ARCWISE CONNECTEDNESS OF THE SET OF ERGODIC MEASURES OF HEREDITARY SHIFTS

JAKUB KONIECZNY, MICHAL KUPSA, AND DOMINIK KWIECNIK

(Communicated by Nimish Shah)

Abstract. We show that the set of ergodic invariant measures of a shift space with a safe symbol (this includes all hereditary shifts) is arcwise connected when endowed with the \( d \)-bar metric. As a consequence the set of ergodic measures of such a shift is also arcwise connected in the weak-star topology, and the entropy function over this set attains all values in the interval between zero and the topological entropy of the shift (inclusive). The latter result is motivated by a conjecture of A. Katok.

A shift space \( X \) over the alphabet \( \Lambda = \{0,1,\ldots,n-1\} \) for some \( n \geq 2 \) is **hereditary** if \( x \in X \) and \( y \leq x \) (coordinate-wise) imply \( y \in X \). Hereditary shifts were introduced by Kerr and Li in [20, p. 882], and their basic properties are presented in [25]. We say that \( a \in \Lambda \) is a **safe symbol** for a shift space \( X \subset \Lambda^\Z \) (see [31]) if for every \( z \in \Lambda^\Z \) obtained by replacing some entries in \( y \in X \) by \( a \), we have that \( z \in X \). By definition 0 is a safe symbol for every hereditary shift.

The family of hereditary shifts includes: spacing shifts, beta shifts, bounded density shifts, \( \mathcal{B} \)-admissible shifts; also, many examples of \( \mathcal{B} \)-free shifts and some shifts of finite type are hereditary. All classes on that list have been extensively studied, with the \( \mathcal{B} \)-free shifts attracting much attention recently (see Section 3 for more details). In the setting of actions of \( \Z \), as in our paper, shifts with a safe symbol seem to be less important. The notion is useful in the context of higher dimensional shifts (\( \Z^d \) actions with \( d \geq 2 \); see [31] and references therein). It is easy to find examples of shift spaces over \( \{0,1,2\} \) which have 0 as a safe symbol but are not hereditary.

It should come as no surprise that there are very few theorems applicable to all members of such a diverse family of shift spaces. Nevertheless, the main result of this note implies that there is a common feature of all hereditary shift spaces: for any hereditary shift \( X \) and for every \( t \geq 0 \) the set of ergodic invariant measures with entropy less than or equal to \( t \), denoted \( \mathcal{M}_\sigma(t)(X) \), and endowed with the \( d \)-bar metric...
metric \( \bar{d}_\mathcal{M} \) is arcwise connected (Theorem 6). In fact, we only need to assume that there exists a safe symbol, and our proof actually shows that there is a \( d_{\Omega} \)-continuous arc in \( X \) consisting of generic points for ergodic measures from the arc in \( \mathcal{M}_\sigma^{(t)}(X) \). This is a single orbit result very much in the spirit of [11]. Here \( \bar{d}_\Omega \) is a pseudometric on \( X \) given by the upper asymptotic density of the set of indices where two sequences in \( X \) differ.

The \( d \)-bar metric \( \bar{d}_\mathcal{M} \) induces a stronger topology than the usual weak* topology on the space of ergodic invariant measures. It follows that in the latter topological space the set of ergodic invariant measures with entropy less than or equal to \( t \) is also arcwise connected for every \( t \geq 0 \) (Corollary 7). The same holds for sets of ergodic measures with entropy strictly less than \( t \). Furthermore, since the entropy function \( h \) taking an ergodic measure \( \mu \) to its metric entropy \( h(\mu) \) is \( \bar{d}_\mathcal{M} \)-continuous, it has the Darboux (intermediate-value) property over every arc in \( \bar{d}_\mathcal{M} \). Denoting the topological entropy of a shift space \( X \) by \( h_{\text{top}}(X) \) we say that \( X \) has the intermediate entropy property over ergodic measures if for every \( \alpha \in [0, h_{\text{top}}(X)] \) there is an ergodic measure \( \mu \) with \( h(\mu) = \alpha \). In particular, every shift with a safe symbol has the intermediate entropy property, and its set of ergodic measures is either a singleton or is uncountable (Corollary 8). The former case occurs if and only if \( h_{\text{top}}(X) = 0 \) (Corollary 9).

Our result about the intermediate entropy property is motivated by the following conjecture due to A. Katok: If \( r > 1 \) and \( F: M \to M \) is a \( C^r \) diffeomorphism of a smooth compact manifold \( M \), then for every \( \alpha \in [0, h_{\text{top}}(F)] \), there is an \( F \)-invariant ergodic measure \( \mu \) such that the metric entropy of \( F \) with respect to \( \mu \) equals \( \alpha \). Katok proved that this is the case if \( M \) is a compact surface where every ergodic measure of positive entropy is hyperbolic ([18]; for a detailed proof see [19, Theorem 8.5.9]). Katok’s result was extended to certain skew product cases by Sun [38,39]. The conjecture also holds for every ergodic linear automorphism of the torus as a result of work of Quas and Soo [29] and for some partially hyperbolic diffeomorphism by Ures [40].

The approach presented here is different from the methods used in [18,29,38,40]. After publishing this paper on the arXiv we learned that similar techniques were applied in [28].

We also describe the construction of some examples illustrating that our main theorem may not hold without the assumption that the shift space is hereditary. By the same examples one can see that the conclusions of Corollary 7 and Corollary 8 are independent: neither of them implies the other. This is a manifestation of the well-known fact that the metric entropy function \( h \) on the set of invariant measures \( \mathcal{M}_\sigma(X) \) of a shift space \( X \) endowed with the weak* topology is, in general, only upper semi-continuous. On the other hand every Polish topological space is homeomorphic to a set of ergodic measures of some shift space endowed with the weak* topology (see [9, Theorem 5] and [17]). This suggests that there should be plenty of examples of shift spaces without the intermediate entropy property. It is obvious that this is the case if the shift space has at most countably many ergodic invariant measures. It is less obvious if the set of ergodic measures is arcwise connected in the weak* topology. Using a characterization of possible entropy functions due to Downarowicz and Serafin [10] we show that for every non-trivial Polish topological space \( P \) there is a shift space \( X \) whose set of ergodic measures \( \mathcal{M}_\sigma^*(X) \) endowed with the weak* topology is homeomorphic with \( P \) and the entropy
function has an isolated positive value (see Theorem 11). By the same technique, we show that for every uncountable Polish topological space \( P \) (not necessarily connected!) there is a shift space \( X \) with \( \mathcal{M}_P(X) = P \) and a non-constant entropy function with the intermediate entropy property (see Theorem 12).

Finally, our first proof of Theorem 6 was based on a relative version of the Furstenberg unique ergodicity theorem (Theorem 16 below). We no longer need this to prove our main theorem, but as we hope it is a result of independent interest we attach it with a proof in Appendix A.

1. Definitions

By a dynamical system we mean a pair \((X,T)\), where \( X \) is a compact metric space and \( T: X \to X \) is a homeomorphism. If \((Y,S)\) is another dynamical system and \( \pi: X \to Y \) is a continuous map onto with \( \pi \circ T = S \circ \pi \), then we call \( \pi \) a factor map, \( Y \) a factor of \( X \), and \( X \) an extension of \( Y \).

The set of all Borel probability measures on \( X \) is denoted by \( \mathcal{M}(X) \). The usual weak* topology makes \( \mathcal{M}(X) \) a compact metrizable space.

Let \( \mathcal{M}_T(X) \) denote the set of \( T \)-invariant measures in \( \mathcal{M}(X) \). We write \( \mathcal{M}_P(X) \) for the set of all ergodic measures in \( \mathcal{M}_T(X) \), and \( h(\mu) \) denotes the Kolmogorov-Sinai entropy of \( \mu \). For each \( \mu \in \mathcal{M}_T(X) \) we call the triple \((X,T,\mu)\) a measure-preserving system. Note that this is more restrictive than the usual definition.

A point \( x \in X \) is generic for a measure \( \mu \in \mathcal{M}_T(X) \) if for each continuous function \( \varphi: X \to \mathbb{R} \) we have

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{j=0}^{N-1} \varphi(T^j(x)) = \int_X \varphi \, d\mu.
\]

The set of all generic points for \( \mu \) is denoted by \( \text{Gen}_T(\mu) \). We write \( \text{Gen}(X,T) \) for the set of points that are generic for some \( T \)-invariant measure. For \( x \in \text{Gen}(X,T) \) we denote by \( \hat{\mu}(x) \) the measure for which \( x \) is generic.

Fix \( n \geq 2 \) and let \( \Lambda \) be a finite set with \( n \) elements; without loss of generality we assume that \( \Lambda = \{0,1,\ldots,n-1\} \). Let \( \Omega = \Lambda^\mathbb{Z} \) be the set of all two-sided \( \Lambda \)-valued sequences. We equip \( \Omega \) with the product (Tikhonov) topology induced by the discrete topology on \( \Lambda \). The shift map \( \sigma: \Omega \to \Omega \) is given by \( \sigma(x)_i = x_{i+1} \).

By a shift space (for short a shift) we mean any non-empty closed set \( X \subset \Omega \) such that \( \sigma(X) = X \).

2. Disjointness and spectral theory

In this section we will review some basic facts from the spectral theory of Koopman operators. Since these results are classical we have not attempted to document the source of every one of them (see [16,26]). We need them for a proof of Theorem 4 which may be known among aficionados, but we were unable to find it in the literature.

Using the map \( \psi(t) = \exp(2\pi it) \) we identify the quotient group \( \mathbb{T}^1 = \mathbb{R}/\mathbb{Z} \) equipped with addition mod 1 and the circle, that is, the multiplicative group \( 
\mathbb{S}^1 = \{ z \in \mathbb{C} : |z| = 1 \} 
\). By \( \lambda \) we denote the Lebesgue measure on \( \mathbb{T}^1 \) and its image (pushforward) through \( \psi \) on \( \mathbb{S}^1 \). Given \( \alpha \in \mathbb{R} \setminus \mathbb{Q} \) the map \( R_\alpha: \mathbb{T}^1 \to \mathbb{T}^1 \) given by \( R_\alpha(t) = t + \alpha \) mod 1 is called an irrational rotation of the circle. It is well known that \( \lambda \) is the unique invariant measure for \( R_\alpha \).
Given a measure-preserving system \((X, T, \mu)\) we define the Koopman operator \(U_T : L^2(X, \mu) \to L^2(X, \mu)\) by \(U_T(f) = f \circ T\). The Koopman operator is unitary; hence for each \(f \in L^2\) the bi-infinite sequence \(r_n = \langle U_T^n(f), f \rangle\) is a positive definite sequence of complex numbers. By the Herglotz theorem the positive definite sequence \(\{U_T^n(f), f\}\) determines uniquely a finite non-negative Borel measure on circle \(\mathbb{T}\), called the spectral measure of \(f\) and denoted by \(\sigma_f\).

A cyclic subspace determined by \(f \in L^2(X, \mu)\), denoted by \(Z(f)\), is the closure of the linear span of \(\{U^n(f) : n \in \mathbb{Z}\}\). A cyclic subspace is maximal if it is not contained in any larger cyclic subspace. The spectral theorem for unitary operators allows us to define the spectral type \(\sigma_T\), which is (up to equivalence) the spectral measure determined by any vector determining a maximal cyclic subspace. The spectral type \(\sigma_T\) can be written as \(\sigma_T = \sigma_d + \sigma_c\), where \(\sigma_d\) is purely atomic measure and \(\sigma_c\) is a continuous (non-atomic) measure.

The point spectrum of a measure-preserving transformation \((X, T, \mu)\) is the set of eigenvalues for the Koopman operator \(U_T(f) = f \circ T\), i.e.,

\[
H(T, \mu) = \{\lambda \in \mathbb{C} : \lambda f = f \circ T \text{ for some } f \in L^2(X, \mu) \text{ with } f \neq 0\}.
\]

The point spectrum is a subset of the unit circle \(S^1 \subset \mathbb{C}\). Since \(L^2(X, \mu)\) is separable, the point spectrum is at most countable and, in the ergodic case, it forms a multiplicative subgroup of \(S^1\). Furthermore, the spectral measure of each eigenvector is discrete. The purely atomic part of the spectral type is closely connected with spectral measures associated to eigenvectors; in particular, \(\sigma_d\) is supported on \(H(T, \mu)\).

Let \((X, T, \mu)\) and \((Y, S, \nu)\) be measure-preserving systems. A joining of \((X, T, \mu)\) and \((Y, S, \nu)\) is a \(T \times S\)-invariant measure \(\eta\) on \(X \times Y\) with marginals \(\mu\) and \(\nu\). We write \(J(\mu, \nu)\) for the set of all joinings of \((X, T, \mu)\) and \((Y, S, \nu)\). We say that \((X, T, \mu)\) and \((Y, S, \nu)\) are disjoint if \(\mu \times \nu\) is the only joining of these two systems.

We will also use in the proofs of our theorems the description of the spectrum for rotations and the characterization of the ergodicity of the product of two transformations via their spectra. If \(R_\alpha\) is an irrational rotation, then the spectrum \(H(R_\alpha, \lambda)\) with respect to the invariant Lebesgue measure \(\lambda\) is generated by \(e^{2\pi i \alpha}\), i.e.,

\[
H(R_\alpha, \lambda) = \{\exp(2\pi ik\alpha) : k \in \mathbb{Z}\}.
\]

Furthermore, the spectral type of the Koopman operator associated with \((\mathbb{T}^1, R_\alpha, \lambda)\) is purely atomic.

**Lemma 1.** The product of two ergodic measure-preserving systems \((X, T, \mu)\) and \((Y, S, \nu)\) is ergodic if and only if their point spectra have trivial intersection, i.e.,

\[
H(T, \mu) \cap H(S, \nu) = \{1\}.
\]

This fact might not be a part of classical textbooks on ergodic theory, but can be found e.g. in [26, p. 46, Exercise 2]. In a similar spirit, we have the following criterion for disjointness.

**Theorem 2** (Theorem 6.28 in [16]). If \((X, T, \mu)\) and \((Y, S, \nu)\) are measure-preserving systems whose spectral types are mutually singular except for the common atom at 1, then \((X, T, \mu)\) and \((Y, S, \nu)\) are disjoint.

The following lemma was suggested to us by Lemańczyk and Przytycki.

**Lemma 3.** Let \(\alpha \in \mathbb{R} \setminus \mathbb{Q}\). The measure-preserving systems \((\mathbb{T}^1, R_\alpha, \lambda)\) and \((X, T, \mu)\) are disjoint if and only if \(\exp(2\pi ik\alpha) \notin H(T, \mu)\) for \(k \in \mathbb{Z} \setminus \{0\}\).
Proof. For the “if” direction, the assumptions imply that $H(T, \mu) \cap H(R_\alpha, \lambda) = \{1\}$. Since the spectral type of $R_\alpha$ is purely atomic and the purely atomic part of the spectral type of $T$ is supported on $H(T, \mu)$, the spectral types of $(X, T, \mu)$ and $(\mathbb{T}^1, R_\alpha, \lambda)$ are mutually singular (except the common atom at 1). By Theorem \ref{thm:spec_type} this implies that $(\mathbb{T}^1, R_\alpha, \lambda)$ and $(X, T, \mu)$ are disjoint. For the “only if” direction, we first use Lemma \ref{lem:joinings} to note that if $\exp(2\pi ik\alpha) \in H(T, \mu)$ for some $k \in \mathbb{Z} \setminus \{0\}$, then $\lambda \times \mu$ is not ergodic. This implies that the systems $(\mathbb{T}^1, R_\alpha, \lambda)$ and $(X, T, \mu)$ cannot be disjoint because the product of ergodic and disjoint systems is always ergodic by Theorem 6.2 in \cite{16}.

**Theorem 4.** For every ergodic measure-preserving system $(X, T, \mu)$ there exists an irrational $\alpha$ such that $(\mathbb{T}^1, R_\alpha, \lambda)$ and $(X, T, \mu)$ are disjoint.

**Proof.** The point spectrum of $(X, T, \mu)$ is a countable subgroup of the circle; hence there is an $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ such that $\exp(2\pi ik\alpha) \notin H(T, \mu)$ for all $k \in \mathbb{Z} \setminus \{0\}$. It remains to apply Lemma \ref{lem:irrational}

3. **Arcwise connectedness**

In this section we prove our main result. But first we prepare some notation and state an auxiliary lemma.

Let $Z$ be a topological space and let $x, y \in Z$. A path (resp. arc) from $x$ to $y$ in $Z$ is a continuous function (resp. homeomorphism onto the image) $\gamma : [0, 1] \to Z$ such that $\gamma(0) = x$ and $\gamma(1) = y$. The space $Z$ is pathwise connected (resp. arcwise connected) if for every $x, y \in Z$ there is a path (resp. an arc) from $x$ to $y$.

By $d(A)$ we denote the upper asymptotic density of a set $A \subset \mathbb{Z}$. Recall that

$$d(A) = \limsup_{n \to \infty} \frac{|A \cap \{1, \ldots, n\}|}{n}.$$ 

Given $x, y \in \Omega = \Lambda^Z$ with $x = (x_j)_{j \in \mathbb{Z}}$ and $y = (y_j)_{j \in \mathbb{Z}}$ the formula

$$d_\Omega(x, y) = d\{n \in \mathbb{N} \mid x_n \neq y_n\}$$

defines a pseudometric on $\Omega$. Note that $d_\Omega(x, y) = 0$ implies that $x$ and $y$ differ on a set of coordinates of zero upper asymptotic density. The $d$-bar pseudometric $\bar{d}_\Omega$ is closely connected with a metric on the set of shift-invariant measures $\Omega$ stronger than the usual metric determining the weak* topology. This metric, denoted by $\bar{d}_\mathcal{M}$, is defined for $\mu, \nu \in \mathcal{M}_\sigma(\Omega)$ by

$$\bar{d}_\mathcal{M}(\mu, \nu) = \inf_{\eta \in J(\mu, \nu)} \eta\{(x, y) \in \Omega \times \Omega : x_0 \neq y_0\},$$

whereas in the previous section $J(\mu, \nu)$ denotes the set of all joinings of $\mu$ and $\nu$ (see \cite{34} Theorem I.9.7). The link between the $d$-bar metric $\bar{d}_\mathcal{M}$ on measures and the pseudometric $\bar{d}_\Omega$ on generic points needed here is the following corollary of Lemma I.9.8 in \cite{34}:

$$\bar{d}_\mathcal{M}(\bar{\mu}(x), \bar{\mu}(y)) \leq \bar{d}_\Omega(x, y), \quad x, y \in \text{Gen}(\Omega, \sigma).$$

For $x \in \Omega$ and $y \in \{0, 1\}^\mathbb{Z}$ we denote by $x \ast y$ the coordinate-wise product: $(x \ast y)_j = x_j \cdot y_j$. If $X \subset \Omega$ and $Y \subset \{0, 1\}^\mathbb{Z}$ are shift spaces, then we denote the image of $X \times Y \subset \Omega$ through $\ast$ as $X \ast Y$. Observe that 0 is a safe symbol for a shift space $X$ over $\{0, 1, \ldots, n-1\}$ if and only if $X \ast \{0, 1\}^\mathbb{Z} = X$. It is clear that $\ast : X \times Y \to X \ast Y$ is a factor map of $(X \times Y, \sigma \times \sigma)$ onto $(X \ast Y, \sigma)$. It follows that if $(x, y) \in \text{Gen}(X \times Y, \sigma \times \sigma)$, then $x \ast y \in \text{Gen}(X \ast Y, \sigma)$. Unfortunately, it
is not always the case that if \( x \) and \( y \) are generic points, then \( (x, y) \) is generic; see [13, p. 22]. But it holds if we assume that the corresponding measures are disjoint. The proof follows the same lines as the proof of Theorem I.6 in [13].

**Lemma 5.** If \( (X, T, \mu) \) and \( (Y, S, \nu) \) are disjoint measure-preserving systems, \( x \in \text{Gen}_T(\mu) \), and \( y \in \text{Gen}_S(\nu) \), then \( (x, y) \in \text{Gen}_{T \times S}(\mu \times \nu) \).

Finally, given a shift space and \( t \geq 0 \) we write \( \mathcal{M}_\sigma(t)(X) \) for the set of ergodic measures on \( X \) with entropy less than or equal to \( t \), that is, \( \mathcal{M}_\sigma(t)(X) = \{ \mu \in \mathcal{M}_\sigma^e(X) : h(\mu) \leq t \} \). This was the last piece we needed for the proof of our main result.

**Theorem 6.** If \( X \) is a shift space with a safe symbol (in particular, if \( X \) is a hereditary shift) and \( t \geq 0 \), then \( \mathcal{M}_\sigma^e(t)(X) \) endowed with \( d_\Lambda \) is arcwise connected.

**Proof.** Without loss of generality we assume that \( 0 \) is a safe symbol for \( X \). By [42, Cor. 31.6] every pathwise connected Hausdorff space is also arcwise connected. Therefore it suffices to show that for any \( \mu \in \mathcal{M}_\sigma^e(X) \) there exists a path from \( \mu \) to the Dirac measure \( \delta_0 \), where \( 0 \) denotes the bi-infinite sequence of 0’s. That is, we need to define a continuous function \( \hat{\Phi} : [0, 1] \to \mathcal{M}_\sigma^e(X) \) with \( \hat{\Phi}(0) = \delta_0 \) and \( \hat{\Phi}(1) = \mu \) which is continuous when we endow \( \mathcal{M}_\sigma^e(X) \) with the \( d_\Lambda \) metric. To do that we construct a \( d_\Omega \)-continuous path of generic points. Since the topology introduced by \( d_\Omega \) on \( X \) is not Hausdorff we consider an equivalence relation on \( X \) defined by \( x \equiv y \) if \( d_\Omega(x, y) = 0 \). The resulting set of equivalence classes endowed with the metric induced by \( d_\Omega \) is called the Besicovitch space of \( X \) and is denoted \( X_B \). It is easy to see that a \( d_\Omega \) continuous path in \( X \) leads to a path in \( X_B \) and the arc in \( X_B \) leads to a \( d_\Omega \)-continuous arc in \( X \). Hence it remains to define a \( d_\Omega \)-continuous path of generic points of ergodic measures in \( X \).

For \( \alpha, \beta \in [0, 1] \), define the point \( y_{\alpha, \beta} \in \Omega \) by

\[
y_{\alpha, \beta} = (\chi_{[0, \beta]}(j \alpha \mod 1))_{j \in \mathbb{Z}} \in \Omega,
\]

where \( \chi_{[0, \beta]} \) denotes the characteristic function of \( [0, \beta] \). Note that \( y_{\alpha, 0} = 0 \) and \( y_{\alpha, 1} = 1 \), where \( 1 \) denotes the bi-infinite sequence of 1’s. Write \( Y_{\alpha, \beta} \) for the closure of the orbit of \( y_{\alpha, \beta} \) with respect to \( \sigma \).

It is a well-known fact that if \( \alpha \) is irrational and \( 0 < \beta < 1 \), then the dynamical system \((Y_{\alpha, \beta}, \sigma)\) is minimal and has a unique invariant measure, which we denote by \( \nu_{\alpha, \beta} \). Furthermore, the measure-preserving system \((Y_{\alpha, \beta}, \sigma, \nu_{\alpha, \beta})\) is isomorphic (in the category of measure-preserving systems) to the circle rotation \((\mathbb{T}^1, R_\alpha, \lambda)\). In addition, let \( \nu_{\alpha, 0} = \delta_0 \) be the Dirac measure concentrated on 0 and similarly \( \nu_{\alpha, 1} = \delta_1 \). With this notation \( y_{\alpha, \beta} \) is generic for \( \nu_{\alpha, \beta} \) for any \( \beta \in [0, 1] \). Therefore, for \( \beta \in (0, 1) \), measure-preserving systems \((Y, S, \nu)\) and \((Y_{\alpha, \beta}, \sigma, \nu_{\alpha, \beta})\) are disjoint if and only if \((Y, S, \nu)\) and \((\mathbb{T}^1, R_\alpha, \lambda)\) are disjoint.

Fix the choice of an irrational \( \alpha \in [0, 1] \) such that \((X, \sigma, \mu)\) and \((\mathbb{T}^1, R_\alpha, \lambda)\) are disjoint, whose existence is assured by Theorem 4. By Lemma 5 for any \( \mu \)-generic point \( x \in X \) and any \( \beta \in [0, 1] \), the point \((x, y_{\alpha, \beta})\) is generic for the ergodic measure \( \mu \times \nu_{\alpha, \beta} \). It follows that \( x \star y_{\alpha, \beta} \) is a generic point for some ergodic measure \( \mu_\beta = \hat{\mu}(x \star y_{\alpha, \beta}) \) for \( \beta \in [0, 1] \). Furthermore, \( x \star y_{\alpha, 0} = 0 \) and \( x \star y_{\alpha, 1} = x \); hence \( \mu_0 = \delta_0 \) and \( \mu_1 = \mu \). Note that \( h(\mu \times \nu_{\alpha, \beta}) = h(\mu) \); thus \( h(\mu_\beta) \leq h(\mu) \leq t \) for all \( 0 \leq \beta \leq 1 \).
For $0 \leq \beta < \beta' \leq 1$ we have $(y_{\alpha, \beta})_j \neq (y_{\alpha, \beta'})_j$ if and only if $R^j_\alpha(0) \in [\beta, \beta')$. Therefore
\[\bar{d}_\Omega(y_{\alpha, \beta}, y_{\alpha, \beta'}) = \bar{d}(\{j \in \mathbb{Z} : (y_{\alpha, \beta})_j \neq (y_{\alpha, \beta'})_j \}) = \bar{d}(\{j \in \mathbb{Z} : R^j_\alpha(0) \in [\beta, \beta') \}).\]
Since $R_\alpha$ is a uniquely ergodic transformation, we have
\[\bar{d}(\{j \in \mathbb{Z} : R^j_\alpha(0) \in [\beta, \beta') \}) = \lambda([\beta, \beta')) = \beta' - \beta.\]
It follows that the map $\hat{\Phi} : [0, 1] \ni \beta \mapsto x \ast y_{\alpha, \beta} \in \Omega$ is $\bar{d}_\Omega$-continuous, because
\[\bar{d}_\Omega(x \ast y_{\alpha, \beta}, x \ast y_{\alpha, \beta'}) \leq \bar{d}_\Omega(y_{\alpha, \beta}, y_{\alpha, \beta'}) = |\beta - \beta'|.\]
Furthermore since $0$ is a safe symbol for $X$ we have $\Phi(X) \subset X$. Hence, by (11), the map $\hat{\Phi} : [0, 1] \ni \beta \mapsto \mu_\beta = \mu \circ \Phi(\beta) \in M^e_\sigma(\Omega) \subset M^e_\sigma(X)$ is also $\bar{d}_M$-continuous and establishes a path from $\hat{\Phi}(0) = \mu_0 = \delta_0$ to $\hat{\Phi}(1) = \mu(x) = \mu$ in $M^e_\sigma(X)$. \(\square\)

By [32] Thm. 7.7] the topology of $\bar{d}_M$ metric is stronger than the weak* topology on $M^e_\sigma(\Omega)$. Together with Theorem 6 it yields the following result:

**Corollary 7.** If $X$ is a shift space with a safe symbol (in particular, if $X$ is a hereditary shift), then for any $t \geq 0$ the set $M^e_\sigma(X)$ (in particular, $M^e_\sigma(X)$) is arcwise connected in the weak* topology.

As the entropy function on $M^e_\sigma(\Omega)$ endowed with the $\bar{d}_M$ metric is continuous (see [32] Theorem, 7.9 or [34] Theorem, I.9.16)) we conclude also the following:

**Corollary 8.** If $X$ is a shift space with a safe symbol (in particular, if $X$ is a hereditary shift), then $\{h(\mu) : \mu \in M^e_\sigma(X)\} = [0, h_{top}(X)]$ (possibly degenerate to a point).

If $X$ is a hereditary shift, then the bi-infinite sequence of 0’s denoted by 0 is a fixed point for the shift map and belongs to $X$. Hence the atomic measure $\delta_0$ carried by 0 is invariant for $X$. There are hereditary shifts for which $\delta_0$ is the only invariant measure and the existence of another invariant measure has many consequences (see [25] for more details). Thus we divide all hereditary shifts into two disjoint classes:

I. uniquely ergodic hereditary shifts ($\delta_0$ is the unique invariant measure),

II. non-uniquely ergodic hereditary shifts.

Hereditary shift spaces in class (1) are characterized as those in which for every point $x \in X$ the symbols other than 0 appear in $x$ on a set of coordinates of zero asymptotic density. Class (1) coincides also with the hereditary shifts having zero topological entropy. Although this is not stated explicitly in [25], the proof applies verbatim to shift spaces with a safe symbol. Hence we may note the following corollary.

**Corollary 9.** A shift space with a safe symbol (in particular, a hereditary shift) has positive topological entropy if and only if it has uncountably many ergodic measures.

4. EXAMPLES OF HEREDITY SHIFTS

Here we list some notable examples of hereditary shifts to which our main result can be applied.

The primary example of a $B$-free shift is the square-free shift considered by Sarnak [33], that is, a shift space whose structure reflects the statistical properties
of square-free numbers. Recall that $n \in \mathbb{N}$ is square-free if there is no prime number $p$ such that $p^2$ divides $n$. Let $\eta$ be the characteristic function of the square-free numbers treated as a point in $\{0,1\}^\mathbb{Z}$. The square-free shift is the closure of the orbit of $\eta$ with respect to the shift map, and it turns out that it is hereditary \[27,33\]. The study of the square-free shift has been recently extended \[1,5\] to the general $\mathcal{B}$-free shifts $X_b$ induced in the same manner by the characteristic function $b$ of $\mathcal{B}$-free numbers, that is, integers with no factor in a given set $\mathcal{B} \subset \mathbb{N}$. If $\mathcal{B}$ is an Erdős set, that is, it consists of pairwise co-prime integers and $\sum_{b \in \mathcal{B}} 1/b < \infty$, then the $\mathcal{B}$-free shift is also hereditary; see \[1\]. In general this is not the case, but the smallest hereditary shift containing $X_b$ still has some interesting properties (see \[5\]). These systems were also investigated by Avdeeva \[2\], Cellarosi and Sinai \[6\], Kulaga-Przymus, Lemańczyk, and Weiss \[21,22\], and Peckner \[27\].

Another shift space related to $\mathcal{B}$-free integers is the $\mathcal{B}$-admissible shift \[5\]. We say that a sequence $x = (x_j)_{j \in \mathbb{Z}} \in \{0,1\}^\mathbb{Z}$ is $\mathcal{B}$-admissible if for every $b \in \mathcal{B}$ the set $\{j \in \mathbb{Z} : x_j = 1\}$ is disjoint with a set $b\mathbb{Z} + r$ for some $0 \leq r < b$. It is not hard to see that the set of $\mathcal{B}$-admissible sequences in $\Omega$ is a hereditary shift space $X_{\mathcal{B}}$ called the $\mathcal{B}$-admissible shift. Because $b$ is clearly a $\mathcal{B}$-admissible sequence we see immediately that $X_b \subset X_{\mathcal{B}}$ and the equality holds if $\mathcal{B}$ is an Erdős set.

Beta shifts introduced by Rényi \[30\] are related to number theory, tilings, and dynamics of discontinuous transformations. For $\beta > 1$ the beta shift $\Omega_\beta$ is the closure of the set of sequences in $\{0,1,\ldots, \lfloor \beta \rfloor\}^\mathbb{N}$ arising as greedy $\beta$-expansions of numbers from $[0,1]$. All beta shifts are hereditary (see \[25\]).

Spacing shifts were introduced by Lau and Zame in \[23\]. A spacing shift $\Omega_P$, where $P \subset \mathbb{N}$, is the set of all $x = (x_i) \in \{0,1\}^\mathbb{Z}$ such that $x_i = x_j = 1$ and $i \neq j$ imply $|i - j| \in P$. Spacing shifts were studied in \[3,4\]. It is easy to see that they are hereditary.

Bounded density shifts were recently introduced by Stanley \[37\]. They are defined by fixing a function $f : \mathbb{N} \to [0,\infty)$ and considering the set of all bi-infinite sequences such that for each $p \in \mathbb{N}$ the sum of the entries of any finite subword of length $p$ does not exceed $f(p)$. Since for any word coordinate-wise smaller than a given word the sum of the entries can only decrease, these shifts are hereditary.

5. Intermediate Entropy Property vs. Arcwise Connectedness

The following results show that the conclusions of Corollary \[7\] and Corollary \[8\] are independent of each other.\footnote{We are grateful to Tomasz Downarowicz for drawing our attention to the theory presented in \[11\] and generously sharing his insight on these matters.}

Recall that for every dynamical system $(X,T)$ the set $\mathcal{M}_T(X)$ endowed with the weak$^*$ topology has the structure of a Choquet simplex (see \[11\]). We say that a non-empty metrizable compact subset $K$ of a locally convex topological vector space is a Choquet simplex if every point of $K$ is the barycenter of a unique probability measure supported on the set $\text{ext} K$ of extreme points of $K$.

Lemma 10. Let $K$ be a convex subset of a vector space (over $\mathbb{R}$ or $\mathbb{C}$). If $P$ is a closed subset of the set $\text{ext} K$ of extreme points of $K$, then its characteristic function $\chi_P$ is upper semi-continuous and convex; that is, for $x,y \in K$ and $0 < \alpha < 1$ we have $\chi_P(\alpha x + (1-\alpha)y) \leq \alpha \chi_P(x) + (1-\alpha)\chi_P(y)$.\footnote{We are grateful to Tomasz Downarowicz for drawing our attention to the theory presented in \[11\] and generously sharing his insight on these matters.}
Proof. The upper semi-continuity is obvious. For the proof of convexity, note that if \( x, y \in K \), \( x \neq y \), and \( 0 < \alpha < 1 \), then \( \alpha x + (1 - \alpha)y \) is not an extreme point, so \( \chi_P(\alpha x + (1 - \alpha)y) = 0 \).

**Theorem 11.** For every Polish topological space \( P \) there exists a minimal shift space \( X \) (a Toeplitz shift) such that \( M_p^e(X) \) with the weak* topology and \( P \) are homeomorphic and there is a unique measure \( \mu \in M_p^e(X) \) with positive metric entropy. (In particular, if \( P \) has more than one point, then \( X \) does not have the intermediate entropy property.)

Proof. Let \( K_P \) be the Choquet simplex whose set of extreme points \( \text{ext} \, K_P \) is homeomorphic to \( P \) (such a Choquet simplex exists by [17]). Fix an extreme point \( z \in \text{ext} \, K_P \). Let \( \delta_z \) denote the Dirac measure concentrated on \( \{z\} \) and let \( \chi_{\{z\}} \) denote the characteristic function of \( \{z\} \subset \text{ext} \, K_P \). Given a point \( x \in K_P \), we denote by \( \xi^x \) the unique probability measure concentrated on \( \text{ext} \, K_P \) such that \( x \) is the barycenter of \( \xi^x \); that is, \( x \) is equal to the Pettis integral of the identity with respect to \( \xi^x \):

\[
x = \int_{\text{ext} \, K_P} y \, d\xi^x(y).
\]

(See Appendix in [11] for more details.) We define a function \( \varphi : K_P \to \mathbb{R} \) by

\[
\varphi(x) = \int_{\text{ext} \, K_P} \chi_{\{z\}}(y) \, d\xi^x(y) = \xi^x(\{z\}).
\]

This function is the harmonic prolongation of \( \chi_{\{z\}} \) [11 Definition A.2.18]. The characteristic function of a closed set is upper semi-continuous and convex by Lemma [10]. By [11] Fact A.2.10 every upper semi-continuous harmonic function on a Choquet simplex is affine. Therefore, \( \varphi \) is bounded, affine, non-negative, upper semi-continuous on \( K_P \), and \( \varphi|_{\text{ext} \, K_P} = \chi_{\{z\}} \).

By [10] Theorem 1, there exists a minimal Toeplitz shift \((X, \sigma)\) and an affine (onto) homeomorphism \( \psi : K_P \to M_e(X) \) such that for every \( x \in K_P \), \( \varphi(x) = h(\psi(x)) \), where \( h \) denotes the entropy function. This proves the proposition with \( \mu = \varphi(z) \).

**Theorem 12.** For every uncountable Polish topological space \( P \) there exists a minimal shift space \( X \) (a Toeplitz shift) such that \( M_p^e(X) \) with the weak* topology and \( P \) are homeomorphic and the metric entropy function \( h \) restricted to \( M_p^e(X) \) is not constant and has the Darboux (intermediate value) property.

Proof. Assume that \( P \) is uncountable. Then there is \( C \subset P \) homeomorphic to the usual Cantor set in \([0, 1]\) [36 Theorem 3.2.7]. Let \( g : P \to [0, 1] \) be any continuous function such that \( g|_C \) maps \( C \) onto \([0, 1]\) (say, use the Tietze extension theorem to extend the Devil’s staircase function on \( C \) to the whole \( P \) ). Then take \( \eta = g \cdot \chi_C \) where \( \chi_C \) is the characteristic function of \( C \) in \( P \). The function \( \eta : P \to [0, 1] \) is upper semi-continuous and has the Darboux (intermediate-value) property. Reasoning as above, we can find a Toeplitz minimal shift \((X, T)\) such that (up to affine homeomorphism) \( M_T(X) = K_P \) and \( \eta = h|_{M_T^e(X)} \).

6. Final remarks

We find the following question intriguing: Let \( \mu \) denote the Möbius function extended to \( \mathbb{Z} \) in an obvious way (say \( \mu(0) = 0 \) and \( \mu(-n) = \mu(n) \) for \( n \in \mathbb{N} \)). Let
Let $X_M$ be the orbit closure of $\mu$ in $\{-1, 0, 1\}^\mathbb{Z}$. Is 0 the safe symbol for $X_M$? Recall that the study of the square-free system in [33] was motivated by questions about Möbius function $\mu$.

Some special cases of our results were known to hold for some time. It follows from the general theory of Choquet simplices that if the set of ergodic measures is dense in the simplex of all invariant measures in the weak* topology, then it is an arcwise connected set [22]. Therefore the set of ergodic measures is arcwise connected in the weak* topology for all hereditary $\mathcal{B}$-free shifts $X_b$, as $\mathcal{M}_\sigma^0(X_b) = \mathcal{M}_\sigma(X_b)$ by [5][22]. For a concrete example of a shift space with dense set of ergodic measures but without the intermediate entropy property see [15]. Kulaga-Przymus, Lemańczyk, and Weiss showed also in [22] that ergodic measures need not be dense among all invariant measures for a general hereditary system, but must be arcwise connected by our result. This phenomenon was previously observed in the Dyck shift by Climenhaga [7].

All the results stated in this work remain valid for unilateral shift spaces, that is, closed $\sigma$-invariant subsets of $\Lambda^\mathbb{N}$. This follows directly from [8, Proposition 2.1].

After we finished writing this paper Ayşe Şahin kindly shared with us the article [28], where $d_{13}$ continuous arcs of ergodic measures are constructed for some examples of $\mathbb{Z}$ and $\mathbb{Z}^2$ shift spaces. Anthony Quas observed that one can use the construction from [29] (loosely related to the “grand coupling” that shows up in probability theory) to get another proof of Theorem 6 for the case $t = h_{\text{top}}(X)$, that is, for $\mathcal{M}_\sigma^0(X)$. This approach does not work for smaller $t$. The reader interested in the result about the measures only can also use the idea from [28] and coupling with irrational rotations to simplify the proof of Theorem 6 at the cost of losing the explicit construction of generic points. The latter result is in the spirit of “single orbit dynamics” (see remarks about differences between almost everywhere results and single orbit theorems in [41, pp. 8 and 89]), and we hope the reader will also find it interesting. Besides, it gives some information about topological properties of the Besicovitch space of the shift (see the proof of Theorem 6). We are grateful to Anthony Quas for sharing his idea with us.

**Added in proof.** We note that by Corollary 7, Theorem A from Shinoda’s paper [35] applies to any shift space with a safe symbol. This yields the following corollary:

**Corollary 13.** If $X$ is a non-trivial shift space with a safe symbol, then there exists a dense set of continuous functions $\mathcal{D} \subset \mathcal{C}(X)$ such that for every $f \in \mathcal{D}$ there exist uncountably many ergodic maximizing measures; that is, the set 

$$\mathcal{M}_{\text{max}}(f) = \{\mu \in \mathcal{M}_\sigma(X) : \int f \, d\mu = \max_{\nu \in \mathcal{M}_\sigma(X)} \int f \, d\nu\}$$

is uncountable.

**Appendix A. Uniquely ergodic extensions**

Originally, we proved Theorem 6 using tools developed in this section. After a talk one of us presented of these results during a seminar at the Institute of Mathematics of the Polish Academy of Sciences, Professors Lemańczyk and Przytycki kindly suggested another approach for this proof. As we feel that the following observations might be of independent interest and allow us to show Theorem 6 avoiding spectral theory of unitary operators we decided to attach them here.
Suppose that \((X,T)\) is a topological factor of \((Y,S)\) and \(\pi: Y \to X\) is a factor map. Let \(\mu \in \mathcal{M}_T^p(X)\). We say that \((Y,T)\) is a uniquely ergodic extension of \((X,T)\) over \(\mu\) if there is a unique measure \(\nu \in \mathcal{M}_S^p(Y)\) such that \(\pi_*\nu = \nu\). We also say that \((Y,S,\nu)\) uniquely extends \((X,T,\mu)\) through \(\pi: Y \to X\). The following lemma and its corollary explain our interest in uniquely ergodic extensions.

**Lemma 14** (Weiss, [11], Proposition 3.4). If the ergodic system \((Y,S,\nu)\) uniquely extends \((X,T,\mu)\) over \(\pi: Y \to X\), then \(y_0 \in \text{Gen}_T(\mu)\), and \(y_0 \in \pi^{-1}\{\{x_0\}\}\), then \(y_0 \in \text{Gen}_S(\nu)\).

The lemma above implies a useful criterion for genericity of a pair of generic points with respect to a product measure.

**Corollary 15.** Let \((X,T,\mu)\) and \((Y,S,\nu)\) be two ergodic systems with ergodic product. If \((X \times Y,T \times S,\mu \times \nu)\) uniquely extends \((X,T,\mu)\) over the projection onto the first coordinate \(\pi: X \times Y \to X\), then for every \(x_0 \in \text{Gen}_T(\mu)\) and \(y_0 \in Y\) the pair \((x_0,y_0)\) is generic for \(\mu \times \nu\).

Note that in the situation of the above corollary, \((Y,S,\nu)\) is necessarily uniquely ergodic, so \(\text{Gen}_S(\nu) = Y\). Lemma 14 was used in place of Lemma 5 in the first version of our proof of Theorem 6.

The following relative version of the unique ergodicity theorem of Furstenberg can be used to show that certain extensions are uniquely ergodic over group rotations. The proof can be obtained through an easy modification of the standard argument, which can be found in e.g. [12, Theorem 4.21]; we include it here for the convenience of the reader.

**Theorem 16.** Assume that \(T: X \to X\) is a homeomorphism of a compact metric space and \(\mu \in \mathcal{M}_T^p(X)\). Let \(G\) be a compact group, let \(\lambda_G\) be the Haar measure for \(G\), and let \(\varphi: X \to G\) be a continuous map. Define \(Y = X \times G\) and \(S: Y \to Y\) by \(S(x,g) = (T(x),\varphi(x)g)\). If \(S\) is ergodic with respect to \(\mu \times \lambda_G\), then \((Y,S,\mu \times \lambda_G)\) uniquely extends \((X,T,\mu)\) over the projection \(\pi: Y \ni (x,g) \mapsto x \in X\).

**Proof.** The \(S\)-invariance of \(\mu \times \lambda_G\) is an immediate consequence of Fubini’s theorem, since for any \(f \in \mathcal{C}(Y)\) we have

\[
\int_Y f \circ S \, d(\mu \times \lambda_G) = \int_X \int_G f(Tx,\varphi(x)g) \, d\lambda_G(g) \, d\mu(x) = \int_G \int_X f(Tx,g) \, d\mu(x) \, d\lambda_G(g) = \int_Y f \, d(\mu \times \lambda_G).
\]

Assume that \((Y,S,\mu \times \lambda_G)\) is an ergodic measure-preserving system. We claim that for every \(h \in G\) the set \(\text{Gen}_S(\mu \times \lambda_G)\) is invariant under the map \(R_h: Y \ni (x,g) \mapsto (x,gh) \in Y\). Note that \((x,g) \in \text{Gen}_S(\mu \times \lambda_G)\) if and only if for every continuous function \(f: Y \to \mathbb{R}\) we have

\[
(2) \quad \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(S^n(x,g)) = \int_Y f \, d(\mu \times \lambda_G).
\]

We fix \(f \in \mathcal{C}(Y)\) and \(h \in G\). We want to show that (2) holds with \((x,gh)\) in place of \((x,g)\). First note that

\[
(3) \quad \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(S^n(x,gh)) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f \circ R_h(S^n(x,g)).
\]
Now use the fact that \((x, g)\) is generic to conclude that the right hand side of (x) converges to
\[
\int_Y f \circ R_h d(\mu \times \lambda_G) = \int_Y f d((R_h)_*(\mu \times \lambda_G)) = \int_Y f d(\mu \times \lambda_G),
\]
where the last equality holds because \(\mu \times \lambda_G\) is \(R_h\) invariant. This shows that for all \(h \in G\) we have \(R_h(Gen_S(\mu \times \lambda_G)) \subset Gen_S(\mu \times \lambda_G)\). Repeating the same argument with \(h^{-1}\) in place of \(h\) yields the reverse inclusion, and hence for any \(h \in G\) we have
\[
R_h(Gen_S(\mu \times \lambda_G)) = Gen_S(\mu \times \lambda_G).
\]
This completes the proof of our claim.

Let \(E_1 \subset X\) be the projection of \(Gen_S(\mu \times \lambda_G)\) onto the first coordinate. Clearly, \(\mu(E_1) = 1\). If \((x, g) \in Gen_S(\mu \times \lambda_G)\) and \(g' \in G\), then taking \(h = g^{-1}g'\) and using our claim we get that \((x, gh) = (x, g') \in Gen_S(\mu \times \lambda_G)\). This shows that \(Gen_S(\mu \times \lambda_G) = E_1 \times G\).

Let \(\nu\) be an ergodic \(S\)-invariant measure such that \(\pi_*(\nu) = \mu\). Let \(E_2 = \pi(Gen_S(\nu))\). Then \(\mu(E_2) = 1\); hence \(\mu(E_1 \cap E_2) = 1\), and in particular there exists \(x_0 \in E_1 \cap E_2\). Let \(y_0 \in Gen_S(\nu)\) be such that \(\pi(y_0) = x_0\). But then \(y_0 \in \{x_0\} \times G \subset Gen_S(\mu \times \lambda_G)\), and \(y_0\) is generic for both \(\nu\) and \(\mu \times \lambda_G\), implying that \(\nu = \mu \times \lambda_G\).

\(\square\)

Theorem 16 yields the following corollary, which may replace Theorem 4 in the proof of our main result.

**Corollary 17.** If \(T: X \to X\) is a homeomorphism of a compact metric space and \(\mu \in \mathcal{M}_0^e(X)\), then there exists \(\alpha \in \mathbb{R} \setminus \mathbb{Q}\) such that the product system \((T \times R_\alpha, X \times \mathbb{T}^1, \mu \times \lambda)\) is a uniquely ergodic extension of \((X, T, \mu)\) over the projection \(\pi: X \times \mathbb{T}^1 \to X\), where \(R_\alpha\) is the rotation by \(\alpha\) and \(\lambda\) denotes the Lebesgue measure on \(\mathbb{T}^1\).

**Proof.** By Theorem 16, it is enough to find \(\alpha \in \mathbb{R} \setminus \mathbb{Q}\) such that \(T \times R_\alpha\) is ergodic with respect to \(\mu \times \lambda\). To this end, using spectral theory discussed in Section 2, it will suffice to ensure that the only common element of \(H(T, \mu)\) and \(H(R_\alpha, \lambda) = \{\exp(2\pi i k \alpha) : k \in \mathbb{Z}\}\) is 1. Because \(H(T, \mu)\) is at most countable, this will be the case for all but countably many choices of \(\alpha\).

\(\square\)

**References**


