Shift Spaces with Specification are Intrinsically Ergodic

Gregory Hemenway

University of Houston

Abstract

We give a self-contained account of sufficient conditions for a shift space (Σ, σ) to be intrinsically ergodic. In particular, we adapt uniqueness results by Bowen to the setting of symbolic dynamics. We intend for this document to be an introduction to topics in the area of symbolic dynamics and thermodynamic formalism.

Keywords: Symbolic Dynamics, Shift Spaces, Specification, Entropy, Thermodynamic Formalism

1. Introduction

For a compact metric space X, the variational principle for topological entropy states that if $f\colon X \to X$ is continuous, then $h_{\text{top}}(f) = \sup_\mu h_\mu(f)$, where h_{top} denotes the topological entropy and the supremum is taken over all f-invariant probability measures on X. A measure μ that achieves this supremum is called a ***measure of maximal entropy*** (MME). A system which has a unique MME is called ***intrinsically ergodic***.

In the 1960s, Parry proved that transitive subshifts of finite type are intrinsically ergodic. A presentation of this result is given by Sherman [1]. In 1974, Bowen showed that expansive systems satisfying the *specification property* have unique MMEs (*see [2]*). In particular, Bowen's arguement can be applied to *shift spaces*. In this paper, we will show that specification is a sufficient condition for a shift space to be intrinsically ergodic.

Theorem 1.1. *If A is a finite alphabet and $\Sigma \subset A^{\mathbb{N}}$ is a shift space with specification, then Σ has a unique measure of maximal entropy.*

In particular, we will adapt Bowen's argument for expansive homeomorphisms with specification to prove theorem 1.1. Since for any distint $x, y \in \Sigma$, there is a minimal $n \in \mathbb{N}$ such that $x_n \neq y_n$. Then, as we'll see in section 2.1, $d(\sigma^{n-1}(x), \sigma^{n-1}(y)) = \frac{1}{2}$. Therefore, the symbolic structure of a shift space is indeed analogous to the expansive property used by Bowen and there is no potential function to consider. The proof of theorem 1.1 is presented in five parts and follows the structure of Bowen's proof.

1. Show that there is a $Q_1 > 0$ such that $e^{nh} \leq \#\mathcal{L}_n \leq Q_1 e^{nh}$ for every $n \in \mathbb{N}$.
2. Explicitly construct a MME μ using the variational principle.
3. Use specification to show that μ satisfies a Gibbs property.
4. Show that μ is ergodic.
5. Show that the existence of another ergodic MME ν contradicts the Gibbs property for μ. Thereby, showing the uniqueness of μ.

For the proofs of the lemmas to be considered, we will follow Climenhaga and Thompson's paper [3] and a blog post by Climenhaga [4] since these are written in the setting of symbolic dynamics albeit for a broader class of systems.

2. Definitions and Results

2.1. Symbolic Dynamics

We provide the necessary background in symbolic dynamics that a reader will need to understand the material discussed in this paper. For a more complete presentation of this topic, we refer the reader to Kitchens [5] and Lind [6].

Let $A^{\mathbb{N}}$ denote the set of all infinite sequences over a finite alphabet $A = \{1, ..., p\}$; i.e. $A^{\mathbb{N}} = \{x_1 x_2 x_3 \cdots \mid x_i \in A, i \in \mathbb{N}\}$. This is a compact metric space with distance function $d(x,y) = 2^{-\min\{\mathrm{n:\ x_n \neq y_n}\}}$. Let σ denote the shift map on $A^{\mathbb{N}}$; i.e. for $x \in A^{\mathbb{N}}$, $\sigma^i(x) = x_{i+1}$. A ***shift space*** is a closed σ-invariant subset $\Sigma \subset A^{\mathbb{N}}$. Since Σ is closed, it is compact. Given $n \in \mathbb{N}$, consider the set $\mathcal{L}_n := \{w \in A^n \mid [w] \neq \emptyset\}$, where $[w]$ denotes the set of all $x \in \Sigma$ starting with w. That is, $[w] = \{x \in X \mid x_1 \cdots x_n = w\}$. We call $[w]$ the ***cylinder set*** for the ***word*** $w \in \mathcal{L}_n$. The union $\mathcal{L} := \bigcup_{n \in \mathbb{N}} \mathcal{L}_n$ is called the ***language*** of the shift space Σ. We will denote the collection of n-cylinders by $\mathcal{U}_n$ and the collection of all cylinder sets by $\mathcal{U}$.

A ***topological dynamical system*** is a pair (X,T) where X is a compact metric space and T is a continuous mapping of X to itself. Therefore, (Σ, σ)

is a topological dynamical system since the shift map is continuous. Since each position in an element of Σ is discretely chosen, it is obvious that Σ inherits the discrete-product topology. It is also easy to verify that $\mathcal{U}$ is a basis for the topology on Σ. Also since a cylinder set is the complement of a union of cylinder sets, cylinder sets are both open and closed. This fact will be useful later.

In particular, we will consider shift spaces with a certain transitivity property. For symbolic dynamics, the *topological transitivity property* takes the following form: Σ is transitive if and only if for every pair of words $u, v \in \mathcal{L}$, there exists a word $w \in \mathcal{L}$ such that $uwv \in \mathcal{L}$. A shift space is said to have ***specification*** if the gluing word w in the transitivity property can be chosen to have a fixed length for all u and v. That is, a shift space satisfies the specification property if there exists $\tau \in \mathbb{N}$ such that for every $u, v \in \mathcal{L}$, there exists a word $w \in \mathcal{L}_\tau$ such that $uwv \in \mathcal{L}$.

2.2. Topological Entropy

In this section and the next, we introduce entropy. As we cannot include every detail on the subject, we refer the reader to chapters 4 and 7 in [7] for further background.

Let X be a compact topological space. The ***join*** of open covers α and β of X is the set $\alpha \vee \beta = \{A \cap B \colon A \in \alpha, B \in \beta\}$. For a finite collection of open covers $\{\alpha_1, ..., \alpha_n\}$, we denote their join by $\bigvee_{k=1}^{n} \alpha_k$. Note that the preimage of an open cover under a continuous tranformation $T \colon X \to X$ is also an open cover. In particular, we are interested in joins of the form $\bigvee_{k=0}^{n-1} T^{-k}\alpha = \alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-(n-1)}\alpha$. Observe that $\mathcal{U}_1$ is an open cover of Σ and $\mathcal{U}_n = \bigvee_{k=0}^{n-1} \sigma^{-k}\mathcal{U}_1$. We will see in a moment that the collections of n-cylinders are part of an important class of open covers.

Let α be an open cover of X. Since X is compact, there exists a finite subcover of α. Let $N(\alpha)$ be the smallest cardinality of subcover of α. The ***entropy of*** α is defined by $H(\alpha) = \log N(\alpha)$. The ***entropy of*** T ***relative to*** α is defined by

$$h(T, \alpha) = \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{k=0}^{n-1} T^{-k}\alpha\Big),$$

where $T^{-k}\alpha = \{T^{-k}A\}_{A\in\alpha}$. Finally, the ***topological entropy,*** $h_{\text{top}}(T)$***,*** of a continuous transformation T is defined to be the supremum of $h(T, \alpha)$ over all open covers α of X. That is, $h_{\text{top}}(T) = \sup\{h(T, \alpha) |\, \alpha \text{ is an open cover of } X\}$.

Throughout this paper, topological entropy will be denoted as h if the context is clear.

An advantage to working in the setting of symbolic dynamics is the fact that the covers $\mathcal{U}_n$ are *generators* of σ. A ***generator*** of a continuous transformation of X, T, is an open cover α of X in which for any sequence of sets $\{A_n\}_{n=0}^{\infty}$ in α, $\bigcap_{n=0}^{\infty} T^{-n}\overline{A_n}$ contains exactly one point of X. Indeed, for any n-cylinder $[w]$, $\bigcap_{n=0}^{\infty} \sigma^{-n}[w]$ is a single word of Σ. Proposition 2.3 below will show that the topological entropy of the shift map σ is equal to $h(\sigma,\mathcal{U}_1)$, but we will first state two results that we'll need for the proof.

Proposition 2.1. *If (X, d) is a compact metric space and α is an open cover of X, then there exists $\delta > 0$ such that each subset of X of diameter less than or equal to δ lies in some member of α. This proposition is called Lebesgue's covering lemma and the constant δ is called a Lebesgue number of α*

Proof. Note that since X is compact, it suffices to consider a finite subcover of α. Let $\alpha_0 = \{A_1, ..., A_k\}$ where $A_j \in \alpha$ for $1 \leq j \leq k$. Assume the result is false. For each $n \geq 1$, there exists $B_n \subseteq X$ such that $\operatorname{diam}(B_n) \leq \frac{1}{n}$ and B_n is not contained in any A_j of α_0. Choose $x_n \in B_n$ and then a convergent subsequence $\{x_{n_i}\}$. Suppose $x \in A_j \in \alpha$ where $x_{n_i} \to x$. Let $a = d(x, X \backslash A_j) > 0$. Choose n_i such that $n_i > 2/a$ and $d(x_{n_i}, x) < a/2$. Then if $y \in B_{n_i}$,

$$d(y, x) \leq d(y, x_{n_i}) + d(x_{n_i}, x) \leq \frac{1}{n_i} + \frac{a}{2} < a.$$

Hence $B_{n_i} \subseteq A_j$, a contradiction. □

We say that α ***refines*** β, denoted $\beta < \alpha$, if for each $A \in \alpha$, there exists a $B \in \beta$ such that $A \subset B$.

Proposition 2.2. *If $\alpha > \beta$, then $h(T, \alpha) \geq h(T, \beta)$.*

Proof. For any subcover $\alpha_0 = \{A_1, ..., A_n\}$ of α, each A_i is contained in some $B_i \in \beta$. Note that $\beta_0 = \{B_1, ..., B_n\}$ is also an open cover of X of cardinality n. Thus, $N(\alpha) \geq N(\beta)$ and $H(\alpha) \geq H(\beta)$. Now suppose $A = \bigcap_{i=0}^{n-1} T^{-i}A_i \in \bigvee_{i=0}^{n-1} T^{-i}\alpha$. Since α refines β, for each i $A_i \subset B_i$ for some $B_i \in \beta$. So if $B = \bigcap_{i=0}^{n-1} T^{-i}B_i \in \bigvee_{i=0}^{n-1} T^{-i}\beta$, then $A \subset B$. Thus, $\bigvee_{i=0}^{n-1} T^{-i}\alpha < \bigvee_{i=0}^{n-1} T^{-i}\beta$ implying $H(\bigvee_{i=0}^{n-1} T^{-i}\alpha) \geq H(\bigvee_{i=0}^{n-1} T^{-i}\beta)$ and $h(T, \alpha) \geq h(T, \beta)$. □

Proposition 2.3. *Consider the generator $\mathcal{U}_1$ of our dynamical system (Σ, σ). Then $h_{\text{top}}(\sigma) = h(\sigma,\mathcal{U}_1)$.*

Proof. Let β be an open cover with a Lebesgue number δ. We can choose $N \in \mathbb{N}$ such that $\operatorname{diam}(\bigvee_{i=0}^{N-1} \sigma^{-i}\mathcal{U}_1) < \delta$. Then $\beta < \bigvee_{i=0}^{N-1} \sigma^{-i}\mathcal{U}_1$ and

$$\begin{aligned} h(\sigma, \beta) &\leq h(\sigma, \bigvee_{i=0}^{N-1} \sigma^{-i}\mathcal{U}_1) \\ &= \lim_{k\to\infty} \frac{1}{k} H\Big(\bigvee_{l=0}^{k-1} \sigma^{-l} \Big(\bigvee_{i=0}^{N-1} \sigma^{-i}\mathcal{U}_1 \Big)\Big) \\ &= \lim_{k\to\infty} \frac{1}{k} H\Big(\bigvee_{i=0}^{k+N-1} \sigma^{-i}\mathcal{U}_1 \Big) \\ &= \lim_{k\to\infty} \frac{k+N-1}{k} \cdot \frac{1}{k+N-1} H\Big(\bigvee_{i=0}^{k+N-1} \sigma^{-i}\mathcal{U}_1 \Big) \\ &= h(\sigma, \mathcal{U}_1). \end{aligned}$$

Thus, $h(\sigma, \beta) \leq h(\sigma, \mathcal{U}_1)$ for all open covers β and so $h_{\text{top}}(\sigma) = h(\sigma, \mathcal{U}_1)$. □

Thus, for a shift space, we need only consider the open cover $\mathcal{U}_1$ to calculate the topological entropy of σ. That is, the topological entropy of σ is given by

$$h_{\text{top}}(\sigma) = h(\sigma, \mathcal{U}_1) = \lim_{n\to\infty} \frac{1}{n} \log \#\mathcal{L}_n.$$

2.3. Measure-Theoretic Entropy

The build up of measure-theoretic entropy is very similar to that of topological entropy. Consider a probability space $(X, \mathcal{B}, m)$ where $\mathcal{B}$ denotes the Borel σ-algebra. The definition of the *join* of a collection of partitions of X is analogous to that of a collection of open covers. Let $\zeta = \{A_1, ..., A_k\}$ be a finite partition of X. We define the ***entropy of*** ζ to be

$$H_m(\zeta) = -\sum_{i=1}^{k} m(A_i) \log m(A_i).$$

Now suppose $T\colon X \to X$ is a measure-preserving transformation of (X, m). We call

$$h_m(T, \zeta) = \lim_{n\to\infty} \frac{1}{n} H_m\Big(\bigvee_{k=0}^{n-1} T^{-k}\zeta \Big)$$

the ***entropy of T with respect to*** ζ. Finally, the ***entropy of T*** is defined to be $h_m(T) = \sup\{h_m(T, \zeta) |\ \zeta \text{ is a finite partition of } X\}$.

2.4. Useful Results

The following results will be used numerous times throughout this paper. The first we state without proof since its details require more structured results than which are relevant to the purpose of this paper. For details, see theorem 4.3 in Walters.

Proposition 2.4. *For finite partitions ζ and η, $H_m(\zeta \vee \eta) \leq H_m(\zeta) + H_m(\eta)$.*

Proposition 2.5. *If $\{a_n\}_{n\geq 1}$ is a sequence of real numbers such that*

$$a_{n+p} \leq a_n + a_p \text{ for all } n \in \mathbb{N},$$

then $\lim_{n\to\infty} \frac{a_n}{n}$ exists and is equal to $\inf_n \frac{a_n}{n}$.

Proof. Fix a natural number p. We can write each $n \in \mathbb{N}$ as $n = kp + i$ where $0 \leq i < p$. Then

$$\frac{a_n}{n} = \frac{a_{i+kp}}{i+kp} \leq \frac{a_i}{kp} + \frac{a_{kp}}{kp} \leq \frac{a_i}{kp} + \frac{ka_p}{kp} = \frac{a_i}{kp} + \frac{a_i}{p}.$$

As $n \to \infty$, k must also approach infinity which implis that $\overline{\lim} \frac{a_n}{n} \leq \frac{a_p}{p}$. Since p was arbitrary, $\overline{\lim} \frac{a_n}{n} \leq \inf \frac{a_p}{p}$. But $\inf \frac{a_p}{p} \leq \underline{\lim} \frac{a_n}{n}$ so that $\lim \frac{a_n}{n}$ exists and equals $\inf \frac{a_n}{n}$. □

We use these two propositions to prove a useful result about the entropy of T.

Proposition 2.6. *Let $\zeta = \{A_1, ..., A_k\}$ be a finite partition of Σ. The measure-theoretic entropy*

$$H_m(T, \zeta) = \inf_{n\geq 1} \frac{1}{n} H_m\Big(\bigvee_{k=0}^{n-1} T^{-k}\zeta \Big).$$

Proof. Let $a_n = H_m\big(\bigvee_{k=0}^{n-1} T^{-k}\zeta\big)$. Note that $H_m(\zeta) = H_m(T^{-1}\zeta)$ since T is measure-preserving. Then by proposition 2.4, we see that

$$\begin{aligned} a_{n+p} = H_m\Big(\bigvee_{k=0}^{n+p-1} T^{-k}\zeta \Big) &\leq H_m\Big(\bigvee_{k=0}^{n-1} T^{-k}\zeta \Big) + H_m\Big(\bigvee_{k=n}^{n+p-1} T^{-k}\zeta \Big) \\ &= a_n + H_m\Big(\bigvee_{k=0}^{p-1} T^{-k}\zeta \Big) \\ &= a_n + a_p. \end{aligned}$$

Applying proposition 2.5 finishes the proof. □

We will need to estimate the entropy of ζ. Thus we will need to prove some properties for functions of the form $\phi(t) = -t \log t$. Note that ϕ is a concave function.

Proposition 2.7. *Consider a probability vector $\{p_1, ..., p_k\}$, i.e. $p_i \geq 0$ for $1 \leq i \leq k$ and $\sum_{i=1}^k p_i = 1$. Then $\sum_{i=1}^k -p_i \log p_i \leq \log k$.*

Proof.

$$\frac{1}{k} \sum_{i=1}^{k} -p_i \log p_i = \frac{1}{k} \sum_{i=1}^{k} \phi(p_i) \leq \phi\left(\frac{1}{k} \sum_{i=1}^{k} p_i\right) = \frac{1}{k} \log k,$$

where the inequality is a consequence of the concavity of $\phi(p_i)$. Note that equality holds only if $p_i = \frac{1}{k}$ for each $1 \leq i \leq k$. □

Proposition 2.8. *Let $\{p_1, ..., p_k\}$ be an arbitrary collection of nonnegative real numbers. Then there exists a probability vector $\{p_i', ..., p_k'\}$ such that*

$$\sum_{i=1}^{k} -p_i \log p_i = p \sum_{i=1}^{k} -p_i' \log p_i' - p \log p.$$

Proof. Let $p = \sum_{i=1}^k p_i$ and $p_i' = \frac{p_i}{p}$ for $1 \leq i \leq k$. Then $\sum_{i=1}^k p_i' = 1$ and $p_i = p_i' p$ so that

$$\sum_{i=1}^{k} -p_i \log p_i = \sum_{i=1}^{k} -p_i' p (\log p_i' + \log p) = p \sum_{i=1}^{k} -p_i' \log p_i' - p \log p.$$

□

Let X be a compact metric space and $T\colon X \to X$ be continuous. Denote by $M(X)$ the collection of all probability measures on $(X, \mathcal{B}(X))$ and $M(X,T)$ the collection of probability measures that make T a measure-preserving transformation. For background on this topic, see chapter 6 in Walters' book on ergodicity [7]. In particular, we will need a result about convergence in the weak*-topology on $M(X)$ for a space X. A sequence converges in the weak*-topology if $\int f \, d\mu_n = \int f \, d\mu$ for all $f \in C(X)$.

Proposition 2.9. *Let X be a compact metric space and $T\colon X \to X$ be continuous. Define a sequence $\{\mu_n\}_{n=1}^\infty$ by $\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} \sigma_n T^{-i}$, where $\{\sigma_n\}$ is a sequence in M(X). Then any limit point μ of $\{\mu_n\}$ is a member of $M(X,T)$.*

Proof. Let $\mu_{n_j} \to \mu$ in $M(X)$ and $f \in C(X)$. Then

$$\begin{aligned}
\left| \int f \circ T^{-1} \, d\mu - \int f \, d\mu \right| &= \lim_{j\to\infty} \left| \int f \circ T^{-1} \, d\mu_{n_j} - \int f \, d\mu_{n_j} \right| \\
&= \lim_{j\to\infty} \left| \frac{1}{n_j} \int \sum_{i=0}^{n_j-1} (f \circ T^{-(i+1)} \;\; - f \circ T^{-i}) \, d\sigma_{n_j} \right| \\
&= \lim_{j\to\infty} \left| \frac{1}{n_j} \int (f \circ T^{n_j} \;\; - f \;) \, d\sigma_{n_j} \right| \\
&\leq \lim_{j\to\infty} \frac{1}{n_j} \int \left| (f \circ T^{n_j} \;\; - f \;) \right| \, d\sigma_{n_j} \\
&\leq \lim_{j\to\infty} \frac{2\|f\|}{n_j} = 0.
\end{aligned}$$

Therefore, $\mu(X, T)$. □

Let $\triangle$ denote the symmetric difference of two sets; i.e. for sets A and B, $A \triangle B = (A \backslash B) \cup (B \backslash A)$.

Proposition 2.10. *Let $(X, \mathcal{B}, m)$ be a measure space. If $A, B \in \mathcal{B}$ have finite measure, then $|m(A) - m(B)| \leq m(A \triangle B)$.*

Proof. Note that for any two sets X and Y, $X = (X \backslash Y) \cup (X \cap Y)$ and vice-versa. Let $A, B \in \mathcal{B}$. Therefore,

$$\begin{aligned}
|m(A) - m(B)| &= |m(A \backslash B) + m(A \cap B) - m(B \backslash A) - m(A \cap B)| \\
&= |m(A \backslash B) - m(B \backslash A)| \\
&\leq m(A \backslash B) + m(B \backslash A) \\
&= m((A \backslash B) \cup (B \backslash A)) = m(A \triangle B).
\end{aligned}$$

□

3. Counting Estimates

In this section, we use the topological entropy and specification property of (Σ, σ) to provide bounds on $\#\mathcal{L}_n$, the cardinality of $\mathcal{L}_n$. Then we use the bound to also bound the the number of words of a given length in a subset $\mathcal{D} \subset \mathcal{L}$. These results will be used later to show that a MME of Σ has a Gibbs property and is unique.

Lemma 3.1. *There exists a constant $Q_1 > 0$ such that for every $n \in \mathbb{N}$*

$$e^{nh} \leq \#\mathcal{L}_n \leq Q_1 e^{nh}.$$

Proof. Denote by $\mathcal{L}_m\mathcal{L}_n$ the concatenation of $\mathcal{L}_m$ and $\mathcal{L}_n$; that is, $\mathcal{L}_m\mathcal{L}_n$ is the set of all words of the form vw where $v \in \mathcal{L}_m$ and $w \in \mathcal{L}_n$. If $u \in \mathcal{L}_{m+n}$, then $[u] = \{x \in X|\ u = x_1 \cdots x_{m+n}\}$. Let $v = x_1 \cdots x_m \in \mathcal{L}_m$ and $w = x_{m+1} \cdots x_{m+n} \in \mathcal{L}_n$. Thus, $u = vw \in \mathcal{L}_m\mathcal{L}_n$ so $\mathcal{L}_{m+n} \subset \mathcal{L}_m\mathcal{L}_n$. This implies that $\#\mathcal{L}_{m+n} \leq (\#\mathcal{L}_m)(\#\mathcal{L}_n)$ and $\log \#\mathcal{L}_{m+n} \leq \log \#\mathcal{L}_m + \log \#\mathcal{L}_n$. By proposition 2.5,

$$h(\sigma) = \lim_{n\to\infty} \frac{1}{n} \log \#\mathcal{L}_n = \inf_n \frac{1}{n} \log \#\mathcal{L}_n \implies h \leq \frac{1}{n} \log \#\mathcal{L}_n,$$

for all $n \in \mathbb{N}$. Therefore, $e^{nh} \leq \#\mathcal{L}_n$ for every natural n.

To establish the upper bound, define a map $(\mathcal{L}_n)^k \to \mathcal{L}_{k(n+\tau)}$ by

$$(w_1, ..., w_k) \mapsto w_1u_1 \cdots w_ku_k$$

where $u_i \in \mathcal{L}_\tau$, $1 \leq i \leq k$ is provided by specification. If $(x_1, ..., x_k)$ is another element of $(\mathcal{L}_n)^k$ such that $w_1u_1 \cdots w_ku_k = x_1v_1 \cdots x_kv_k$, then $w_i = x_i$ since $|u_i| = |v_i| = \tau$ for $1 \leq i \leq k$. Thus, γ is injective implying that $\#\mathcal{L}_{k(n+\tau)} \geq (\#\mathcal{L}_n)^k$. So for any n,

$$\frac{1}{k(n+\tau)} \log \#\mathcal{L}_{k(n+\tau)} \geq \frac{1}{n+\tau} \log \#\mathcal{L}_n.$$

Sending $k \to \infty$ gives $h \geq \frac{1}{n+\tau} \log \#\mathcal{L}_n \implies e^{h(n+\tau)} \geq \#\mathcal{L}_n$ providing the upper bound. Setting $Q_1 = e^{\tau h}$ gives the desired result. $\square$

Later we will need to bound the number of elements of a given length in a subset of the language of Σ. For a subset $\mathcal{D} \subset \mathcal{L}$, denote $\mathcal{D}_n = D \cap \mathcal{L}_n$ for each $n \in \mathbb{N}$. Given a measure ν, we write $\nu(\mathcal{D}_n)$ for $\nu(\bigcup_{w\in\mathcal{D}_n}[w])$ and $\nu(w)$ for $\nu([w])$.

Lemma 3.2. *For $\gamma \in (0,1)$, there exists $K_1 > 0$ such that if ν is a MME, $n \in \mathbb{N}$, and $\mathcal{D}_n \subset \mathcal{L}_n$ has $\nu(\mathcal{D}) \geq \gamma$, then*

$$\#\mathcal{D}_n \geq K_1 e^{nh}.$$

Proof. Recall that $h_\nu(\sigma) = \inf_{n\geq 1} \frac{1}{n} H_\nu\left(\bigvee_{k=0}^{n-1} T^{-k}\zeta\right)$. Since $h_\nu(\sigma) = h_{\text{top}}(\sigma)$,

$$\begin{aligned} nh_{\text{top}}(\sigma) &\leq \sum_{w\in\mathcal{L}_n} -\nu(w)\log\nu(w) \\ &= \sum_{w\in\mathcal{D}_n} -\nu(w)\log\nu(w) + \sum_{w\in\mathcal{D}_n^c} -\nu(w)\log\nu(w), \end{aligned}$$

where $\mathcal{D}_n^c$ denotes the complement of $\mathcal{D}_n$ in $\mathcal{L}_n$. Applying proposition 2.8 to both sums yields

$$\begin{aligned} nh_{\text{top}}(\sigma) \leq\ & \nu(\mathcal{D}_n)\left(\sum_{w\in\mathcal{D}_n} -\frac{\nu(w)}{\nu(\mathcal{D}_n)}\log\frac{\nu(w)}{\nu(\mathcal{D}_n)}\right) - \nu(\mathcal{D}_n)\log\nu(\mathcal{D}_n) \\ &+ \nu(\mathcal{D}_n^c)\left(\sum_{w\in\mathcal{D}_n^c} -\frac{\nu(w)}{\nu(\mathcal{D}_n^c)}\log\frac{\nu(w)}{\nu(\mathcal{D}_n^c)}\right) - \nu(\mathcal{D}_n^c)\log\nu(\mathcal{D}_n^c) \\ \leq\ & \nu(\mathcal{D}_n)\log\#\mathcal{D}_n + \nu(\mathcal{D}_n^c)\log\#\mathcal{D}_n^c + \log 2. \end{aligned}$$

By lemma 3.1, we see that $\#\mathcal{D}_n^c \leq \#\mathcal{L}_n \leq Q_1 e^{nh}$ which in turn implies that $\log\#\mathcal{D}_n^c \leq \log Q_1 + nh$ and

$$nh_{\text{top}}(\sigma) \leq \nu(\mathcal{D}_n)\log\#\mathcal{D}_n + (1-\nu(\mathcal{D}_n))(\log Q_1 + nh) + \log 2.$$

Therefore by using the assumption that $\nu(\mathcal{D}_n) \geq \gamma$, we see that

$$\begin{aligned} \nu(\mathcal{D}_n)\log\#\mathcal{D}_n &\geq \nu(\mathcal{D}_n)(\log Q_1 + nh) - \log(2Q_1) \\ \log\#\mathcal{D}_n &\geq \log Q_1 + nh - \frac{\log(2Q_1)}{\nu(\mathcal{D}_n)} \\ &\geq \log Q_1 + nh - \gamma^{-1}\log(2Q_1). \end{aligned}$$

Exponentiating both sides of the last inequality and setting $K_1 = 2^{1-\gamma}Q_1^{1-\gamma^{-1}}$ finishes the proof. □

4. Construction of MME

As stated above, the variational principle applies to shift spaces since they are compact metrics spaces and σ is continuous. The construction of the MME is a byproduct of the proof of the variational principle.

Lemma 4.1. *Let Σ be a shift space. The topological entropy of σ is equal to the supremum over measures $\mu \in M(\Sigma, \sigma)$. That is,*

$$h_{\text{top}}(\sigma) = \sup\{h_\mu(\sigma)|\ \mu \in M(\Sigma, \sigma)\}.$$

Proof. Let $\mu \in M(\Sigma, \sigma)$ and $\zeta = \{[1], ...[p]\}$. Recall that for each $n \geq 1$, $\bigvee_{k=0}^{n-1} \sigma^{-k}\zeta$ is just the collection of n-cylinders $\mathcal{U}_n$. By noting that the collection $\{\mu([w])\colon [w] \in \mathcal{U}_n\}$ is a probability vector of $\#\mathcal{L}_n$ elements, we can apply proposition 2.7 and see that $H_\mu(\bigvee_{k=0}^{n-1} \sigma^{-k}\zeta) \leq \log(\#\mathcal{L}_n)$ for each $n \in \mathbb{N}$. Dividing by n and sending n to infinity gives $h_\mu(\sigma) \leq h_{\text{top}}(\sigma)$ for any $\mu \in M(\Sigma, \sigma)$. Therefore, $\sup\{h_\mu(\sigma)|\ \mu \in M(\Sigma, \sigma)\} \leq h_{\text{top}}(\sigma)$.

For each $n \in \mathbb{N}$, construct a set $E_n \subset X$ by arbitrarily choosing an element from each n-cylinder in $\mathcal{U}_n$ Note that each E_n has the cardinality of $\mathcal{L}_n$ and that for every $[w]$ contains exactly one element of E_n. Define measures

$$\nu_n = \frac{1}{\#E_n} \sum_{x \in E_n} \delta_x \text{ and } \mu_n = \frac{1}{n} \sum_{k=0}^{n-1} \nu_n \circ \sigma^{-k}.$$

By the compactness of $M(\Sigma)$, there exists a subsequence $\{n_j\}$ such that $\lim_{j\to\infty} \log \#\mathcal{L}_{n_j} = h_{\text{top}}(\sigma)$ and μ_{n_j} converges to a measure μ. By proposition 2.9, $\mu \in M(X, \sigma)$.

We will show that $h_\mu(\sigma) \geq h_{\text{top}}(\sigma)$ which will finish the proof and imply that μ is a MME. Recall that the measure-theoretic entropy of μ is given by $h_\mu(\sigma) = \lim_{M\to\infty} \frac{1}{M} H_\mu\big(\bigvee_{k=0}^{M-1} \sigma^{-k}\zeta\big)$. Thus, establishing the inequality amounts to estimating $H_\mu(\bigvee_{k=0}^{M-1} \sigma^{-k}\zeta)$ from below. First, note that

$$H_\mu\left(\bigvee_{k=0}^{M-1} \sigma^{-k}\zeta\right) = \lim_{j\to\infty} H_{\mu_{n_j}}\left(\bigvee_{k=0}^{M-1} \sigma^{-k}\zeta\right),$$

where $\{n_j\}$ is the subsequence on which μ_{n_j} converges to μ. Fix $M \in \mathbb{N}$. Define a new sequence $n'_j := \lfloor \frac{n_j}{M} \rfloor$ and note that $n'_j M \leq n_j < (n'_j + 1)M$. Therefore by the convergence of $\{n_j\}$, we see that convergence also happens along the subsequence $\{n'_j M\}$. Hence, we need to estimate $H_{\mu_{n'_j M}}\big(\bigvee_{k=0}^{M-1} \sigma^{-k}\zeta\big)$. So,

$$H_{\mu_{n'_j M}}\left(\bigvee_{p=0}^{M-1} \sigma^{-p}\zeta\right) = \sum_{w \in \mathcal{L}_M} -\mu_{n'_j M}(w) \log \mu_{n'_j M}(w)$$

$$\geq \frac{1}{n'_j M} \sum_{k=0}^{n'_j M-1} \sum_{w \in \mathcal{L}_M} -\nu_{n'_j M}(\sigma^{-k}(w)) \log \nu_{n'_j M}(\sigma^{-k}(w))$$

$$= \frac{1}{n'_j M} \sum_{k=0}^{n'_j M-1} H_{\nu_{n'_j M}}\Big(\sigma^{-k} \bigvee_{p=0}^{M-1} \sigma^{-p}\zeta\Big)$$

$$= \frac{1}{n'_j M} \sum_{i=0}^{M-1} \sum_{l=0}^{n-1} H_{\nu_{n'_j M}}\Big(\sigma^{-(lM+i)} \bigvee_{p=0}^{M-1} \sigma^{-p}\zeta\Big),$$

since finite sums are interchangeable and $-x \log x$ is a concave function. For $i = 0$, note that

$$\sum_{l=0}^{n-1} H_{\nu_{n'_j M}}(\sigma^{-lM} \bigvee_{p=0}^{M-1} \sigma^{p}\zeta) = \sum_{l=0}^{n-1} H_{\nu_{n'_j M}}\Big(\bigvee_{p=lM}^{(l+1)M-1} \sigma^{-p}\zeta\Big) \geq H_{\nu_{n'_j M}}\Big(\bigvee_{p=0}^{n'_j M-1} \sigma^{-p}\zeta\Big).$$

It follows that for all $0 \leq i \leq M-1$,

$$\sum_{l=0}^{n-1} H_{\nu_{n'_j M}}(\sigma^{-(lM+i)} \bigvee_{p=0}^{M-1} \sigma^{-p}\zeta) \geq H_{\nu_{n'_j M}}(\sigma^{-i} \bigvee_{p=0}^{n'_j M-1} \sigma^{-p}\zeta)$$

which in turn implies that

$$H_{\mu_{n'_j M}}(\bigvee_{k=0}^{M-1} \sigma^{-k}\zeta) \geq \frac{1}{n'_j M} \sum_{i=0}^{M-1} \sum_{j=0}^{n-1} H_{\nu_{n'_j M}}(\sigma^{-(jM+i)} \bigvee_{k=0}^{M-1} \sigma^{-k}\zeta)$$

$$\geq \frac{1}{n'_j M} \sum_{i=0}^{M-1} H_{\nu_{n'_j M}}(\sigma^{-i} \bigvee_{k=0}^{n'_j M-1} \sigma^{-k}\zeta)$$

$$= \frac{1}{n'_j M} \sum_{i=0}^{M-1} H_{\nu_{n'_j M}}\Big(\bigvee_{k=i}^{i+n'_j M-1} \sigma^{-k}\zeta\Big).$$

Note that $H_{\nu_{n'_j M}}\big(\bigvee_{k=i}^{i+n'_j M-1} \sigma^{-k}\zeta\big) \geq H_{\nu_{n'_j M}}\big(\bigvee_{k=i}^{n'_j M-1} \sigma^{-k}\zeta\big)$. Using propositions 2.4 and 2.7, we see that

$$\begin{aligned} H_{\nu_{n'_jM}}\Big(\bigvee_{k=0}^{n'_jM-1}\sigma^{-k}\zeta\Big) &\le H_{\nu_{n'_jM}}\Big(\bigvee_{k=0}^{i-1}\sigma^{-k}\zeta\Big) + H_{\nu_{n'_jM}}\Big(\bigvee_{k=i}^{n'_jM-1}\sigma^{-k}\zeta\Big) \\ &\le \log\#\mathcal{L}_i + H_{\nu_{n'_jM}}\Big(\bigvee_{k=i}^{n'_jM-1}\sigma^{-k}\zeta\Big) \\ &\le M\log\#\zeta + H_{\nu_{n'_jM}}\Big(\bigvee_{k=i}^{n'_jM-1}\sigma^{-k}\zeta\Big).^{1} \end{aligned}$$

Thus, $H_{\mu_{n'_jM}}(\bigvee_{k=0}^{M-1}\sigma^{-k}\zeta) \ge \frac{1}{n'_j}H_{\nu_{n'_jM}}\big(\bigvee_{k=0}^{n'_jM-1}\sigma^{-k}\zeta\big) - \frac{M}{n'_j}\log\#\zeta$.

Since $\mathcal{U}_{n'_jM} = \bigvee_{k=0}^{n'_jM-1}\sigma^{-k}\zeta$ and $\nu_{n'_jM}(w) = (\#\mathcal{L}_{n'_jM})^{-1}$ for all $w\in\mathcal{L}_{n'_jM}$, apply proposition 2.7 again gives $H_{\nu_{n'_jM}}(\bigvee_{k=0}^{n'_jM-1}\sigma^{-k}\zeta) = \log\#\mathcal{L}_{n'_jM}$. Hence,

$$\begin{aligned} H_\mu\Big(\bigvee_{k=0}^{M-1}\sigma^{-k}\zeta\Big) &= \lim_{j\to\infty} H_{\mu_{n'_jM}}\Big(\bigvee_{k=0}^{M-1}\sigma^{-k}\zeta\Big) \\ &\ge \lim_{j\to\infty}\left[\frac{1}{n'_j}\log\#\mathcal{L}_{n'_jM} - \frac{M}{n'_j}\log\#\zeta\right] \end{aligned}$$

yielding $H_\mu\big(\bigvee_{k=0}^{M-1}\sigma^{-k}\zeta\big) \ge Mh_{\text{top}}(\sigma)$. Dividing by M and then letting M tend towards ∞ gives $h_\mu(\sigma) = \lim_{M\to\infty}\frac{1}{M}H_\mu(\bigvee_{k=0}^{M-1}\sigma^{-k}\zeta) \ge h_{\text{top}}(\sigma)$ as desired. Hence finishing the proof of the variational principle for shift spaces. □

5. Gibbs property for MME

Here we prove a Gibbs property for the MME constructed in the previous section. It will be used later to show help show the uniqueness of μ.

Lemma 5.1. *There exists a constant $Q_2 > 0$ such that for every $n\in\mathbb{N}$ and $w\in\mathcal{L}_n$,*

$$Q_2^{-1} \le \frac{\mu([w])}{e^{-nh}} \le Q_2.$$

[1]The last inequality is a consequence of the fact that $\log\#\mathcal{L}_i \le \log(\#\zeta)^i \le M\log\#\zeta$.

Proof. Fix $n \in \mathbb{N}$. Choose $m \gg n$ and let $k < m - n$. Consider $w \in \mathcal{L}_n$. We begin by estimating $\nu_m \circ \sigma^{-k}[w]$. If $v \in \mathcal{L}_m$ and $x \in [v]$, then $x \in \sigma^{-k}[w] \iff \sigma^k x \in [w] \iff x_k x_{k+1} \cdots x_{k+n-1} = w$. This last statement is equivalent to $v_k \cdots v_{k+n-1} = w$ so if we let $\mathcal{H}_m(w) = \{v \in \mathcal{L}_m | \ v_k \cdots v_{k+n-1} = w\}$,

$$\nu_m(\sigma^{-k}[w]) = \frac{1}{\#\mathcal{L}_m} \sum_{x \in E_m} \delta_x(\sigma^{-k}[w]) = \frac{\#\mathcal{H}_m(w)}{\#\mathcal{L}_m}$$

since the cylinder corresponding to a word in $\mathcal{L}_m$ contains only one element of E_m.

For the lower bound, we consider a map that uses the specification property to glue admissible words of lengths $\mathcal{L}_{k-\tau}$ and $\mathcal{L}_{m-k-n-\tau}$ to their respective ends of w creating the collection $\mathcal{H}_m(w)$ of words in $\mathcal{L}_m$ where w appears in the k-th position. Define a map $\mathcal{L}_{k-\tau} \times \mathcal{L}_{m-k-n-\tau} \to \mathcal{H}_m(w)$ by $(u, v) \mapsto u s_u w s_v v$, where $s_u, s_v \in \mathcal{L}_\tau$ are given by specification. Note that map is not surjective since the gluing words given by specification do not need to be unique. If two points $(u^1, v^1), (u^2, v^2) \in \mathcal{L}_{k-\tau} \times \mathcal{L}_{m-k-n-\tau}$ get mapped to equivalent points $u^1 s_u^1 w s_v^1 v^1$ and $u^2 s_u^2 w s_v^2 v^2$, then $u^1 = u^2$ and $v^1 = v^2$. That is, the map is injective so $\#\mathcal{L}_{k-\tau}\#\mathcal{L}_{m-k-n-\tau} \leq \#\mathcal{H}_m(w)$. By Lemma 3.1, we see that if $K_2 = \frac{e^{-2\tau h}}{Q_1}$

$$\nu_m(\sigma^{-k}[w]) \geq \frac{\#\mathcal{L}_{k-\tau}\#\mathcal{L}_{m-k-n-\tau}}{\#\mathcal{L}_m} \geq \frac{(e^{(k-\tau)h})(e^{(m-k-n-\tau)h})}{Q_1 e^{mh}} = K_2 e^{-nh}.$$

Therefore, by summing over k we find that $\sum_{k=0}^{m-1} \nu_m(\sigma^{-k}[w]) \geq m K_{21} e^{-nh} \implies \mu_m([w]) \geq K_2 e^{-nh}$. So sending $m \to \infty$ establishes a lower bound.

For the upper bound, we consider the map $\mathcal{H}_m(w) \to \mathcal{L}_k \times \mathcal{L}_{m-k-n}$ defined as $v \mapsto (v_1 \cdots v_k, v_{k+n+1} \cdots v_m)$. This map is clearly injective so $\mathcal{H}_m(w) \leq \#\mathcal{L}_k\#\mathcal{L}_{m-k-n}$. Using Lemma 3.1 again gives

$$\nu_m(\sigma^{-k}[w]) \leq \frac{\#\mathcal{L}_k\#\mathcal{L}_{m-k-n}}{\#\mathcal{L}_m} \leq \frac{(Q_1 e^{kh})(Q_1 e^{(m-k-n)h})}{e^{mh}} = Q_1^2 e^{-nh}.$$

Thus, $\mu_m([w]) \leq Q_1^2 e^{-nh}$. Sending $m \to \infty$ gives a upper bound. Let $Q_2 = \max\{Q_1^2, K_2^{-1}\}$. Then $Q_2^{-1} \leq \frac{\mu([w])}{e^{-nh}} \leq Q_2$ as desired. □

6. Ergodicity of MME

Consider a probability space $(X, \mathcal{B}, m)$. A measure-preserving transformation T of $(X, \mathcal{B}, m)$ is called ***ergodic*** if $B \in \mathcal{B}$ and $T^{-1}B = B$ implies that $m(B) = 0$ or $m(B) = 1$. A measure $\nu \in M(X,T)$ is say to be ergodic if T is ergodic on (X, ν). We will use the following proposition to show that the MME μ constructed in section 4 is ergodic. For more background in ergodicity, we refer the reader to chapter 1 of [7].

Proposition 6.1. *Let T be a measure-preserving transformation of the probability space $(X, \mathcal{B}, m)$. If for every $A, B \in \mathcal{B}$ with $m(A) > 0$, $m(B) > 0$, there exists $n > 0$ such that $m(T^{-n}A \cap B) > 0$, then T is ergodic.*

Proof. Suppose $B \in \mathcal{B}$ with $T^{-1}B = B$. If $0 < m(B) < 1$, then $m(X \backslash B) \neq 0$ and $0 = m(B \cap (X \backslash B)) = m(T^{-n}B \cap (X \backslash B))$ for all $n \geq 1$, a contradiction to the hypothesis. □

Therefore, the ergodicity of μ is a consequence of the following lemma. We will need to build up a few results to prove it.

Lemma 6.2. *If two measurable sets $P, Q \subset \Sigma$ both have positive μ-measure, then*

$$\overline{\lim_{n\to\infty}} \mu(P \cap \sigma^{-n}(Q)) > 0.$$

Proof. First, we show that there exists a subsequence $m_j \nearrow \infty$ and constant $K_3 > 0$ such that for cylinder sets of $u, v \in \mathcal{L}$, $\mu([u] \cap \sigma^{-m_j}[v]) \geq K_3 \mu([u])\mu([v])$ for sufficiently large j.

Consider $u, v \in \mathcal{L}$ and denote their lengths as n_u and n_v. Let $\{n_j\}$ be an increasing sequence of natural numbers. Let $m \in \mathbb{N}$ be large and fix $k \leq m - n_u - n_v - n_j$. Define another sequence by letting $m_j = n_j + n_u + 2\tau$. Similar to how we estimated the μ-measure of a cylinder set for the Gibbs property, we consider $\nu_m(\sigma^{-k}[u] \cap \sigma^{-(k+m_j)}[v])$. Note that

$$\sigma^{-k}[u] = \{x \in \mathcal{L}_m : x_k \cdots x_{k+n_u-1} = u\}$$

and

$$\sigma^{-(k+m_j)}[v] = \{x \in \mathcal{L}_m : x_{k+m_j} \cdots x_{k+m_j+n_v-1} = v\}.$$

Thus, the intersection of the two sets are all the words $x \in \Sigma$ such that u appears in the k-th position of x and v appears m_j positions later. That is,

$$\sigma^{-k}[u] \cap \sigma^{-(k+m_j)}[v] = \{x \in \mathcal{L}_m : x_k \cdots x_{k+n_u-1} = u,\ x_{k+m_j} \cdots x_{k+m_j+n_v-1} = v\}.$$

Therefore, $\nu_m(\sigma^{-k}[u] \cap \sigma^{-(k+m_j)}[v]) = \frac{\#(\sigma^{-k}[u] \cap \sigma^{-(k+m_j)}[v])}{\#\mathcal{L}_m}$ and the desired bound can be obtained by estimating $\#(\sigma^{-k}[u] \cap \sigma^{-(k+m_j)}[v])$.

Define a map $\mathcal{L}_{k-\tau} \times \mathcal{L}_{n_j} \times \mathcal{L}_{m-k-m_j-n_v-\tau} \to \sigma^{-k}[u] \cap \sigma^{-(k+m_j)}[v]$ by sending points $(w_1, w_2, w_3) \longmapsto w_1 s_1 u s_2 w_2 s_3 v s_4 w_3$ where each s_i has length τ provided by specification. This map is clearly injective so

$$(\#\mathcal{L}_{k-\tau})(\#\mathcal{L}_{n_j})(\#\mathcal{L}_{m-k-m_j-n_v-\tau}) \leq \#(\sigma^{-k}[u] \cap \sigma^{-(k+m_j)}[v]).$$

Thus by using lemmas 3.1 and 5.1, we see that

$$\begin{aligned}\nu_m(\sigma^{-k}[u] \cap \sigma^{-(k+m_j)}[v]) &\geq \frac{\#\mathcal{L}_{k-\tau}\#\mathcal{L}_{n_j}\#\mathcal{L}_{m-k-m_j-n_v-\tau}}{\mathcal{L}_m} \\ &\geq \frac{e^{m-n_u-n_v-4\tau}}{Q_1 e^{mh}} = K_2 e^{-2\tau h} e^{-(n_u+n_v)h} \\ &\geq k_2 Q_2^{-2} e^{-2\tau h} \mu([u])\mu([v]).\end{aligned}$$

Writing $K_3 = CQ_2^{-2}e^{-2\tau h}$ yields $\mu_m([u] \cap \sigma^{-m_j}[v]) \geq K_3\mu([u])\mu([v])$. Therefore passing to the convergent subsequence gives

$$\mu([u] \cap \sigma^{-m_j}[v]) \geq K_3\mu([u])\mu([v]).$$

We can immediately extend this result to the case when P and Q are the unions of cylinders of the same length by noting that if $[P] = \bigcup_{w \in P}[w]$,

$$\begin{aligned}\mu([P] \cap \sigma^{-m_j}[Q]) &\geq \sum_{x \in P,\ y \in Q} \mu([x] \cap \sigma^{-m_j}[y]) \\ &\geq \sum_{x \in P,\ y \in Q} K\mu([x])\mu([y]) = K\mu([P])\mu([Q]),\end{aligned}$$

where we again used the notation $\mu(x) \sim \mu([w])$ and $\mu(P) \sim \mu([P])$.

Finally, let P and Q be measurable subsets of Σ. Fix $\varepsilon > 0$ and choose sets U, V that are unions of cylinders of the same length satisfying $\mu(U \triangle P) < \varepsilon$ and $\mu(V \triangle Q) < \varepsilon$. Observe that for every natural number n,

$$\begin{aligned}|\mu(U \cap \sigma^{-n}(V) - \mu(P \cap \sigma^{-n}(Q))| &\leq \mu((U \cap \sigma^{-n}(V) \triangle (P \cap \sigma^{-n}(Q))) \\ &\leq \mu((U \triangle P) \cap (V \triangle Q)) < \varepsilon,\end{aligned}$$

where the first inequality is given by proposition 2.10 and the second by monotinicity. This along with result above yields

$$\overline{\lim}_{n\to\infty}\mu(P \cap \sigma^{-n}(Q) \geq K_3\mu(P)\mu(Q) - \varepsilon.$$

Therefore, $\overline{\lim}_{n\to\infty}\mu(P\cap\sigma^{-n}(Q) \geq K_3\mu(P)\mu(Q)$ since ε was arbitrary. Hence if P and Q both have positive measure, then $K_3\mu(P)\mu(Q) > 0$ finishing the proof of the lemma.

□

7. Uniqueness of MME

Finally, we finish the proof of theorem 1.1 by showing that the existence of another ergodic MME on Σ violates the Gibbs property of section 5.

The proof of uniqueness requires a result about mutual singularity. Consider two probability measures m and ν. We say that these measures are ***mutually singular*** if there exists disjoint measurable sets E and F such that $E \cup F = X$ and $m(E) = \nu(F) = 0$. We omit the proof of the following proposition because it is beyond the scoop of this paper. Instead we refer the reader to theorem 6.10 in Walters' book [7].

Proposition 7.1. *Let T be a continuous transformation of the compact metric space X. If $\mu, \nu \in M(X,T)$ are both ergodic and $\nu \neq \mu$, then they are mutually singular.*

Lemma 7.2. *The measure of maximal entropy μ constructed in section 4 is unique.*

Proof. Let μ be the ergodic MME as constructed in section 4. Suppose there exists another ergodic $\nu \in M(\Sigma,\sigma)$ such that $h_\nu(\sigma) = h_{\text{top}}(\sigma)$. By proposition 7.1, ν and μ are mutually singular. Consider a collection of words $\mathcal{D} \subset \mathcal{L}$ such that $\mu(\mathcal{D}_n) \to 0$ and $\nu(\mathcal{D}_n) \to 1$. Then if we fix $\varepsilon > 0$, there exits $N \in \mathbb{N}$ such that $\nu(\mathcal{D}_n) > 1 - \varepsilon$ for all $n \geq N$. Applying lemma 3.2 gives a constant $K_1 > 0$ such that $\#\mathcal{D}_n \geq K_1 e^{nh}$ for all n. Choose $n \geq N$. Using the Gibbs property proven in section 5, we see that for every $n \in \mathbb{N}$,

$$\mu(\mathcal{D}_n) = \mu\Bigg(\bigcup_{w\in\mathcal{D}_n}[w]\Bigg) = \sum_{w\in\mathcal{D}_n}\mu([w]) \geq \#\mathcal{D}_n Q_2^{-1}e^{-nh} \geq K_1 Q_2^{-1} > 0.$$

This is a contradiction the fact that $\mu(\mathcal{D}_n) \to 0$. Thus establishing uniqueness. □

References

[1] Ben Sherman. Measures of maximal entropy for shifts of finite type. 2013.

[2] Rufus Bowen. Some Systems with Unique Equilibrium States. *Mathematical Systems Theory*, 8(3):193–202, 1975.

[3] Vaughn Climenhaga Daniel Thompson. Intrinsic Ergodicity Beyond Specificatoin β-Shifts, S-Gap Shifts, and their Factors. *Israel Journal of Mathematics*, (192):785–817, 2012.

[4] Vaughn Climenhaga. Shifts with specification have a unique MME. 2015.

[5] Bruce P. Kitchens. *Symbolic Dynamics*. Springer, 1998.

[6] Douglas Lind Brian Marcus. *An Introduction to Symbolic Dynamics and coding*. Cambridge University Press, 1995.

[7] Peter Walters. *An Introduction to Ergodic Theory*. Springer-Verlag New York Inc., 1982.