to $C_c(\Sigma)$. Thus,

$$\lim_{n \to \infty} \int 1_C d\mu_n = \int 1_C d\mu.$$

Therefore, the sequence $(\mu_n)_n$ converges on cylinders to μ.

Suppose now that $(\mu_n)_n$ converges on cylinders to μ and let $f \in C_c(\Sigma)$. We will prove that $\lim_{n\to\infty} \int f d\mu_n = \int f d\mu$. The function f is uniformly continuous, in particular for every $\varepsilon > 0$ there exists $n = n(\varepsilon) \in \mathbb{N}$ such that $\mathrm{var}_n(f) \leq \varepsilon$. Since the support of f is a compact set there exists $M \in \mathbb{N}$ such that it is contained in $\bigcup_{i=1}^{M} [i]$. By the locally compactness of Σ there are finitely many (non-empty) cylinders of length n, that we denote by $(C_i)_{i=1}^{q}$, intersecting $\bigcup_{i=1}^{M} [i]$. Note that if a cylinder of length n intersects a cylinder of length one then it is contained in it, therefore $\bigcup_{i=1}^{q} C_i = \bigcup_{i=1}^{M} [i]$. We now define a locally constant function $\tilde{f} : \Sigma \to \mathbb{R}$ that approximates f. For every $i \in \{1, \ldots, q\}$ choose a point $x_i \in C_i$ and let $\tilde{f} : \Sigma \to \mathbb{R}$ be the function defined by

$$\tilde{f}(x) := \begin{cases} f(x_i) & \text{if } x \in C_i, \text{ for } i \in \{1, \ldots, q\}; \\ 0 & \text{if } x \notin \bigcup_{i=1}^{q} C_i. \end{cases}$$

By construction the function $\tilde{f}$ is locally constant depending only on the first n coordinates, thus $\mathrm{var}_n(\tilde{f}) = 0$. Moreover, it is zero on the complement of the set $\bigcup_{i=1}^{M} [i]$. In particular $\tilde{f} = \sum_{i=1}^{q} a_i 1_{C_i}$, for some sequence of real numbers $(a_i)_{i=1}^{q}$. By the definition of the topology of convergence on cylinders we know that $\lim_{m\to\infty} \int \tilde{f} d\mu_m = \int \tilde{f} d\mu$. Moreover, it follows from our construction that

$$\|f - \tilde{f}\| = \sup_{x \in \Sigma} |f(x) - \tilde{f}| \leq \varepsilon.$$

The construction of $\tilde{f}$ can be made for every $\varepsilon > 0$. In particular, we can construct a sequence $(\tilde{f}_k)_k$ such that $\|f - \tilde{f}_k\| \leq 1/k$, and $\lim_{m\to\infty} \tilde{f}_k d\mu_m = \int \tilde{f}_k d\mu$. This immediately implies that $\lim_{m\to\infty} \int f d\mu_m = \int f d\mu$, which completes the proof. $\square$

3.4   The space of test functions for the topology of convergence on cylinders.   The space of test functions for the weak* topology is $C_b(\Sigma)$. Similarly, the space of test functions for the vague topology is $C_c(\Sigma)$. For duality reasons it is actually convenient to have a Banach space as the space of test functions.   For the vague topology this is not a serious issue, we can simply consider the closure of $C_c(\Sigma)$ in $C_b(\Sigma)$; this gives us the space $C_0(\Sigma)$ of functions that vanish at infinity. More precisely, a function $f \in C_0(\Sigma)$ if it is a continuous function such that for every $\varepsilon > 0$ there exists a compact set $K \subset \Sigma$ such that for every $x \in \Sigma \smallsetminus K$ we have $|f(x)| < \varepsilon$.

It is a natural question to determine what is the space of test functions for the topology of convergence on cylinders. More precisely, determine a Banach space V such that $(\mu_n)_n$ converges on cylinders to μ if and only if $\lim_{n\to\infty} \int f d\mu_n = \int f d\mu$ for every $f \in V$. From the definition of the topology of convergence on cylinders (Definition 3.14) it is clear that the space V should contain the characteristic function of a cylinder and finite linear combination of those. Define

$$H := \left\{ f \in C_b(\Sigma) : f = \sum_{i=1}^{n} a_i 1_{C_i}, \text{ where } a_i \in \mathbb{R} \text{ and } C_i \text{ is a cylinder for each } i \right\}.$$

As with the vague topology, our space of test functions will be the closure of H in $C_b(\Sigma)$, that we denote by $\overline{H}$. The following is direct from the definition of the topology of convergence on cylinders.

Lemma 3.19. *Let (Σ, σ) be a countable Markov shift and $\mu, (\mu_n)_n \subset \mathcal{M}_{\leq 1}(\Sigma)$. Then $(\mu_n)_n$ converges on cylinders to the measure μ if and only if for every $f \in \overline{H}$ we have*

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu.$$

In what follows we will characterize the space $\overline{H}$; in order to do so we will require the following notions.

Definition 3.20. Let (Σ, σ) be a countable Markov shift and $f : \Sigma \to \mathbb{R}$ a function. If C is a cylinder of length m, denote

$$C(\geq n) := \left\{ x \in C : \sigma^m(x) \in \bigcup_{k\geq n}[k] \right\}.$$

For a non-empty set $A \subset \Sigma$ we define

$$\text{var}^A(f) := \sup\{ |f(x) - f(y)| : (x, y) \in A \times A\}.$$

We declare $\text{var}^A(f) = 0$ if A is the empty set.

Lemma 3.21. *A function $f \in C_b(\Sigma)$ belongs to $\bar{H}$ if and only if the following three conditions hold:*

  (a) *f is uniformly continuous.*
  (b) *$\lim_{n \to \infty} \sup_{x \in [n]} |f(x)| = 0$.*
  (c) *$\lim_{n \to \infty} \mathrm{var}^{C(\geq n)}(f) = 0$, for every cylinder $C \subset \Sigma$.*

Moreover, if Σ is locally compact, then $\bar{H}$ coincides with $C_0(\Sigma)$, the space of functions that vanish at infinity.

Proof. Denote by H_0 the space of bounded functions satisfying conditions (a), (b) and (c). The space H_0 is a closed subset of $C_b(\Sigma)$. The inclusion $H \subset H_0$ follows directly from the definition of H. Since H_0 is closed we obtain that $\bar{H} \subset H_0$.

We will now prove that $H_0 \subset \bar{H}$. Fix $f \in H_0$ and $\varepsilon > 0$. We will construct a function $g \in H$ such that $\|f - g\| < \varepsilon$. This would imply that $f \in \bar{H}$.

Since f is uniformly continuous, there exists $q \in \mathbb{N}$ such that $\mathrm{var}_q(f) < \varepsilon$. Let $n_1 \in \mathbb{N}$ be such that

$$\sup_{x \in [m]} |f(x)| < \varepsilon,$$

whenever $m > n_1$. Choose $n_2 \in \mathbb{N}$ such that $\mathrm{var}^{[i](\geq m)}(f) < \varepsilon$, whenever $i \in \{1, \dots, n_1\}$ and $m > n_2$. Similarly, choose $n_3 \in \mathbb{N}$ such that $\mathrm{var}^{[ij](\geq m)}(f) < \varepsilon$, whenever $(i, j) \in \prod_{s=1}^2 \{1, \dots, n_s\}$ and $m > n_3$. Inductively, we obtain a sequence $\{n_1, \dots, n_q, n_{q+1}\}$ such that for every $k \in \{1, \dots, q\}$ we have $\mathrm{var}^{[i_1, \dots, i_k](\geq m)}(f) < \varepsilon$, whenever $(i_1, \dots, i_k) \in \prod_{s=1}^k \{1, \dots, n_s\}$ and $m > n_{k+1}$.

Let $f^* \in H_0$ and C be a non-empty cylinder of length q. We will define a number that depends on f^* and C, which we denote by $l_{f^*}(C)$, as follows. Let us first assume that $C(\geq n) = \emptyset$, for some $n \in \mathbb{N}$. In this case we define $l_{f^*}(C) = 0$. Now assume there exists a strictly increasing sequence $(n_k)_k \subset \mathbb{N}$ and points $x_k \in C \cap \sigma^{-q}[n_k]$. In this case we define $l_{f^*}(C) := \lim_{k \to \infty} f^*(x_k)$. It follows from condition (c) that $l_{f^*}(C)$ is well defined: it is independent of the sequences $(n_k)_k$ and $(x_k)_k$.

A point $(a_1, \dots, a_k) \in \prod_{s=1}^k \{1, \dots, n_s\}$ defines the cylinder $[a_1, \dots, a_k]$. Collect all the non-empty cylinders that arise in this way and call this set Ω_k. Define $\Omega = \bigcup_{k=1}^q \Omega_k$.

Let us first prove the result when $l_f(C) = 0$ holds for every $C \in \Omega$. By our choice of $\{n_1, \dots, n_{q+1}\}$ and the assumption $l_f(C) = 0$, we know that for every $k \in \{1, \dots, q\}$ we have

(3.2)
$$\sup_{x \in [i_1, \dots, i_k](\geq m)} |f(x)| < \varepsilon,$$

whenever $(i_1, \dots, i_k) \in \prod_{s=1}^k \{1, \dots, n_s\}$, and $m > n_{k+1}$ (for consistency define the supremum over the empty set as zero). For every $C \in \Omega_q$ choose a point $x_C \in C$.

Define $g := \sum_{C \in \Omega_q} f(x_C) 1_C$. Observe that if $x \in \bigcup_{C \in \Omega_q} C$, then $|f(x) - g(x)| < \varepsilon$ (recall that $\text{var}_q(f) < \varepsilon$). If x does not belong to $\bigcup_{C \in \Omega_q} C$, then x belongs to a cylinder of the form $[i_1, \dots, i_k, b]$, where $(i_1, \dots, i_k) \in \prod_{s=1}^{k} \{1, \dots, n_s\}$ and $b > n_{k+1}$ (if $k = 0$, then $x \in [b]$ where $b > n_1$). By (3.2) and our choice of n_1 we obtain that $|f(x)| < \varepsilon$. We therefore have $\|f - g\| < \varepsilon$.

We will now explain how to reduce the general case to the situation where $l_f(C) = 0$ holds for every $C \in \Omega$.

Define

$$h_1 := \sum_{C \in \Omega_1} l_f(C) 1_C \quad \text{and} \quad f_1 := f - h_1.$$

We claim that $l_{f_1}(C) = 0$, for every $C \in \Omega_1$. First suppose that there exists n such that $C(\geq n) = \emptyset$. In this case, by definition, we have that $l_{f_1}(C) = 0$. It remains to consider the non-trivial case in the definition of $l_{f_1}(C)$. Let $(n_k)_k \subset \mathbb{N}$ be a strictly increasing sequence and $(x_k)_k$ points in Σ such that $x_k \in C \cap \sigma^{-1}[n_k]$. Then

$$l_{f_1}(C) = \lim_{k \to \infty} f_1(x_k) = \lim_{k \to \infty} (f(x_k) - h_1(x_k)) = l_f(C) - \lim_{k \to \infty} h_1(x_k).$$

Since $h_1 = \sum_{C \in \Omega_1} l_f(C) 1_C$, we know that $h_1(x_k) = l_f(C)$ (observe that all cylinders in Ω_1 have length 1). In particular $l_{f_1}(C) = 0$.

Now define

$$h_2 := \sum_{C \in \Omega_2} l_{f_1}(C) 1_C, \quad \text{and} \quad f_2 := f_1 - h_2.$$

We claim that $l_{f_2}(C) = 0$, for every $C \in \Omega_1 \cup \Omega_2$. As before, first suppose that there exists n such that $C(\geq n) = \emptyset$. In this case, by definition, we have that $l_{f_2}(C) = 0$. It remains to consider the non-trivial case in the definition of $l_{f_2}(C)$. Let $C_1 \in \Omega_1$ and $C_2 \in \Omega_2$. Choose a strictly increasing sequence $(n_k)_k \subset \mathbb{N}$ and points $(x_k)_k$ in Σ such that $x_k \in C_2 \cap \sigma^{-2}[n_k]$. Then

$$l_{f_2}(C_2) = \lim_{k \to \infty} f_2(x_k) = \lim_{k \to \infty} (f_1(x_k) - h_2(x_k)) = l_{f_1}(C_2) - \lim_{k \to \infty} h_2(x_k).$$

As before $h_2(x_k) = l_{f_1}(C_2)$, because $x_k \in C_2 \cap \sigma^{-2}[k]$. We conclude that $l_{f_2}(C_2) = 0$. Similarly, choose sequences $(m_k)_k \subset \mathbb{N}$ and $(y_k)_k \subset \Sigma$ such that $y_k \in C_1 \cap \sigma^{-1}[m_k]$; then

$$l_{f_2}(C_1) = \lim_{k \to \infty} f_2(y_k) = \lim_{k \to \infty} (f_1(y_k) - h_2(y_k)) = l_{f_1}(C_1) - \lim_{k \to \infty} h_2(y_k).$$

Since h_2 is a finite linear combination of indicators of cylinders of length 2 we obtain that $\lim_{k \to \infty} h_2(y_k) = 0$. By construction we have that $l_{f_1}(C_1) = 0$. We conclude that $l_{f_2}(C) = 0$, for every $C \in \Omega_1 \cup \Omega_2$.

Continue this process and define

$$(h_k)_{k=1}^q \quad \text{and} \quad (f_k)_{k=1}^q$$

such that $f_k = f_{k-1} - h_k$, for every $k \in \{1, \dots, q\}$ (where we set $f_0 = f$). By construction $l_{f_k}(C) = 0$ for every cylinder $C \in \bigcup_{i=1}^k \Omega_i$. In particular $l_{f_q}(C) = 0$, for every cylinder $C \in \Omega$. Finally, observe that $f_q = f - \sum_{k=1}^q h_k$. Since $h := \sum_{k=1}^q h_k$ is a function in H, it is enough to approximate f_q (we can add back the function h afterwards, H is a vector space).

We will now assume that Σ is locally compact. Since Σ is locally compact the set $C(\geq m)$ is empty for large enough m. In particular condition (c) is always satisfied. By definition a function $f \in C_b(\Sigma)$ belongs to $C_0(\Sigma)$ if and only if condition (b) holds. In particular condition (b) implies condition (a). Therefore $\overline{H}$ coincides with $C_0(\Sigma)$.   □

By abuse of notation we denote by $C_0(\Sigma)$ the space of test functions for the topology of convergence on cylinders; this is reasonable because $\overline{H} = C_0(\Sigma)$ in the locally compact case. We say that f vanishes at infinity if $f \in C_0(\Sigma)$. To summarize, a function $f \in C_b(\Sigma)$ vanishes at infinity if it satisfies conditions (a), (b) and (c). It follows from the discussion above that the Banach space $C_0(\Sigma)$ is the space of test functions for the topology of convergence on cylinders. In particular, the map $\mu \mapsto \int f d\mu$ is continuous in $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$, whenever $f \in C_0(\Sigma)$.

4   The space of invariant sub-probability measures is compact

We already noticed in Lemma 3.11 that if (Σ, σ) is a locally compact transitive countable Markov shift, the space of invariant sub-probability measures $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is compact with respect to the vague topology. It is a consequence of Lemma 3.18 that $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is also compact with respect to the topology of convergence on cylinders. In this section we prove that the space $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is compact with respect to the topology of convergence on cylinders for a larger class of transitive countable Markov shifts, that is, for countable Markov shifts with the $\mathcal{F}$-property (see Definition 4.9). Our results are also sharp: if (Σ, σ) does not satisfy the $\mathcal{F}$-property, then there are sequences of periodic measures that converge to a finitely additive measure that is not countably additive (see Proposition 4.19). Our next result states that invariance is preserved under limits provided the limiting object is a countably additive measure.

Lemma 4.1. *Let* (Σ, σ) *be a countable Markov shift. Let* $(\mu_n)_n$ *be a sequence of invariant probability measures converging on cylinders to a sub-probability measure* μ. *Then* μ *is an invariant measure.*

Proof. The measure μ is invariant if it is equal to $\sigma_* \mu := \mu(\sigma^{-1})$. Note that, in order to prove the invariance, it is enough to prove that $\sigma_* \mu(D) = \mu(D)$, for every cylinder D. Observe that if D is a cylinder then $\sigma^{-1} D = \bigcup_{i \geq 1} D_i$, where $(D_i)_i$ is a finite or countable collection of cylinders. Since μ_n is invariant we have that $\mu_n(D) = \mu_n(\sigma^{-1} D)$. If $\sigma^{-1} D = \bigcup_{i=1}^m D_i$ is a finite union of cylinders we obtain that

$$\mu(D) = \lim_{n \to \infty} \mu_n(D) = \lim_{n \to \infty} \mu_n(\sigma^{-1} D) = \lim_{n \to \infty} \mu_n\left(\bigcup_{i=1}^m D_i \right) = \mu(\sigma^{-1} D).$$

If $\sigma^{-1} D$ is a union of infinitely many cylinders we have

$$\mu(D) = \lim_{n \to \infty} \mu_n(D) = \lim_{n \to \infty} \mu_n(\sigma^{-1} D) \geq \lim_{n \to \infty} \mu_n\left(\bigcup_{i=1}^M D_i \right) = \mu\left(\bigcup_{i=1}^M D_i \right),$$

for every $M \in \mathbb{N}$. We conclude that $\mu(D) \geq \mu(\sigma^{-1} D)$. We have therefore proved that for every cylinder D we have $\mu(D) \geq \mu(\sigma^{-1} D)$. Suppose D is a cylinder of length s and enumerate all cylinders of length s by $(E_k)_k$ with $E_1 = D$. Since $\mu(E_k) \geq \mu(\sigma^{-1} E_k)$, for every $k \in \mathbb{N}$, we obtain

$$(4.1) \qquad \mu\left(\bigcup_{k \geq 1} E_k \right) \geq \mu\left(\bigcup_{k \geq 1} \sigma^{-1} E_k \right).$$

Observe that $(E_k)_k$ and $(\sigma^{-1} E_k)_k$ are partitions of Σ, in particular

$$\Sigma = \bigcup_{k \geq 1} E_k = \bigcup_{k \geq 1} \sigma^{-1} E_k.$$

This implies that (4.1) is an equality, therefore $\mu(E_k) = \mu(\sigma^{-1} E_k)$, for every $k \in \mathbb{N}$. In particular, we obtained that $\mu(D) = \mu(\sigma^{-1} D)$, as desired. $\qquad\square$

Remark 4.2. Recall that if (Σ, σ) is transitive, then the periodic measures are dense in $\mathcal{M}(\Sigma, \sigma)$ with respect to the weak* topology (see Theorem 3.8). It is a consequence of Lemma 3.17 that the same holds for the topology of convergence on cylinders. In other words, given an invariant probability measure μ, there exists a sequence $(\mu_n)_n$ of periodic measures such that $\lim_{n \to \infty} d(\mu, \mu_n) = 0$.

Remark 4.3. Recall that every element of $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is of the form $\lambda \mu$, where $\mu \in \mathcal{M}(\Sigma, \sigma)$ and $\lambda \in [0, 1]$.

In order to study the space $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ we will model it with a space of functions.

Definition 4.4. Denote by *FinCyl*(Σ) the collection of non-empty finite unions of cylinders in Σ. Let $M(\Sigma)$ be the space of functions $F : FinCyl(\Sigma) \to [0, 1]$ and $L(\Sigma) \subset M(\Sigma)$ the space of functions satisfying the following conditions:

  (a) If $C \subset C'$ are cylinders, then $F(C) \leq F(C')$.
  (b) $F(\bigcup_{k=1}^{n} C_k) = \sum_{k=1}^{n} F(C_k)$, for every (finite) collection of disjoint cylinders $(C_k)_k$.

Remark 4.5. Observe that *FinCyl*(Σ) is a countable set. Fix a bijection between *FinCyl*(Σ) and $\mathbb{N}$. This bijection allow us to identify $M(\Sigma)$ with $[0, 1]^{\mathbb{N}}$. From now on we consider $L(\Sigma)$ as a subset of $[0, 1]^{\mathbb{N}}$. We endow $M(\Sigma)$ with the product topology. Observe that $(R_n)_n \subset M(\Sigma)$ converges to $R \in M(\Sigma)$ if and only if

$$\lim_{n \to \infty} R_n(D) = R(D),$$

for every $D \in FinCyl(\Sigma)$.

Remark 4.6. Observe that the metric d, see equation (3.1), also defines a metric on $L(\Sigma)$. Indeed, if $F, G \in L(\Sigma)$ then

$$(4.2) \qquad d(F, G) = \sum_{n \geq 1} \frac{1}{2^n} |F(C_n) - G(C_n)|,$$

where $(C_n)_n$ is some enumeration of the cylinders on Σ, is a metric on $L(\Sigma)$. It is important to observe that the topology induced by d on $L(\Sigma)$ is compatible with the product topology on $M(\Sigma) = [0, 1]^{\mathbb{N}}$. Indeed, $\lim_{n \to \infty} d(F_n, F) = 0$, if and only if $\lim_{n \to \infty} F_n(C) = F(C)$, for every cylinder C. By condition (b) in the definition of $L(\Sigma)$ this is equivalent to $\lim_{n \to \infty} F_n(D) = F(D)$, for every $D \in FinCyl(\Sigma)$. In other words, there exists a continuous injective map from the set $L(\Sigma)$, endowed with the topology generated by the metric d, into the space $[0, 1]^{\mathbb{N}}$ endowed with the product topology.

Lemma 4.7. *The set $L(\Sigma)$ is compact with respect to the topology induced by d.*

Proof. We will first prove that $L(\Sigma)$ is a closed subset of $M(\Sigma)$. Let $(F_n)_n$ be a sequence of functions in $L(\Sigma)$ that converges to $F \in M(\Sigma)$. Let C and D be cylinders such that $C \subset D$. Then $F_n(C) \leq F_n(D)$, for every $n \in \mathbb{N}$. We conclude that

$$F(C) = \lim_{n \to \infty} F_n(C) \leq \lim_{n \to \infty} F_n(D) = F(D).$$

Similarly, if $(C_k)_{k=1}^m$ is a finite collection of disjoint cylinders we have that

$$F\left(\bigcup_{k=1}^m C_k\right) = \lim_{n\to\infty} F_n\left(\bigcup_{k=1}^m C_k\right) = \lim_{n\to\infty} \sum_{k=1}^m F_n(C_k) = \sum_{k=1}^m F(C_k).$$

We conclude that $F \in L(\Sigma)$. It follows that $L(\Sigma)$ is a closed subset of $M(\Sigma)$. Since $[0,1]^{\mathbb{N}}$ is compact, by virtue of Remark 4.6 we have that $L(\Sigma)$ is compact with the topology induced by d. $\square$

Remark 4.8. Observe that every sub-probability measure on Σ can be identified with a unique function $F \in L(\Sigma)$. More precisely, given $\mu \in \mathcal{M}_{\leq 1}(\Sigma, \sigma)$ we define $F_\mu \in L(\Sigma)$ by $F_\mu(D) := \mu(D)$, for every $D \in FinCyl(\Sigma)$. The map $\mu \mapsto F_\mu$ defines a continuous embedding $\mathcal{M}(\Sigma, \sigma) \hookrightarrow L(\Sigma)$, when we endow $\mathcal{M}(\Sigma, \sigma)$ with the topology of convergence on cylinders. We say that a sequence $(\mu_n)_n \subset \mathcal{M}(\Sigma, \sigma)$ converges to $F \in L(\Sigma)$ if $(F_{\mu_n})_n \subset L(\Sigma)$ converges to F.

In light of Remark 4.8, in order to prove that the space of invariant sub-probability measures is compact with respect to the cylinder topology, it suffices to prove that $\overline{\mathcal{M}(\Sigma, \sigma)} \subset L(\Sigma)$ consists of invariant sub-probability measures. At this point we will make a further assumption on the countable Markov shifts considered.

Definition 4.9. A countable Markov shift (Σ, σ) is said to satisfy the $\mathcal{F}$-**property** if for every element of the alphabet **i** and natural number n, there are only finitely many admissible words of length n starting and ending at **i**.

Remark 4.10. Every countable Markov shift (Σ, σ) of finite topological entropy and every locally compact countable Markov shift satisfies the $\mathcal{F}$-property. There also exist infinite entropy non-locally compact countable Markov shifts satisfying the $\mathcal{F}$-**property**. Indeed, let $(a_n)_n$ be a sequence of positive integers such that

$$\lim_{n\to\infty} \frac{1}{n} \log a_n = \infty.$$

Consider the countable Markov shift defined by a graph made of a_n simple loops of length n which are based at a common vertex and otherwise do not intersect. This system has the desired properties.

Proposition 4.11. *Let (Σ, σ) be a transitive countable Markov shift satisfying the $\mathcal{F}$-property. If $(\mu_n)_n$ is a sequence of periodic measures converging to a function $F \in L(\Sigma)$, then F extends to an invariant sub-probability measure.*

Proof. We start by proving that F extends to a measure. Fix a cylinder $C = [a_1, \ldots, a_m]$, and denote by Ck the cylinder $[a_1, \ldots, a_m, k]$. We will need the following lemma.

Lemma 4.12.

(4.3)
$$F(C) = \sum_{k \geq 1} F(Ck).$$

We assume that $F(C) > 0$, otherwise there is nothing to prove (both the left- and right-hand side would be zero). From now on assume that n is sufficiently large so that $\mu_n(C) > 0$. Let p_n be a periodic point associated to μ_n such that $p_n \in C$.

Proof of Lemma 4.12.   Observe that

$$F(C) - \sum_{s=1}^{k-1} F(Cs) = \lim_{n \to \infty} \left(\mu_n(C) - \sum_{s=1}^{k-1} \mu_n(Cs) \right) = \lim_{n \to \infty} \mu_n \left(\bigcup_{s \geq k} Cs \right),$$

therefore Lemma 4.12 is equivalent to proving that

$$\lim_{k \to \infty} \lim_{n \to \infty} \mu_n \left(\bigcup_{s \geq k} Cs \right) = 0.$$

We will argue by contradiction and assume that

$$\lim_{k \to \infty} \lim_{n \to \infty} \mu_n \left(\bigcup_{s \geq k} Cs \right) = A > 0.$$

Observe that $(\lim_{n \to \infty} \mu_n(\bigcup_{s \geq k} Cs))_k$ decreases as k goes to infinity. We obtain that

$$\lim_{n \to \infty} \mu_n(\bigcup_{s \geq k} Cs) \geq A,$$

for every $k \in \mathbb{N}$.

Recall that $C = [a_1, \ldots, a_m]$ and define the set $Q \subset \mathbb{N}$ by the following rule: $q \in Q$ if and only if $\mathbf{a_m q}$ is an admissible word. Define a function $p : Q \to \mathbb{Z}$ as follows: $p(i) = k$ if there exists an admissible word starting at i and ending at a_1 of length $k + 1$, but there is not any such word of length less than or equal to k. The map p is proper, in other words, $p^{-1}([a, b])$ is finite for every $a, b \in \mathbb{R}$. Indeed, assume by contradiction that $p^{-1}([a, b])$ is infinite; this would imply that $p^{-1}(c)$ is infinite for some $c \in \mathbb{N}$. For each $w \in p^{-1}(c)$ we have an admissible word of length $c + 1$ connecting w and a_1; this will create an admissible word (with length $m + c + 1$) of the form $\mathbf{a_1} \cdots \mathbf{a_m w} \cdots \mathbf{a_1}$. This contradicts the fact that (Σ, σ) satisfies the $\mathcal{F}$-property. We conclude that $p : Q \to \mathbb{Z}$ is proper.

Choose $k_0 \in \mathbb{N}$ such that $p(s) \geq \lfloor \frac{4}{A} \rfloor + 1$, for every $s \in Q$ satisfying $s \geq k_0$. Recall that p_n is a periodic point associated to the measure μ_n. We denote the minimal period of p_n by $m + t_n$, and let $[a_1 \cdots a_m b_1 \cdots b_{t_n}]$ be a neighborhood

of p_n. By the definition of μ_n we know that $\mu_n(Cs)$ is approximately the number of times that the word $\mathbf{a_1} \cdots \mathbf{a_m s}$ appears in $\mathbf{w_n} := \mathbf{a_1} \cdots \mathbf{a_m} \mathbf{b_1} \cdots \mathbf{b_{t_n}} \mathbf{a_1} \cdots \mathbf{a_m}$, divided by $m + t_n$. If $s \geq k_0$, then each block $\mathbf{a_1} \cdots \mathbf{a_m s}$ appearing in $\mathbf{w_n}$ is contained in a longer block of the form $\mathbf{a_1} \cdots \mathbf{a_m s r_1} \cdots \mathbf{r_B}$, where $B \geq \lfloor \frac{4}{A} \rfloor + 1$, and $r_t \neq a_1$, for all $t \in \{1, \ldots, B\}$. In particular, for $s \geq k_0$, each block of the form $\mathbf{a_1} \cdots \mathbf{a_m s}$ generates B letters that do not contribute to the number of blocks $\mathbf{a_1} \cdots \mathbf{a_m}$ in $\mathbf{w_n}$. Choose n_0 such that

$$\mu_n\left(\bigcup_{s \geq k_0} Cs \right) \geq \frac{A}{2},$$

for every $n \geq n_0$. This implies that the number of blocks of the form $\mathbf{a_1} \cdots \mathbf{a_m s}$, where $s \geq k_0$, in $\mathbf{w_n}$ is at least $(m+t_n)A/2$. As explained above, each of those blocks generates a disjoint block of length $(m+B+1)$. The number of letters used in those disjoint blocks add up to $(m+B+1)(m+t_n)A/2$. Observe that $(m+B+1)A/2 > 1$, which contradicts that the total number of letters is $(m + t_n)$. We conclude that $A = 0$. □

We will now use Kolmogorov's extension theorem to prove that F comes from a measure on Σ. To each $I_n := \{1, \ldots, n\} \subset \mathbb{N}$ we associate a measure on $\mathbb{N}^{I_n}$: this is the atomic measure ν_n that assigns to $[m_1, \ldots, m_n]$ the number $F([m_1, \ldots, m_n])$. We remark that if $\mathbf{m_1} \cdots \mathbf{m_n}$ is not an admissible word of Σ, then $F([m_1, \ldots, m_n]) = 0$. In order to use Kolmogorov's extension theorem and obtain a measure on $\mathbb{N}^{\mathbb{N}}$ we need to verify the consistency of the family $(\nu_n)_n$, in other words, that

$$\nu_n((m_1, \ldots, m_n)) = \nu_{n+1}((m_1, \ldots, m_n) \times \mathbb{N}).$$

By definition of the family $(\nu_n)_n$ this is equivalent to the formula

$$F(D) = \sum_{k \geq 1} F(Dk),$$

for $D = [m_1, \ldots, m_n]$. Lemma 4.12 implies the consistency of $(\nu_n)_n$. It follows from Kolmogorov's extension theorem that F extends to a measure μ on the full shift $\mathbb{N}^{\mathbb{N}}$. Observe that by definition of F the measure μ is supported on $\Sigma \subset \mathbb{N}^{\mathbb{N}}$. The invariance of μ follows from Lemma 4.1. □

Remark 4.13 (Limits are not always measures)**.** We now exhibit examples of countable Markov shifts that do not satisfy the $\mathcal{F}$-property for which sequences of measures converge to a function F that can not be extended to a measure. Let $\Sigma = \mathbb{N}^{\mathbb{N}}$ be the full shift. Consider the periodic point $p_n = \overline{1n}$, and denote by μ_n the periodic measure associated to p_n. Observe that $(\mu_n)_n$ converges

to $F \in L(\Sigma)$, where F is given by $F([1]) = 1/2$, and $F(C) = 0$, for any other cylinder $C \subsetneq \Sigma$. In this case it is clear that F does not come from a measure: use the decomposition $[1] = \bigcup_{s \geq 1}[1s]$, and the definition of F. Equivalently, the formula $F([1]) = \sum_{s \geq 1} F([1s])$ does not hold. In the full shift we can not expect to always have a measure as the limit of probability measures (in the topology of convergence on cylinders). Similar examples are easy to construct. For instance, consider the countable Markov shift defined by the matrix $M = (M_{ij})$, where $M_{1k} = 1 = M_{k1}$, for all $k \in \mathbb{N}$, and $M_{ij} = 0$ for the remaining entries. In this case the same choice of measures $(\mu_n)_n$ would provide a sequence of invariant probability measures that do not converge to a countably additive measure.

Proposition 4.14. *Let* (Σ, σ) *be a transitive countable Markov shift satisfying the $\mathcal{F}$-property. Then any sequence of invariant probability measures $(\mu_n)_n$ has a sub-sequence that converges on cylinders to an invariant sub-probability measure.*

Proof. Since $\mathcal{M}(\Sigma, \sigma) \subset L(\Sigma)$, by compactness of $L(\Sigma)$ there exists a sub-sequence $(\mu_{n_k})_k$ converging to a function $F \in L(\Sigma)$. Since the periodic measures are dense in $\mathcal{M}(\Sigma, \sigma)$ (see Remark 4.2) we can find a sequence of periodic measures $(\nu_k)_k$ such that $d(\mu_{n_k}, \nu_k) \leq \frac{1}{k}$. It follows that $\lim_{k \to \infty} d(\nu_k, F) = 0$. We can now use Proposition 4.11 and conclude that F corresponds to an invariant sub-probability measure. $\qquad\square$

Remark 4.15. The proof of Proposition 4.14 also implies that

$$\overline{\mathcal{M}(\Sigma, \sigma)} \subset \mathcal{M}_{\leq 1}(\Sigma, \sigma).$$

Indeed, if $F \in \overline{\mathcal{M}(\Sigma, \sigma)}$, then we have a sequence of invariant probability measures $(\mu_n)_n$ converging to F. As in the proof of Proposition 4.14 we conclude that F can be approximated by periodic measures, and therefore Proposition 4.11 implies the result.

As mentioned in the introduction, to prove that $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is affine homeomorphic to the Poulsen simplex we need to prove the existence of a sequence of invariant measures that converges on cylinders to the zero measure. In our next result we obtain such a property. We emphasize that if (Σ, σ) does not satisfy the $\mathcal{F}$-property, then this is not necessarily true (see Example 4.17).

Lemma 4.16. *Let* (Σ, σ) *be a transitive countable Markov shift satisfying the $\mathcal{F}$-property. Then there exists a sequence of invariant probability measures converging on cylinders to the zero measure.*

Proof. Fix some natural number k. We say that Property (k) holds if there exist arbitrarily long admissible words of the form $\mathbf{a_1} \cdots \mathbf{a_m}$, where $\{a_1, a_m\} \subset \{1, \ldots, k\}$, and $a_i \geq k + 1$, for all $i \in \{2, \ldots, m - 1\}$. If Property (k) holds we can construct a sequence of periodic measures $(\mu_n^{(k)})_n$ such that

$$\lim_{n \to \infty} \mu_n^{(k)} \left(\bigcup_{s=1}^{k} [s] \right) = 0.$$

First observe that there exists $M_0 = M_0(k)$ such that every two letters in $\{1, \ldots, k\}$ can be connected with an admissible word of length less than or equal to M_0. By hypothesis for every $n \in \mathbb{N}$ there exists an admissible word $\mathbf{w_n} = \mathbf{a_1^{(n)}} \cdots \mathbf{a_{m_n}^{(n)}}$, where $\{a_1^{(n)}, a_{m_n}^{(n)}\} \subset \{1, \ldots, k\}$, and $a_i^{(n)} \geq k + 1$, for all $i \in \{2, \ldots, m_n - 1\}$, and $m_n \geq n$. We can extend the word $\mathbf{w_n}$ into an admissible word

$$\mathbf{w_n'} = \mathbf{a_1^{(n)}} \cdots \mathbf{a_{m_n}^{(n)}} \mathbf{b_1^{(n)}} \cdots \mathbf{b_{s_n}^{(n)}} \mathbf{a_1^{(n)}}, \quad \text{where } s_n \leq M_0.$$

The word $\mathbf{w_n'}$ can be used to define a periodic orbit, and therefore a periodic measure, say $\mu_n^{(k)}$, on Σ. Observe that

$$\mu_n^{(k)} \left(\bigcup_{s=1}^{k} [s] \right) \leq \frac{s_n + 2}{s_n + m_n} \leq \frac{M_0 + 2}{n},$$

which readily implies that $\lim_{n \to \infty} \mu_n^{(k)} (\bigcup_{s=1}^{k} [s]) = 0$.

We will now verify that under the hypothesis of Lemma 4.16, Property (k) holds. Assume by contradiction that this is not possible, in other words that any such word has length less than or equal to N_0. Define

$$T : \{n \in \mathbb{N} : n \geq k + 1\} \to \mathbb{N},$$

in the following way: $T(n) = r$, if there exists an admissible word of length r with first letter in $\{1, \ldots, k\}$ and ending at n, but there is no such admissible word of length strictly less than r. Similarly define

$$S : \{n \in \mathbb{N} : n \geq k + 1\} \to \mathbb{N},$$

in the following way: $S(n) = r$, if there exists an admissible word of length $r + 1$ with first letter n and ending at some letter in $\{1, \ldots, k\}$, but there is no such admissible word of length less than or equal to r. By definition of T and S we know that given $n \geq k + 1$, there exists an admissible word

$$\mathbf{y_n} := \mathbf{c_1^{(n)}} \cdots \mathbf{c_{T(n)-1}^{(n)}} \mathbf{n} \mathbf{d_1^{(n)}} \cdots \mathbf{d_{S(n)}^{(n)}},$$

where $\{c_1^{(n)}, d_{S(n)}^{(n)}\} \subset \{1, \dots, k\}$ and the rest of the letters are strictly larger than k. Observe that by assumption we have $T(n) + S(n) \le N_0$, for every $n \ge k + 1$. For $n \ge k + 1$ define $W(n)$ as the biggest letter in the word $\mathbf{y_n}$. We can inductively choose a sequence $(n_t)_t$ such that $W(n_t) < n_{t+1}$, and observe that $(\mathbf{y_{n_t}})_t$ are pairwise distinct. As with the words $(\mathbf{w_n})_n$, we can extend each $\mathbf{y_n}$ to an admissible word $\mathbf{y_n'} := \mathbf{e_1^{(n)}} \cdots \mathbf{e_{s_n}^{(n)}} \mathbf{y_n} \mathbf{f_1^{(n)}} \cdots \mathbf{f_{r_n}^{(n)}}$, where s_n and r_n are less than M_0, and $e_1^{(n)} = 1 = f_{r_n}^{(n)}$. The word $\mathbf{y_n'}$ defines a periodic point of period $\le 2M_0 + N_0$. Since $(\mathbf{y_{n_t}})_t$ are pairwise distinct we found infinitely many periodic points of periods less than or equal to $2M_0 + N_0$ (starting and ending at 1), which contradicts that (Σ, σ) satisfies the $\mathcal{F}$-property. We conclude that Property (k) holds for every $k \in \mathbb{N}$.

For every $k \in \mathbb{N}$ we obtain a sequence of periodic measures $(\mu_n^{(k)})_n$ such that $\lim_{n \to \infty} \mu_n^{(k)}(\bigcup_{s=1}^k [s]) = 0$. Let n_k be such that $\mu_{n_k}^{(k)}(\bigcup_{s=1}^k [s]) \le \frac{1}{k}$. To simplify notation we define $\nu_k := \mu_{n_k}^{(k)}$. We claim that $(\nu_k)_k$ converges on cylinders to the zero measure. Observe that for $k \ge m$, we have

$$\nu_k([m]) \le \nu_k \left(\bigcup_{s=1}^m [s] \right) \le \nu_k \left(\bigcup_{s=1}^k [s] \right) \le \frac{1}{k}.$$

We conclude that $\lim_{k \to \infty} \nu_k([m]) = 0$. Since $m \in \mathbb{N}$ was arbitrary, and every cylinder C is contained in a cylinder of length one, we conclude that $(\nu_k)_k$ converges on cylinders to the zero measure. $\qquad\square$

Example 4.17. We exhibit an example of a countable Markov shift of infinite entropy not satisfying the $\mathcal{F}$-property, for which there is no sequence of measures converging to zero in the cylinder topology. Let (Σ, σ) be the countable Markov shift defined by the graph formed by infinitely many loops of length two rooted at a common vertex. That is, the allowed transitions are of the form $1 \to N$ and $N \to 1$ for every $N \in \mathbb{N}$; this example was also considered in Remark 4.13. The system has infinite entropy, since it has infinitely many periodic orbits of period two intersecting [1]. The frequency of the digit 1 is at least $1/2$ for every element of Σ. Therefore, if μ is an ergodic measure then $\mu([1]) \ge 1/2$. Thus, for any sequence of invariant measures $(\mu_n)_n$ we must have that $\liminf_{n \to \infty} \mu_n([1]) \ge 1/2$. In particular, the sequence $(\mu_n)_n$ does not converge to zero in the cylinder topology.

We can now prove the compactness of the space of sub-probability measures. As explained in the introduction a compactification of the space of invariant probability measures is important for applications (for instance, see [ITV] and [V]).

Theorem 4.18. *If (Σ, σ) is a transitive countable Markov shift satisfying the $\mathcal{F}$-property, then the space $\mathcal{M}_{\le 1}(\Sigma, \sigma)$ is compact with respect to the topology of convergence on cylinders.*

Proof. It is a consequence of Remark 4.15 that $\overline{\mathcal{M}(\Sigma, \sigma)} \subset \mathcal{M}_{\leq 1}(\Sigma, \sigma)$. It is enough to prove that $\overline{\mathcal{M}(\Sigma, \sigma)} = \mathcal{M}_{\leq 1}(\Sigma, \sigma)$. Let $(\mu_n)_n$ be a sequence of invariant probability measures converging on cylinders to the zero measure (see Lemma 4.16). An element in $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ has the form $\lambda\mu$, where μ is an invariant probability measure and $\lambda \in [0, 1]$. Define $\nu_n = \lambda\mu + (1 - \lambda)\mu_n$. Observe that $(\nu_n)_n$ converges on cylinders to $\lambda\mu$. This concludes that $\overline{\mathcal{M}(\Sigma, \sigma)} = \mathcal{M}_{\leq 1}(\Sigma, \sigma)$, and therefore $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is compact. $\qquad\square$

The idea behind Remark 4.13 can be used to prove that Theorem 4.18 is sharp. We will prove that without the $\mathcal{F}$-property it is possible to construct a sequence of invariant measures that converges on cylinders to a finitely additive measure that is not countably additive. In particular, Theorem 4.18 is false without the $\mathcal{F}$-property assumption.

Proposition 4.19. *Suppose that (Σ, σ) does not satisfy the $\mathcal{F}$-property. Then there exists a sequence of periodic measures that converges on cylinders to $F \in L(\Sigma)$, where F can not be extended to a measure.*

Proof. Since (Σ, σ) does not satisfy the $\mathcal{F}$-property there exists a symbol **i** and natural number n such that there are infinitely many admissible words of length n that start and end at **i**. The set of admissible words of length $k + 1$ starting and ending at **i**, where the symbol **i** only appears at the beginning and at the end of the word, is denoted by A_k. By hypothesis there exists $q \leq n$ such that $|A_q| = \infty$. Set $A_q = \{w_k : k \in \mathbb{N}\}$. Observe that each $w_k \in A_q$ defines a periodic measure that we denote by μ_k. Maybe after passing to a sub-sequence we can assume that $(\mu_k)_k$ converges on cylinders to $F \in L(\Sigma)$. By construction we know that $\mu_k([\mathbf{i}]) = \frac{1}{q}$. Observe that $\mu_k([\mathbf{ir}])$ is equal to 0 or $\frac{1}{q}$, for every k and **r**. If $\lim_{k \to \infty} \mu_k([\mathbf{ir}]) = 0$, for every **r**, then F can not come from a measure: $\sum_{r \geq 1} F([\mathbf{ir}]) = 0$, but $F([\mathbf{i}]) = \frac{1}{q}$. We assume there exists $\mathbf{r_1}$ such that $\lim_{k \to \infty} \mu_k([\mathbf{ir_1}]) = \frac{1}{q}$, which is equivalent to saying that $\mu_k([\mathbf{ir_1}]) = \frac{1}{q}$, for every k sufficiently large. We can repeat the process and conclude that if $\lim_{k \to \infty} \mu_k([\mathbf{ir_1 s}]) = 0$, for every **s**, then F does not come from a measure. We can assume that there exists $\mathbf{r_2}$ such that $\mu_k([\mathbf{ir_1 r_2}]) = \frac{1}{q}$, for k sufficiently large. By repeating this process we obtain that F does not come from a measure or that $\mu_k([\mathbf{ir_1 \cdots r_{q-1}}]) = \frac{1}{q}$, for k sufficiently large. This last condition is equivalent to saying that the sequence $(\mu_k)_k$ stabilizes, which contradicts that the measures are pairwise different. $\qquad\square$

In Section 6 we will be interested in countable Markov shifts that do not necessarily have the $\mathcal{F}$-property. Despite Proposition 4.19 we can regain control by imposing an integrability condition on the sequence of probability measures (see Proposition 6.12). This integrability condition will rule out the sequence constructed in Proposition 4.19.

5   The Poulsen simplex

We now prove one of our main results, in which we characterize the spaces $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ and $\mathcal{M}(\Sigma, \sigma)$ for countable Markov shifts satisfying the $\mathcal{F}$-property.

Theorem 5.1. *Let (Σ, σ) be a transitive countable Markov shift satisfying the $\mathcal{F}$-property. Then $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is affine homeomorphic to the Poulsen simplex.*

Proof. An element in $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ has the form $\lambda\mu$, where μ is an invariant probability measure and $\lambda \in [0, 1]$. To prove that $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is the Poulsen simplex it is enough to prove that the extreme points of $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ are dense (we already know that $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is a metrizable convex compact Choquet simplex). We will approximate the measure $\lambda\mu$ with periodic measures. As explained in the proof of Theorem 4.18, we can construct invariant probability measures $(\nu_n)_n$ such that $\lim_{n\to\infty} d(\nu_n, \lambda\mu) = 0$. By Remark 4.2 we can find a sequence $(\overline{\nu}_n)_n$ of periodic measures such that $d(\overline{\nu}_n, \nu_n) \leq \frac{1}{n}$. This implies that $(\overline{\nu}_n)_n$ converges on cylinders to $\lambda\mu$, and therefore, by the main result of [LOS], $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ is affine homeomorphic to the Poulsen simplex.   □

Since the extreme points of $\mathcal{M}_{\leq 1}(\Sigma, \sigma)$ are the ergodic probability measures together with the zero measures, it follows directly from Theorem 5.1 and Lemma 3.17 that

Theorem 5.2. *Let (Σ, σ) be a transitive countable Markov shift satisfying the $\mathcal{F}$-property. Then $\mathcal{M}(\Sigma, \sigma)$ is affinely homeomorphic to the Poulsen simplex minus a vertex and all of its convex combinations.*

Corollary 5.3. *Let (Σ, σ) be a transitive countable Markov shift satisfying the $\mathcal{F}$-property. Then the set $\mathcal{M}_e(\Sigma, \sigma)$ is path connected.*

Proof. The set of extreme points of the Poulsen simplex is path connected [LOS, (4) p. 101]. It follows from Theorem 5.1 that the set $\mathcal{M}_e(\Sigma, \sigma) \cup \{0_m\}$, where 0_m denotes the zero measure, is path connected. Denote by $Q := [-1, 1]^{\mathbb{N}}$ the Hilbert cube and let

$$P := \{(x_1, x_2, \dots) \in Q : |x_n| < 1, \text{ for every } n \in \mathbb{N}\}.$$

It was proved in [LOS, Theorem 3.1] that there exists a homeomorphism h between the Hilbert cube Q and the Poulsen simplex which maps P onto the set of extreme points of the Poulsen simplex. Denote $z = h^{-1}(0_m) \in P$. For any $x, y \in P \setminus \{z\}$ it is clear that there exists a continuous path $p : [0, 1] \to P$ such that $p(0) = x, p(1) = y$ and $p(t) \neq z$ for every $t \in [0, 1]$. Therefore the set $\mathcal{M}_e(\Sigma, \sigma)$ is path connected.   □

6   The space of invariant measures for suspension flows

In this section we study the space of invariant probability measures of a suspension flow defined over a countable Markov shift.

6.1   Suspension flows.   Let (Σ, σ) be a countable Markov shift and $\tau : \Sigma \to \mathbb{R}$ a continuous positive function bounded away from zero, that is, there exists $c = c(\tau) > 0$ such that $\tau(x) \geq c$, for all $x \in \Sigma$. Consider the space

$$Y = \{(x, t) \in \Sigma \times \mathbb{R} : 0 \leq t \leq \tau(x)\},$$

with the points $(x, \tau(x))$ and $(\sigma(x), 0)$ identified for each $x \in \Sigma$. The **suspension flow** over σ with **roof function** τ is the semi-flow $\Theta = (\theta_t)_{t \geq 0}$ on Y defined by

$$\theta_t(x, s) = (x, s + t), \quad \text{whenever } s + t \in [0, \tau(x)].$$

In particular, $\theta_{\tau(x)}(x, 0) = (\sigma(x), 0)$. The space of invariant probability measures for the shift is related to the space of invariant probability measures for the flow, that we denote by $\mathcal{M}(\Sigma, \sigma, \tau)$. Indeed, it follows from a classical result of Ambrose and Kakutani [AK] (see [PP, Chapter 6] for details) that

Lemma 6.1. *Let* (Y, Θ) *be a suspension flow over* (Σ, σ) *with roof function* τ *bounded away from zero. Let*

$$\mathcal{M}_\tau := \left\{ \mu \in \mathcal{M}(\Sigma, \sigma) : \int \tau d\mu < \infty \right\}.$$

The map $\varphi : \mathcal{M}_\tau \to \mathcal{M}(\Sigma, \sigma, \tau)$ *defined by*

$$\mu \mapsto \frac{\mu \times Leb}{\int \tau d\mu},$$

where Leb is the one-dimensional Lebesgue measure, is a bijection.

We denote the inverse of φ by ψ. We will be particularly interested in a special class of roof functions.

Definition 6.2. A positive function $\tau : \Sigma \to \mathbb{R}$ belongs to the class $\mathcal{R}$ if the following properties hold:
  (a) τ is uniformly continuous, bounded away from zero, and $\text{var}_2(\tau)$ is finite,
  (b)
$$\lim_{k \to \infty} \inf_{x : x_1 \geq k} \tau(x) = \infty,$$
  where x_1 is the first coordinate of x.

Remark 6.3. The class $\mathcal{R}$ includes a wealth of interesting examples. For instance, the geodesic flow over the modular surface can be coded as a suspension flow over the full-shift on a countable alphabet, $\Sigma = \mathbb{N}^{\mathbb{N}}$, with a roof function τ belonging to $\mathcal{R}$; see [GK, KU] for details. A large class of examples belonging to the class $\mathcal{R}$ is to be found in one-dimensional dynamics. Indeed, the class of Expanding–Markov–Renyi (EMR) maps is a class of interval maps with infinitely many branches which was introduced by Pollicott and Weiss in [PW] and has been extensively studied. It turns out that if f is an EMR map, then the symbolic version of the corresponding geometric potential $\log |f'|$ belongs to $\mathcal{R}$. These potentials carry the relevant fractal information of the system as well as the coding of relevant equilibrium measures such as Sinai–Ruelle–Bowen measures. An example of an EMR map is the Gauss map.

6.2   The topology of convergence on cylinders.

The space of invariant sub-probability measures of the suspension flow is denoted by $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$. In this section we endow this space with a topology that makes it compact whenever $\tau \in \mathcal{R}$ (see Theorem 6.16). The topology we consider is an adaptation of the cylinder topology defined in sub-section 3.3. Let (Y, Θ) be a suspension flow over (Σ, σ) with roof function τ and $c = c(\tau) > 0$ such that $\inf \tau > c$.

Definition 6.4. Let $(\nu_n)_n$ and ν be measures in $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$. We say that $(\nu_n)_n$ **converges on cylinders** to ν if

$$\lim_{n \to \infty} \nu_n(C \times [0, c]) = \nu(C \times [0, c]),$$

for every cylinder $C \subset \Sigma$.

Recall that by Kac's formula we know that

$$\nu(C \times [0, c]) = \frac{c\mu(C)}{\int \tau d\mu},$$

whenever $\nu \in \mathcal{M}(\Sigma, \sigma, \tau)$ and $\mu = \psi(\nu)$.

Remark 6.5. Let ν and $(\nu_n)_n$ be invariant probability measures for the suspension flow and set $\mu_n = \psi(\nu_n)$, and $\mu = \psi(\nu)$. From the definition of ψ we have that the following statements are equivalent:

(a) The sequence $(\nu_n)_n$ converges on cylinders to $\lambda \nu$, where $\lambda \in [0, 1]$.

(b) The following limit holds:

$$\lim_{n \to \infty} \frac{\mu_n(C)}{\int \tau d\mu_n} = \lambda \frac{\mu(C)}{\int \tau d\mu},$$

for every cylinder $C \subset \Sigma$.

Moreover, $\lambda_1 \nu_1$ and $\lambda_2 \nu_2$ are equal if and only if

$$\lambda_1 \nu_1(C \times [a, b]) = \lambda_2 \nu_2(C \times [a, b]),$$

for every cylinder $C \subset \Sigma$ and $a, b \in \mathbb{R}$. By Kac's formula this is equivalent to

$$\lambda_1 \frac{\mu_1(C)}{\int \tau d\mu_1} = \lambda_2 \frac{\mu_2(C)}{\int \tau d\mu_2},$$

for every cylinder $C \subset \Sigma$.

Lemma 6.6. *The topology of the convergence on cylinders in $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$ is metrizable.*

Proof. Let $\rho : \mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau) \times \mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau) \to \mathbb{R}$ be defined by

$$\rho(\nu_1, \nu_2) = \sum_{k \geq 1} \frac{1}{2^k} |\nu_1(C_i \times [0, c]) - \nu_2(C_i \times [0, c])|,$$

where $(C_i)_i$ is some enumeration of the cylinders of Σ and $c = c(\tau)$. The map ρ is a metric. Indeed, let $\nu_1 = \lambda_1 \varphi(\mu_1)$ and $\nu_2 = \lambda_2 \varphi(\mu_2)$ be in $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$, where λ_1 and λ_2 are in $[0, 1]$. Suppose that $\rho(\nu_1, \nu_2) = 0$. By Remark 6.5 we know that

$$\lambda_1 \frac{\mu_1(C)}{\int \tau d\mu_1} = \lambda_2 \frac{\mu_2(C)}{\int \tau d\mu_2},$$

for every cylinder $C \subset \Sigma$. If $\lambda_1 = 0$ we necessarily have $\lambda_2 = 0$: in this case ν_1 and ν_2 are both the zero measure. Assume that $\lambda_1 \neq 0$; then

$$\mu_1(C) = \frac{\lambda_2}{\lambda_1} \frac{\int \tau d\mu_1}{\int \tau d\mu_2} \mu_2(C),$$

for every cylinder $C \subset \Sigma$. By the outer regularity of Borel measures on a metric space we conclude that $\mu_1 = A\mu_2$, where $A = \frac{\lambda_2}{\lambda_1} \frac{\int \tau d\mu_1}{\int \tau d\mu_2}$. This immediately implies that

$$\lambda_1 \frac{\mu_1}{\int \tau d\mu_1} = \lambda_2 \frac{\mu_2}{\int \tau d\mu_2},$$

and therefore $\nu_1 = \nu_2$. The other properties of a metric are easily verified. Note that $(\nu_n)_n$ converges on cylinders to ν if and only if $\lim_{n \to \infty} \rho(\nu_n, \nu) = 0$, that is, the topology of convergence on cylinders is metrizable. $\square$

Our next result should be compared with Lemma 3.17. It says that the topology of convergence on cylinders coincides with the weak* topology on $\mathcal{M}(\Sigma, \sigma, \tau)$. We emphasize that this result holds for every τ which is bounded below.

Lemma 6.7. *Let ν and $(\nu_n)_n$ be measures in $\mathcal{M}(\Sigma, \sigma, \tau)$. The following assertions are equivalent:*

(a) *The sequence $(\nu_n)_n$ converges in the weak* topology to ν.*
(b) *The sequence $(\nu_n)_n$ converges on cylinders to ν.*

Proof. Let $C \subset \Sigma$ be a cylinder. Observe that

$$\partial(C \times [0, c]) = C \times \{0, c\}.$$

Since ν is a flow invariant probability measure we know that $\nu(C \times \{x\}) = 0$, for each $x \in \mathbb{R}$. We conclude that $\nu(\partial(C \times [0, c])) = 0$. Finally, use Proposition 3.3 (d) to conclude that (a) implies (b).

We will now prove that (b) implies (a). A base for the topology in Y is given by

$$\Omega := \{C \times (a, b) \subset Y : C \text{ cylinder for } \Sigma \text{ and } a, b \in \mathbb{Q} \text{ with } a < b\}.$$

It follows from the flow invariance of the measures that for every set $C \times (a, b) \subset Y$ we have

$$\lim_{n \to \infty} \nu_n(C \times (a, b)) = \nu(C \times (a, b)).$$

Observe that a finite intersection of elements in Ω is still in Ω. Note that each open set $\mathcal{O} \subset Y$ can be written as a countable union of elements in Ω, say $\mathcal{O} = \bigcup_{k \geq 1}(C_k \times (a_k, b_k))$. The result now follows from [Bi, Theorem 2.2]. $\qquad\square$

One of the main properties of the class $\mathcal{R}$ is that we can rule out the escape of mass by imposing a uniform bound on the integral of τ (see Lemma 6.8). This illustrates the importance of part (b) in Definition 6.2.

Lemma 6.8. *Let $(\mu_n)_n \subset \mathcal{M}(\Sigma, \sigma)$ and $\mu \in \mathcal{M}_{\leq 1}(\Sigma, \sigma)$ be such that $(\mu_n)_n$ converges on cylinders to the measure μ. Let $\tau \in \mathcal{R}$ and assume there exists $M \in \mathbb{R}$ such that $\int \tau d\mu_n \leq M$, for every $n \in \mathbb{N}$. Then μ is a probability measure. Moreover, $(\mu_n)_n$ converges to μ in the weak* topology.*

Proof. Observe that for every $k \in \mathbb{N}$ we have

$$\left(\inf_{x:x_0 \geq k} \tau(x) \right) \mu_n \left(\bigcup_{s \geq k} [s] \right) \leq \int \tau d\mu_n \leq M;$$

then

$$\mu_n \left(\bigcup_{s \geq k} [s] \right) \leq \frac{M}{\inf_{x:x_0 \geq k} \tau(x)}.$$

This is equivalent to $\mu_n(\bigcup_{s<k}[s]) \geq 1 - \frac{M}{\inf_{x:x_0 \geq k} \tau(x)}$. By definition of the convergence on cylinders we have

$$\mu \left(\bigcup_{s<k} [s] \right) = \lim_{n \to \infty} \mu_n \left(\bigcup_{s<k} [s] \right) \geq 1 - \frac{M}{\inf_{x:x_0 \geq k} \tau(x)}.$$

Since $\tau \in \mathcal{R}$ we can conclude that

$$\lim_{k \to \infty} \mu \left(\bigcup_{s < k} [s] \right) = 1,$$

and therefore μ is a probability measure. Since the sequence $(\mu_n)_n$ converges on cylinders to a probability measure we conclude that $(\mu_n)_n$ converges in the weak* topology (see Lemma 3.17).   □

Our next two lemmas completely describe the topology of convergence on cylinders in terms of convergence of measures in Σ.

Lemma 6.9. *Let* $\tau \in \mathcal{R}$. *A sequence* $(\nu_n)_n \subset \mathcal{M}(\Sigma, \sigma, \tau)$ *converges on cylinders to the zero measure if and only if*

$$\lim_{n \to \infty} \int \tau d\psi(\nu_n) = \infty.$$

Proof. To simplify notation define $\mu_n = \psi(\nu_n)$. We will first prove that

(6.1) $$\lim_{n \to \infty} \int \tau d\mu_n = \infty$$

implies that $(\nu_n)_n$ converges on cylinders to the zero measure. Note that equation (6.1) implies that

$$\lim_{n \to \infty} \frac{\mu_n(C)}{\int \tau d\mu_n} = 0,$$

for every cylinder C. In virtue of Remark 6.5 we get that $(\nu_n)_n$ converges on cylinders to the zero measure. To prove the other implication we argue by contradiction: suppose that the sequence $(\nu_n)_n$ converges to the zero measure and that there exists a sub-sequence $(n_k)_k$ such that $\int \tau d\mu_{n_k} \le M$, for some $M \in \mathbb{R}$. From Remark 6.5 we obtain

$$0 = \lim_{k \to \infty} \frac{\mu_{n_k}(C)}{\int \tau d\mu_{n_k}} \ge \frac{1}{M} \limsup_{k \to \infty} \mu_{n_k}(C).$$

In particular, $(\mu_{n_k})_k$ converges on cylinders to the zero measure. Lemma 6.8 implies that $\lim_{k \to \infty} \int \tau d\mu_{n_k} = \infty$, which contradicts the choice of the sequence $(n_k)_k$.   □

Lemma 6.10. *Let* $\tau \in \mathcal{R}$ *and* $(\nu_n)_n$, ν *be invariant probability measures for the suspension flow. Define* $\mu_n = \psi(\nu_n)$ *and* $\mu = \psi(\nu)$. *Then the following are equivalent:*

(a) *The sequence* $(\nu_n)_n$ *converges on cylinders to* $\lambda \nu$, *where* $\nu \in \mathcal{M}(\Sigma, \sigma, \tau)$ *and* $\lambda \in (0, 1]$.

(b) *The sequence* $(\mu_n)_n$ *converges to* μ *in the weak* topology and*

$$\lim_{n \to \infty} \frac{\int \tau d\mu}{\int \tau d\mu_n} = \lambda \in (0, 1].$$

Proof. We first prove that (b) implies (a). If $(\mu_n)_n$ converges in the weak*
topology to μ, then $\lim_{n\to\infty} \mu_n(C) = \mu(C)$, for every cylinder C. It follows from
the hypothesis on λ that

$$\lim_{n\to\infty} \frac{\mu_n(C)}{\int \tau d\mu_n} = \lambda \frac{\mu(C)}{\int \tau d\mu}.$$

Remark 6.5 implies that $(\nu_n)_n$ converges on cylinders to $\lambda\nu$.

Now suppose that $(\nu_n)_n$ converges on cylinders to $\lambda\nu$. It follows from Lemma
6.9 and the assumption that $\lambda > 0$ that $(\int \tau d\mu_n)_n$ is a bounded sequence. After
passing to a sub-sequence we can assume that $(\int \tau d\mu_n)_n$ is convergent. Let $L \in \mathbb{R}$
be such that $\lim_{n\to\infty} \int \tau d\mu_n = L$. Remark 6.5 implies that

$$\lim_{n\to\infty} \mu_n(C) = \frac{\lambda L}{\int \tau d\mu} \mu(C),$$

for every cylinder C. We conclude that $(\mu_n)_n$ converges on cylinders to

$$\mu_0 := \frac{\lambda L}{\int \tau d\mu} \mu.$$

Observe that for sufficiently large n we have $\int \tau d\mu_n \leq (L+1)$. Lemma 6.8 implies
that μ_0 is a probability measure and that $(\mu_n)_n$ converges in the weak* topology
for μ_0. Since μ is a probability measure we conclude that $\lambda L = \int \tau d\mu$, and
therefore $\mu_0 = \mu$. This argument shows that every sub-sequence of our initial
sequence $(\mu_n)_n$ has a sub-sub-sequence converging to μ. This readily implies that
the whole sequence converges to μ, and that $\lim_{n\to\infty} \frac{\int \tau d\mu}{\int \tau d\mu_n} = \lambda$. □

Our next result should be compared with Lemma 4.16. As mentioned in the
introduction, this is a necessary ingredient to prove that $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$ is the Poulsen
simplex.

Lemma 6.11. *If $\tau \in \mathcal{R}$, then there exists a sequence of periodic measures
$(\nu_n)_n \subset \mathcal{M}(\Sigma, \sigma, \tau)$ that converges on cylinders to the zero measure.*

Proof. We will separate our analysis into two cases.

Case 1 (assume (Σ, σ) satisfies the $\mathcal{F}$-property). By Lemma 4.16 there
exists a sequence of periodic measures $(\mu_n)_n$ which converges on cylinders to the
zero measure. Observe that every periodic measure belongs to $\mathcal{M}_\tau$, in particular
$\varphi(\mu_n) \in \mathcal{M}(\Sigma, \sigma, \tau)$. Now, by Lemma 6.8 we conclude that $\lim_{n\to\infty} \int \tau d\mu_n = \infty$. It
follows from Lemma 6.9 that the sequence $(\varphi(\mu_n))_n$ converges to the zero measure.

Case 2 (assume (Σ, σ) does not satisfy the $\mathcal{F}$-property). In this case
there exists an element a in the alphabet, $l \in \mathbb{N}$, and a sequence $(p_n)_n$ of distinct
periodic points of length l such that $p_n \in [a]$. The periodic measure associated

to p_n is denoted by η_n. Since $\tau \in \mathcal{R}$, for $N \in \mathbb{R}$ there exists n_N such that for $n \geq n_N$ we have $S_l \tau(p_n) > N$. Here $S_l \tau$ is the Birkhoff sum of τ of length l. In particular, $\int \tau d\eta_n = \frac{1}{l} S_l \tau(p_n) \geq N/l$. This implies that $\lim_{n \to \infty} \int \tau d\eta_n = \infty$. The result then follows from Lemma 6.9. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Recall that the set $L(\Sigma)$ was introduced in Definition 4.4. We will now prove a compactness result similar to Proposition 4.11. The proof of Proposition 6.12 is significantly simpler than the one of Proposition 4.11; it would be interesting if this result can be generalized to a larger class of potentials.

Proposition 6.12. *Assume that $\tau \in \mathcal{R}$. Let $(\mu_n)_n$ be a sequence of periodic measures on Σ. Suppose that $(\mu_n)_n$ converges on cylinders to $F \in L(\Sigma)$, and that there exists $M \in \mathbb{R}$ such that $\int \tau d\mu_n \leq M$, for all $n \in \mathbb{N}$. Then F extends to an invariant probability measure.*

Proof. We will follow the strategy of the proof of Proposition 4.11. It is enough to prove that

$$\lim_{k \to \infty} \lim_{n \to \infty} \mu_n \left(\bigcup_{s \geq k} Cs \right) = 0,$$

where $C = [a_0 \cdots a_{m-1}]$ is a cylinder. Let $p_n \in \Sigma$ be a periodic point of period r_n and μ_n be the periodic measure associated to the point p_n. It is important to observe that

$$(6.2) \qquad \mu_n \left(\bigcup_{s \geq k} [s] \right) \geq \mu_n \left(\bigcup_{s \geq k} Cs \right).$$

Indeed, the probability measure μ_n is equidistributed on the set

$$\{p_n, \sigma(p_n), \ldots, \sigma^{r_n - 1}(p_n)\}.$$

Observe that $\sigma^k(p_n) \in Cs$ implies that $\sigma^{k+m}(p_n) \in [s]$, from which inequality (6.2) follows. Since $\int \tau d\mu_n \leq M$, we obtain that

$$\mu_n \left(\bigcup_{s \geq k} [s] \right) \leq \frac{\int \tau d\mu_n}{\inf_{x:x_1 \geq k} \tau(x)} \leq \frac{M}{\inf_{x:x_1 \geq k} \tau(x)},$$

but this immediately implies that

$$\lim_{k \to \infty} \lim_{n \to \infty} \mu_n \left(\bigcup_{s \geq k} [s] \right) \leq \limsup_{k \to \infty} \frac{M}{\inf_{x:x_1 \geq k} \tau(x)}.$$

Since $\tau \in \mathcal{R}$ we obtain that the right hand side in the last inequality is zero. Finally use inequality (6.2) to conclude that

$$\lim_{k \to \infty} \lim_{n \to \infty} \mu_n \left(\bigcup_{s \geq k} Cs \right) = 0.$$

As in the proof of Proposition 4.11 we have that F extends to a measure on Σ. Lemma 6.8 implies that F is a probability measure. The invariance follows from Lemma 4.1. ∎

Proposition 6.12 states that, assuming an integrability condition, limits of periodic measures are invariant probability measures. It is then of particular importance to know whether it is possible to approximate a sequence of invariant measures $(\mu_n)_n$ by periodic measures such that the assumption $\sup_n \int \tau d\mu_n < \infty$ still remains true for the sequence of periodic measures. Proposition 6.14 addresses this question. In the proof of Proposition 6.14 we will need to approximate a measure in $\mathcal{M}(\Sigma, \sigma)$ by a convex combination of finitely many ergodic probability measures. This result is classical in the compact case: the space of invariant probability measures is a compact convex set, and therefore the result follows from the Krein–Milman theorem. With lack of a good reference we provide a proof of this result that avoids the Krein–Milman theorem.

Lemma 6.13. *Let $\mu \in \mathcal{M}(\Sigma, \sigma)$ and $(f_i)_{i=0}^n$ be real-valued functions in $L^1(\mu)$. Given $\varepsilon > 0$, there exists $\mu_1 \in \mathcal{M}(\Sigma, \sigma)$ that is a convex combination of finitely many ergodic probability measures which satisfies*

$$\left| \int f_i d\mu - \int f_i d\mu_1 \right| \leq \varepsilon,$$

for every $i \in \{0, \ldots, n\}$.

Proof. It is enough to prove the result under the assumption that each f_i is non-negative. Indeed, let f_i^+ and f_i^- be the positive and negative parts of f_i. Applying the result to the set $(f_i^+)_i \bigcup (f_i^-)_i$ and $\varepsilon/2$, we obtain a measure μ_1. By the triangle inequality, the measure μ_1 verifies the inequalities in the statement of Lemma 6.13.

From now on we assume that each f_i is non-negative. Let $\mu = \int \mu_x dm(x)$ be the ergodic decomposition of μ, in particular,

$$\int f_k d\mu = \int \left(\int f_k d\mu_x \right) dm(x),$$

for $k \in \{0, \ldots, n\}$. By ergodic decomposition, the measure μ_x is ergodic for μ-almost every $x \in \Sigma$. We choose a measurable set $S \subset \Sigma$ such that $\mu(S) = 1$, μ_x is ergodic for every $x \in S$, and f_k is μ_x-integrable for every $x \in S$ and $k \in \{0, \ldots, n\}$. Given $x \in S$ we define $F_k(x) := \int f_k d\mu_x$. Observe that

$$\int f_k d\mu = \int_S F_k d\mu.$$

By definition of the integral we know that

$$\int_S F_k d\mu = \sup \left\{ \int_S g d\mu : g \text{ is simple and } g \leq F_k \right\}.$$

This immediately implies that there exists a measurable partition

$$\mathcal{U}_k = \{U_1^{(k)}, \ldots, U_{p_k}^{(k)}\} \quad \text{of } S,$$

such that, given $\varepsilon > 0$,

$$\int_S F_k d\mu - \varepsilon < \sum_{i=1}^{p_k} \left(\inf_{x \in U_i^{(k)}} F_k(x) \right) \mu(U_i^{(k)}) \leq \int_S F_k d\mu.$$

If $\mathcal{P}$ and $\mathcal{Q}$ are partitions, then $\mathcal{P} \wedge \mathcal{Q}$ denotes the common refinement of $\mathcal{P}$ and $\mathcal{Q}$. That is, $A \in \mathcal{P} \wedge \mathcal{Q}$ if $A = P_1 \cap Q_1$, where $P_1 \in \mathcal{P}$ and $Q_1 \in \mathcal{Q}$. Consider the partition $\mathcal{U} = \bigwedge_{i=0}^n \mathcal{U}_i$ of S, and write $\mathcal{U} = \{U_1, \ldots, U_q\}$. Choose a point $x_i \in U_i$, for every $i \in \{1, \ldots, q\}$. Note that

$$\int_S F_k d\mu - \varepsilon < \sum_{i=1}^q F_k(x_i)\mu(U_i) \leq \int_S F_k d\mu,$$

for every $k \in \{0, \ldots, n\}$. We conclude that

$$\left| \int f_k d\mu - \sum_{i=1}^q \mu(U_i) \int f_k d\mu_{x_i} \right| \leq \varepsilon,$$

for every $k \in \{0, \ldots, n\}$. Finally define $\mu_1 = \sum_{i=1}^q \mu(U_i)\mu_{x_i}$. $\qquad \square$

We will now prove a refinement of Theorem 3.8. As mentioned before, this is an important ingredient to increase the applicability of Proposition 6.12. Recall that τ is positive and bounded away from zero.

Proposition 6.14. *Let (Σ, σ) be a transitive countable Markov shift and τ be a uniformly continuous function such that $\mathrm{var}_2(\tau)$ is finite. For every $\mu \in \mathcal{M}_\tau$ there exists a sequence of periodic measures $(\mu_n)_n$ that converges in the weak* topology to μ and such that $\lim_{n \to \infty} \int \tau d\mu_n = \int \tau d\mu$.*

Proof. In virtue of Lemma 6.13 it is possible to find a sequence $(\mu_n)_n$ of measures, each a convex combination of finitely many ergodic probability measures, satisfying $d(\mu_n, \mu) \leq \frac{1}{n}$, and such that $\lim_{n \to \infty} \int \tau d\mu_n = \int \tau d\mu$. In particular, it is enough to prove that the result holds for measures which are a finite convex combination of ergodic measures. We can moreover assume that the weights in the convex combination are rational numbers. Thus, from now on we assume that

$$\mu = \frac{1}{N} \sum_{j=1}^N \mu_j,$$

where each μ_j is ergodic.

Let $\mathcal{F}_0 = \{f_1, \ldots, f_l\} \subset C_b(\Sigma)$ be a collection of bounded uniformly continuous functions on Σ and define $\mathcal{F} := \mathcal{F}_0 \cup \{\tau\}$. By assumption each $f \in \mathcal{F}$ is uniformly continuous. In particular, given $\varepsilon > 0$, there exists $N_0 = N_0(\varepsilon)$ such that $\mathrm{var}_n(f) \leq \frac{\varepsilon}{4}$, for every $f \in \mathcal{F}$ and $n \geq N_0$. Define $C_0 := \max_{f \in \mathcal{F}_0} \max_{x \in \Sigma} |f(x)|$.

Choose M such that $\mu_j(K_M) > 9/10$, for every $j \in \{1, \ldots, N\}$, where

$$K_M = \bigcup_{s=1}^{M} [s].$$

By transitivity of (Σ, σ) there exists a number L such that every pair of numbers in $\{1, \ldots, M\}^2$ can be connected with an admissible word of length at most L. For each pair $(a, b) \in \{1, \ldots, M\}^2$ we choose a point $p_{a,b}$ such that $p_{a,b} \in [a]$ and $\sigma^{c(a,b)-1}(p_{a,b}) \in [b]$, where $c(a, b) \leq L$. Recall that $S_n \tau(x)$ denotes the Birkhoff sum of length n of the point x. Set $C_1 = \max_{a,b} |S_{c(a,b)} \tau(p_{a,b})|$. Define

$$A_{j,\varepsilon}^s = \left\{ x \in \Sigma : \left| \frac{1}{m} \sum_{i=0}^{m-1} f(\sigma^i x) - \int f d\mu_j \right| < \frac{\varepsilon}{4}, \text{ for every } f \in \mathcal{F} \text{ and } m \geq s \right\}.$$

It follows from the Birkhoff ergodic theorem that $\mu_j(A_{j,\varepsilon}^s) \to 1$ as $s \to \infty$. Choose $s_0 \in \mathbb{N}$ such that $\mu_j(A_{j,\varepsilon}^{s_0}) \geq 9/10$, for every $j \in \{1, \ldots, N\}$. We assume that s_0 is sufficiently large (relative to our constants C_0, C_1, N_0 and L) to be determined later.

Observe that

$$\mu_j(A_{j,\varepsilon}^{s_0} \cap K_M \cap \sigma^{-s_0}(K_M)) > \frac{1}{2}.$$

Pick a point $x_j \in A_{j,\varepsilon}^{s_0} \cap K_M \cap \sigma^{-s_0}(K_M)$. We will construct a periodic point x_0 out of the sequence $(x_j)_{j=1}^N$. Let $\mathbf{y_j}$ be the admissible word coming from the first $(s_0 + 1)$ coordinates of x_j. Observe that the first and last letters of $\mathbf{y_j}$ are in $\{1, \ldots, M\}$. We construct an admissible word of the form

$$\mathbf{y} = \mathbf{y_1 w_1 y_2 w_2 \cdots y_N w_N},$$

where $\mathbf{w_i}$ are admissible words of length less than or equal to L that connects $\mathbf{y_i}$ with $\mathbf{y_{i+1}}$ (where we consider $\mathbf{y_{N+1}} = \mathbf{y_1}$). We will moreover assume that the admissible word $\mathbf{w_i}$ is the same one we used to construct the point $p_{a,b}$, for the corresponding a and b. In this case $l(\mathbf{y_i}) = c(a, b)$ and the point associated to $\mathbf{w_i}$ is denoted by $p_i \in (p_{a,b})_{a,b}$. Then define $x = (\mathbf{yy} \cdots)$. We claim that the periodic measure associated to x, say μ_x, belongs to the set

$$\Omega = \left\{ \nu \in \mathcal{M}(\Sigma, \sigma) : \left| \int f d\nu - \int f d\mu \right| < \varepsilon, \text{ for every } f \in \mathcal{F} \right\}.$$

Our construction ensures the following inequalities:

(a) $|S_{s_0-N_0}f(x) - S_{s_0-N_0}f(x_1)| \leq (s_0 - N_0)\mathrm{var}_{N_0}f \leq (s_0-N_0)\varepsilon/4$, for every $f \in \mathcal{F}$.

(b) $|S_{N_0+l(\mathbf{y_1})}f(\sigma^{s_0-N_0}x) - S_{N_0+l(\mathbf{y_1})}f(\sigma^{s_0-N_0}x)| \leq 2(N_0 + l(\mathbf{y_1}))C_0 \leq 2(N_0+L)C_0$, for every $f \in \mathcal{F}_0$.

(c) $|S_{N_0}\tau(\sigma^{s_0-N_0}x) - S_{N_0}\tau(\sigma^{s_0-N_0}x_1)| \leq N_0\mathrm{var}_2(\tau)$.

(d) $|S_{l(\mathbf{y_1})}\tau(\sigma^{s_0}x) - S_{l(\mathbf{y_1})}\tau(p_1)| \leq l(\mathbf{y_1})\mathrm{var}_2(\tau) \leq L\mathrm{var}_2(\tau)$.

We can use the last inequality to obtain that

$$|S_{l(\mathbf{y_1})}\tau(\sigma^{s_0}x)| \leq L\mathrm{var}_2(\tau) + C_1.$$

Combining these inequalities we obtain that

$$|S_{s_0+l(\mathbf{y_1})}\tau(x) - S_{s_0}\tau(x_1)| \leq \frac{1}{4}(s_0 - N_0)\varepsilon + N_0\mathrm{var}_2(\tau) + L\mathrm{var}_2(\tau) + C_1,$$

and that

$$|S_{s_0+l(\mathbf{y_1})}f(x) - S_{s_0+l(\mathbf{y_1})}f(x_1)| \leq \frac{1}{4}(s_0 - N_0)\varepsilon + 2(N_0+L)C_0,$$

where $f \in \mathcal{F}_0$. Similar inequalities can be obtained when comparing the value of our function f at $\sigma^{(k-1)s_0+\sum_{i=1}^{k-1} l(\mathbf{y_i})}(x)$ and at x_k, where $f \in \mathcal{F}$. Using the triangle inequality and the definition of $A_{j,\varepsilon}^{s_0}$ we can estimate $|\int f d\mu_x - \sum_{j=1}^{N}\int f d\mu_j|$ in an effective way. By taking s_0 large enough (in terms of our constants C_0, C_1, N_0 and L) we can ensure that $\mu_x \in \Omega$. We leave the details to the reader. We have now proved that the result holds for convex combinations of ergodic measures and, as explained at the beginning of the proof, the general result follows from Lemma 6.13. $\qquad \square$

Corollary 6.15. *Assume that $\tau \in \mathcal{R}$. Then the space of ergodic measures $\mathcal{M}_e(\Sigma, \sigma, \tau)$ is weak* dense in $\mathcal{M}(\Sigma, \sigma, \tau)$.*

Proof. Fix $\nu \in \mathcal{M}(\Sigma, \sigma, \tau)$. We will prove that ν can be approximated in the topology of convergence on cylinders by ergodic measures. Let $\mu = \psi(\nu)$. By Proposition 6.14 we can find a sequence of periodic measures $(\mu_n)_n \subset \mathcal{M}(\Sigma, \sigma)$ converging to μ in the weak* topology and such that $\lim_{n \to \infty}\int \tau d\mu_n = \int \tau d\mu$. Set $\nu_n = \varphi(\mu_n)$. Lemma 6.10 implies that $(\nu_n)_n$ converges in the cylinder topology to ν. In virtue of Lemma 6.7 we have that $(\nu_n)_n$ converges in the weak* topology to ν. Moreover, each ν_n is ergodic, since each μ_n is ergodic. This concludes the proof of the corollary. $\qquad \square$

We have finally all the ingredients to prove the main result of this section: the compactness of $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$. We already know that $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$ is a metrizable topological space; it is enough to prove it is sequentially compact.

Theorem 6.16. *Assume that $\tau \in \mathcal{R}$. Let $(\nu_n)_n$ be a sequence of invariant probability measures of the suspension flow. Then there exists a sub-sequence $(\nu_{n_k})_k$ converging on cylinders to an invariant sub-probability measure ν.*

Proof. Let $\mu_n = \psi(\nu_n)$. If $\limsup_{n\to\infty} \int \tau d\mu_n = \infty$, there exists a sub-sequence of $(\nu_n)_n$ that converges on cylinders to the zero measure. We will assume that

$$\int \tau d\mu_n \le M,$$

for all $n \in \mathbb{N}$. By compactness of $L(\Sigma)$ there exists a sub-sequence $(\mu_{n_k})_k$ that converges on cylinders to $F \in L(\Sigma)$. Maybe after passing to a sub-sequence we can assume that $\lim_{k\to\infty} \int \tau d\mu_{n_k} = L$, for some $L \in \mathbb{R}$. We can now use Proposition 6.14 to obtain periodic measures η_k satisfying $d(\mu_{n_k}, \eta_k) \le \frac{1}{k}$ and $\int \tau d\eta_k \le (\int \tau d\mu_{n_k} + 1)$. Note that $\lim_{k\to\infty} d(\eta_k, F) = 0$ and $\int \tau d\eta_k \le (M+1)$, for all $k \in \mathbb{N}$. We can now use Proposition 6.12 to conclude that F extends to an invariant probability measure that we denote by μ. It follows that $(\mu_{n_k})_k$ converges on cylinders to μ and that $\lim_{k\to\infty} \int \tau d\mu_{n_k} = L$. Finally, use Lemma 6.10 to obtain that $(\nu_{n_k})_k$ converges on cylinders to $\lambda\nu$, where $\nu = \varphi(\mu)$ and $\lambda = \frac{\int \tau d\mu}{L}$.

6.3 The space of flow invariant sub-probability measures is the Poulsen simplex.

In Section 5 we proved that $\mathcal{M}_{\le 1}(\Sigma, \sigma)$ is affine homeomorphic to the Poulsen simplex if (Σ, σ) has the $\mathcal{F}$-property. In this section we prove an analogous result for the suspension flow, that is, that $\mathcal{M}_{\le 1}(\Sigma, \sigma, \tau)$ is affine homeomorphic to the Poulsen simplex, provided that $\tau \in \mathcal{R}$.

Theorem 6.17. *Let τ be a potential in $\mathcal{R}$. The space $\mathcal{M}_{\le 1}(\Sigma, \sigma, \tau)$ endowed with the topology of convergence on cylinders is affine homeomorphic to the Poulsen simplex.*

Proof. In Theorem 6.16 we proved that the space $\mathcal{M}_{\le 1}(\Sigma, \sigma, \tau)$ is compact with respect to the cylinder topology. In Lemma 6.6 we showed that it is a metrizable space. Since the space is also a convex Choquet simplex (from the ergodic decomposition), it suffices to prove that the set of extreme points is dense. Note that every element of $\mathcal{M}_{\le 1}(\Sigma, \sigma, \tau)$ is of the form $\lambda\nu$, with $\lambda \in [0, 1]$ and $\nu \in \mathcal{M}(\Sigma, \sigma, \tau)$. In Lemma 6.11 we proved that there exists a sequence of flow invariant ergodic measures $(\tilde{\nu}_n)_n$ converging to the zero measure. Set $\hat{\nu}_n := \lambda\nu + (1-\lambda)\tilde{\nu}_n$, and observe that the sequence $(\hat{\nu}_n)_n$ converges on cylinders to $\lambda\nu$. It follows from Corollary 6.15 that the ergodic measures are dense in $\mathcal{M}(\Sigma, \sigma, \tau)$. This allows us to approximate $(\hat{\nu}_n)_n$ by a sequence of ergodic measures that converges in cylinders to $\lambda\nu$. It follows from the main result of [LOS, Theorem 2.3] that $\mathcal{M}_{\le 1}(\Sigma, \sigma, \tau)$ is affine homeomorphic to the Poulsen simplex. $\square$

As in Section 5 we conclude that

Theorem 6.18. *Let* (Σ, σ) *be a transitive countable Markov shift and* $\tau \in \mathcal{R}$. *Then* $\mathcal{M}(\Sigma, \sigma, \tau)$ *is affinely homeomorphic to the Poulsen simplex minus a vertex and all of its convex combinations.*

Proof. Note that the set of extreme points of $\mathcal{M}_{\leq 1}(\Sigma, \sigma, \tau)$ is the zero measure together with the set of ergodic measures in $\mathcal{M}(\Sigma, \sigma, \tau)$. The result now follows from Theorem 6.17 together with the relation between weak* and cylinder topology (see Lemma 6.7). $\qquad\square$

REFERENCES

[AK] W. Ambrose and S. Kakutani, *Structure and continuity of measurable flows*, Duke Math. J. **9** (1942), 25–42.

[Bi] P. Billingsley, *Convergence of Probability Measures*, Second edition, John Wiley & Sons, New York, 1999.

[B] V. I. Bogachev, *Measure Theory. Vol. I, II*, Springer, Berlin, 2007.

[BBG] M. Boyle, J. Buzzi and R. Gómez, *Borel isomorphism of SPR Markov shifts*, Colloq. Math. **137** (2014), 127–136.

[BCS] J. Buzzi, S. Crovisier and O. Sarig *Measures of maximal entropy for surface diffeomorphisms*, arXiv:1811.02240 [math.DS]

[CS] Y. Coudène and B. Schapira, *Generic measures for hyperbolic flows on non-compact spaces*, Israel J. Math. **179** (2010), 157–172.

[Di] J. Dieudonné, *Éléments d'Analyse. Tome II: Chapitres XII–XV*, Gauthier-Villars, Éditeur, Paris 1968.

[D] T. Downarowicz, *The Choquet simplex of invariant measures for minimal flows*, Israel J. Math. **74** (1991), 241–256.

[Gu1] B. M. Gurevič, *Topological entropy for denumerable Markov chains*, Dokl. Akad. Nauk SSSR **10** (1969), 911–915.

[Gu2] B. M. Gurevič, *Shift entropy and Markov measures in the path space of a denumerable graph*, Dokl. Akad. Nauk SSSR **11** (1970), 744–747.

[GK] B. M. Gurevich and S. Katok, *Arithmetic coding and entropy for the positive geodesic flow on the modular surface*, Mosc. Math. J. **1** (2001), 569–582, 645.

[H] F. Hofbauer, *Examples for the nonuniqueness of the equilibrium state*, Trans. Amer. Math. Soc. **228** (1977), 223–241.

[ITV] G. Iommi, M. Todd and A. Velozo, *Escape of mass and entropy of countable Markov shifts*, arXiv:1908.10741 [math.DS]

[KU] S. Katok and I. Ugarcovici, *Symbolic dynamics for the modular surface and beyond.* Bull. Amer. Math. Soc. (N.S.) **44** (2007), 87–132.

[Ki] B.P. Kitchens, *Symbolic Dynamics. One-Sided, Two-Sided and Countable State Markov Shifts*, Springer, Berlin, 1998.

[Kl] A. Klenke, *Probability Theory. A Comprehensive Course*, Second edition, Springer, London, 2014.

[LM] Y. Lima and C. Matheus, *Symbolic dynamics for non-uniformly hyperbolic surface maps with discontinuities*, Ann. Sci. Ec. Norm. Supér. (4) **51** (2018), 1–38.

[LOS]   J. Lindenstrauss, G. Olsen and Y. Sternfeld, *The Poulsen simplex*, Ann. Inst. Fourier (Grenoble) **28** (1978), 91–114.

[MP]   N. Markley and M. Paul, *Equilibrium states of grid functions*, Trans. Amer. Math. Soc. **274** (1982), 169–191.

[O]   J.C. Oxtoby, *On two theorems of Parthasarathy and Kakutani concerning the shift transformation*, in *Ergodic Theory*, Academic Press, New York, 1963, pp. 203–215.

[PP]   W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque **187–188** (1990).

[Par1]   K. R. Parthasarathy, *On the category of ergodic measures*, Illinois J. Math. **5** (1961), 648–656.

[Par2]   K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York–London 1967.

[PW]   M. Pollicott and H. Weiss, *Multifractal analysis of Lyapunov exponent for continued fraction and Manneville-Pomeau transformations and applications to Diophantine approximation*, Comm. Math. Phys. **207** (1999), 145–171.

[Pou]   E. T. Poulsen, *A simplex with dense extreme points*, Ann. Inst. Fourier (Grenoble) **11** (1961), 83–87.

[Sa1]   O. Sarig, *Phase transitions for countable Markov shifts*, Comm. Math. Phys. **217** (2001), 555–577.

[Sa2]   O. Sarig, *Existence of Gibbs measures for countable Markov shifts*, Proc. Amer. Math. Soc. **131** (2003), 1751–1758.

[Sa2]   O. Sarig, *Symbolic dynamics for surface diffeomorphisms with positive entropy*, J. Amer. Math. Soc. **26** (2013), 341–426.

[Si1]   K. Sigmund, *Generic properties of invariant measures for Axiom* A *diffeomorphisms*, Invent. Math. **11** (1970), 99–109.

[Si2]   K. Sigmund, *On the space of invariant measures for hyperbolic flows*, Amer. J. Math. **94** (1972), 31–37.

[Si3]   K. Sigmund, *On mixing measures for axiom a diffeomorphisms*, Proc. Amer. Math. Soc. **36** (1972), 497–504.

[V]   A. Velozo, *The pressure at infinity of countable Markov shifts*, preprint.

[Wa]   P. Walters, *An Introduction to Ergodic Theory*, Springer, Berlin, 1981.

Godofredo Iommi
FACULTAD DE MATEMÁTICAS
   PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE (PUC)
      AVENIDA VICUÑA MACKENNA 4860, SANTIAGO, CHILE
         email: giommi@mat.puc.cl
            http://http://www.mat.uc.cl/˜giommi/

Anibal Velozo
DEPARTMENT OF MATHEMATICS
   YALE UNIVERSITY
      NEW HAVEN, CT 06511, USA
         email: anibal.velozo@yale.edu
            https://gauss.math.yale.edu/˜av578/