COURSE SUMMARY

GABRIELA BROWN
SYMBOLIC DYNAMICS - WOLF
SPRING 2023

Contents

1. Introduction

2. Fundamentals of Shift Spaces

For the entirety of this document we assume a finite alphabet $\mathcal{A}$. [1][C.1,C.2]

Definition 2.0.1 (Shift Space)**.** [2, L.2] We give two equivalent definitions.

- A *shift space* X is a subset of the full shift Σ_d^+ that is σ-invariant and closed.

- Shift spaces and *languages* are the same thing. A language is a set $\mathcal{L}$ of finite words where subwords are also in the language and any word can be extended into a longer word in the language.

2.1. **Shifts of Finite Type.**

Definition 2.1.1 (SFT)**.** [2, L.2] A *shift of finite type* or SFT is a subshift whose language is defined by a finite set of forbidden words, $\mathcal{F} = \mathcal{L}(\Sigma_d^+) \, \mathcal{L}(X)$. An arbitrary set of words can have words added to it to be made into a forbidden word set whose complement is a language. We say an SFT is an m-step SFT if the longest word in $\mathcal{F}$ has length $m + 1$.

e.g. 2.1.2 (Full Shift)**.**

e.g. 2.1.3 (Golden Mean Shift)**.**

e.g. 2.1.4 (Higher Coded Block Shift)**.**

Theorem 2.1.5. *[2, L.2] Every SFT is conjugate to a 1-step SFT.*

Proof. Given an m-step SFT the mth higher block representation is a 1-step SFT that is topologically conjugate by the map that takes X to $X^{[m]}$. $\qquad\qquad\square$

2.2. **Sofic Shifts.**

Definition 2.2.1 (Sofic Shift)**.** [2, L.3] A sofic shift X is a shift that is described by a finite, directed, labeled graph $\mathcal{G} = (G, l)$.

e.g. 2.2.2 (Even Shift)**.**

Definition 2.2.3. Let $f : X \to X$ and $g : Y \to Y$ be dynamical systems. We say g is a *factor* of f if there exists a continuous surjection $h : X \to Y$ such that $g \circ h = h \circ f$. We say that f is an *extension* of g. (TODO:square diagram)

Theorem 2.2.4. *[2, L.4] X is a sofic shift if and only if X is a factor of an SFT.*

e.g. 2.2.5. The labeling map l is a factor map.

2.3. **Coded Shifts.**

Definition 2.3.1 (Coded Shift)**.** [2, L.8] We give two equivalent definitions.

- A coded shift is a shift $X = \overline{X_\mathcal{G}}$ described by the closure of concatenations of a countable set of generating words $\mathcal{G}$.

- A coded shift is a shift $X_\mathcal{G}$ described by a countable, directed labeled graph $\mathcal{G}$.

e.g. 2.3.2 (S-Gap Shift)**.** [2, L.8]

e.g. 2.3.3 (Context-Free Shift)**.** [2, L.4]

e.g. 2.3.4 (Beta Shifts)**.** [2, L.8, L.9]

Theorem 2.3.5. *[2, L.8] Every coded shift is described by a uniquely representable generating set.*

2.4. **Topological Entropy in Shift Spaces.** [2, L5]

3. Computability of entropy of shift spaces

Our goal in this section is to define the notion of computability at a point, and then use that notion to understand the computability of entropy of shift spaces. For SFTs and sofic shifts it is sufficient to work with topological entropy as we defined in the last section. However to understand coded shifts we will need to enter the world of measure-theoretic entropy, and we'll also need a few important results of ergodic theory in order to be able to move between these two version of entropy.

3.1. **Computability in the Real Numbers.** [2, L6, L7] We begin by discussing the ideas of computability in the context of $\mathbb{R}$ with the usual norm $||$.

Definition 3.1.1. Let $x \in \mathbb{R}$. An *oracle* of x is a function $\varphi : \mathbb{N} \to \mathbb{Q}$ such that $|\varphi(n) - x| < 2^{-n}$.

More intuitively, an oracle of x is something that gives an approximation of x to whatever requested degree of precision.

Definition 3.1.2. We say $x \in \mathbb{R}$ is computable if there exists a turing machine χ which is an oracle of x.

For example, I can write a computer program that takes a desired degree of precision n and outputs the first n digits of π. It is a bit tricky to find examples of uncomputable numbers. Rational, algebraic and some transcendental numbers are all computable, but in fact, most numbers are not computable! It is hard to produce an example of one because if we could write it down then we would have an algorithm for computing it. To produce an example, we make use of the halting problem, which states that there is no universal turing machine that decides if any given Turing machine halts. There are a countable number of turing machines (because programs are binary, which are integers), so let T_i be an ordering of all turing machines. Then define

$$t_n = \begin{cases} 0 & T_n \text{ halts} \\ 1 & T_n \text{ doesn't halt.} \end{cases}$$

Let $\alpha = 0.t_1 t_2 t_3 \ldots$. This number α must be uncomputable otherwise it would be a universal turing machine that solves the halting problem.

Definition 3.1.3. We say $x \in \mathbb{R}$ is *upper semi-computable* if there exists a turing machine $\chi : \mathbb{N} \to \mathbb{Q}$ such that $(\chi(n))$ is a nonincreasing sequence with $\lim_{n \to \infty} \chi(n) = x$. Similarly χ is *lower semi-computable* if $-x$ is upper semi-computable.

Proposition 3.1.4. *A real number is computable if and only if it is upper and lower semi-computable.*

The proof idea is to let the upper and lower turing machines run until they are within 2^{-n} of each other. Now that we have defined computability of a real number, we would like to have a notion of computability of a function.

Definition 3.1.5. Let $S \subseteq \mathbb{R}^n$. We say $g : S \to \mathbb{R}$ is *computable* if there exists a turing machine χ such that for all $x \in S$, oracles φ of x, and $n \in \mathbb{N}$ that $\chi(\varphi, n) \in \mathbb{Q}$ and $|\chi(\varphi, n) - g(x)| < 2^{-n}$.

3.2. **Computability in a general metric space.** Now that we have built an intuition for the definitions of computability of a real number, and functions to the real numbers, we'll describe when we can do this in a general metric space. Here the set S acts for X as $\mathbb{Q}$ does for $\mathbb{R}$.

Definition 3.2.1. Let (X, d_x) be a metric space with a countable, dense subset S (which makes the the space complete and separable). Then (X, d_X, S) is a *computable metric space* if there exists a turing machine $\chi : \mathbb{N} \times \mathbb{N} \times \mathbb{Z} \to \mathbb{Q}$ such that $|\chi(i, j, n) - d_X(s_i, s_j)| < 2^{-n}$.

Definition 3.2.2. Let (X, d_X, S) be a computable metric space. We say a $\varphi : \mathbb{N} \to \mathbb{N}$ is an *oracle* of $x \in X$ if $d_X|\varphi(s_n), x| < 2^{-1}$.

Definition 3.2.3. We say a function $g : Y \subseteq X \to \mathbb{R}$ is computable if there exists a turing machine $\chi : \mathbb{N} \times \mathbb{N} \to \mathbb{Q}$ such that for all $x \in Y$ and oracles φ of $x, |\chi(\varphi, n) - g(x)| < 2^{-n}$.

Definition 3.2.4 (Computability at a point)**.** Let (X, d_X, S) be a computable metric space, and $Y \leqslant X$, $x_0 \in X$. We say a function $g : Y \to \mathbb{R}$ is *computable at* x_0 if there exists a turing machine $\chi : \mathbb{N} \times \mathbb{N} \times \mathbb{N} \to \mathbb{Q}$ such that for any oracle φ of x_0:

> Let $l_{\varphi,n}$ be the maximal precision that χ queries in the calculation of $\chi(\varphi, n)$. Then for all $y \in S$ such that there exists an oracle φ' for y which coincides with φ up to precision $l_{\varphi,n}$ we have $\chi(\varphi, n) = \chi(\varphi', n)$ and $|\chi(\varphi, n) - g(y)| < 2^{-n}$.

3.3. **The set of shift spaces is a metric space.** To talk about computability in shift spaces, we need to first set up the space of all shift spaces and verify that it has some reasonable properties. We want to show that the set of all shift spaces Σ_{inf} is a metric space that is complete and compact. Then we can introduce the notion of computability in general metric spaces. From there we will have the notion of Σ_{inf} being computable and we will be able to then discuss the computability of certain properties of shift spaces, like entropy.

Proposition 3.3.1. *Let* $X, Y \in \Sigma_{inf}$ *be shift spaces. Then*

$$d_S(X, Y) = \begin{cases} 2^{-n} & n = \inf \{ n : \mathcal{L}_n(X) \neq \mathcal{L}_n(Y) \} \\ 0 & X = Y \end{cases}$$

is a metric for Σ_{inf}*.*

Proposition 3.3.2. *The metric space* (Σ_{inf}, d_S) *is complete.*

Proposition 3.3.3. *The metric space* (Σ_{inf}, d_S) *is compact.*

These three propositions tell us that Σ_{inf} is a complete, compact metric space with metric d_S. We take it as a fact that it becomes a computable metric space with the sense set S being the set of SFTs.

3.4. **Computability in the metric space of all shifts.** [2, L7]

Definition 3.4.1. Let $X \in \Sigma_{inf}$. We say φ is an *oracle* of X if $\varphi(n) = \mathcal{L}_{2n+1}(X)$. i.e. increasing precision means giving the language for increasingly large words.

Definition 3.4.2. We say a shift $X \in \Sigma_{inf}$ is *computable* if there exists a turing machine χ which is an oracle of X.

Definition 3.4.3. We say $X \in \Sigma_{inf}$ is *upper semi-computable* if there exists a turing machine $\chi : \mathbb{N} \times \mathbb{N} \to \mathbb{Q}$ such that the sequence $\{\chi(n,k)\}_k$ is a nonincreasing finite collection of words of length $2n+1$ such that $\chi(n,k) = \mathcal{L}_{2n+1}(X)$ for k large enough.

Proposition 3.4.4. $h_{top}(X)$ *is upper semi-computable for any* $X \in \Sigma_{inf}$ *given by an oracle.*

Proof. Given φ we can compute $\frac{1}{n}log|\mathcal{L}_n(X)|$ at any precision. We can select a rational number r_n such that
$$\frac{1}{n} \log |\mathcal{L}_n(X)| < r_n \leqslant \frac{1}{n}|\mathcal{L}_n(X)| + 2^{-n}.$$

Define $q_n = \min\{r_1, \ldots, r_n\}$. Then (q_n) is a nonincreasing computable sequence of rational numbers which converges from above to $h_{top}(X)$ (because $h_{top}(X) = \inf_n \frac{1}{n} \log |\mathcal{L}_n(X)|$.) $\qquad\square$

Theorem 3.4.5. *The function* $X \to h_{top}(X)$ *is computable at* X *iff.* $h_{top}(X) = 0$.

Proof. Case 1: $h_{top}(X) > 0$. We will show that the function is not computable at X by finding an SFT arbitrarily close to X that breaks the definition of computability at a point. Fix $n \in \mathbb{N}$ (this is the precision input to the Turing machine). The context we're working in is the computable metric space $(\Sigma_{inf}, d_{inf}, SFTs)$. Since SFTs are dense, there exists an SFT $X_{\mathcal{F}}$ such that $d_{inf}(X, X_{\mathcal{F}}) < 2^{-n}$, i.e. $\mathcal{L}_n(X) = \mathcal{L}_n(X_{\mathcal{F}})$. We can just consider the transitive components $X_1 \ldots X_k$ of $X_{\mathcal{F}}$. Transitive SFTs have dense periodic points so we can find a periodic orbit p_i for each X_i such that $\mathcal{L}_{2n+1}(p_i) = \mathcal{L}_{2n+1}(X_i)$. Now we construct an SFT $Y = \bigcup_{i=1}^k p_i$. We know the entropy of Y because the number of distinguishable orbits is fixed under the iteration of the shift map, so $h_{top}(Y) = 0$. This shift Y is a point that breaks the definition of computability at a point, since oracles of shifts are given by languages and we guaranteed X and Y have the same language up to an arbitrary precision n, yet the function maps $h_{top}X > 0$ and $h_{top}(Y) = 0$. So there can't exist a turing machine according to the definition because we will always be able to produce one of these points with 0 entropy close to X.

Case 2: $h_top(X) = 0$. We will show X satisfies the definition of computability at a point. First define $h(m) = \min\{\frac{1}{k}log|\mathcal{L}_k(X) : k \leqslant m\}$. By the previous proposition $\lim_m \to \infty h(m)$ converges to $h_top(X) = 0$ from above. Now fix $n \in \mathbb{N}$. There exists an $m \in \mathbb{N}$ such that $h(m) \leqslant 2^{-n}$, which can be computed. Let Y be a subshift whose language coincides with the language of X up to n, i.e. $\mathcal{L}_m(Y) = \mathcal{L}_m(X)$. Then
$$0 \leqslant h_{top}(Y) \leqslant h(m) \leqslant 2^{-n}.$$
The middle inequality holds because $h_Y(m)$ converges to $h_top(Y)$ from above, and $h_X(m) = h_Y(m)$. So $h_X(m)$ is an upper bound for $h_{top}(Y)$. This fulfills the last part of the definition of computability at a point. $\qquad\square$

4. Calculating the entropy of coded shifts

4.1. **Crash Course in Ergodic Theory.** [2, L10]

Let f be a generic dynamical system. Ergodic theory is the study of f-invariant measures. We say a measure is an **ergodic measure** if for all f-invariant subsets A, either $\mu(A) = 0$ or $\mu(A) = 1$. These ergodic measures pick out a particular invariant subset to give full measure, which creates a notion of decomposition into smaller dynamical systems. Any measure is a convex sum of ergodic measures, so it also gives us a way to study measures by their more neatly behaved components. We will go about identifying ergodic measures by the fact that they are the extreme points in the space of measures.

Let (X, d) be a compact metric space with continuous dynamical system $f : X \to X$. We use $\boldsymbol{\mathcal{M}_{X,f}}$ to mean the set of f-invariant Borel probability measures on X. (f-invariant means $\mu(f^{-1}(A)) = \mu(A) \ \forall \ A \in \mathcal{B}(X)$ and probability measure means $\mu(X) = 1$.) The set $\mathcal{M}_{X,f}$ has the **Wasserstein-Kantorovitch distance**, $w_1(\mu, \nu) = \sup_{\phi \in 1\text{-Lip}(X,\mathbb{R})} |\int \phi d\mu - \int \phi d\nu|$. This metric induces the weak* topology (the coursest topology such that all the integrals of any $\phi \in L^1(X, \mathbb{R})$ with respect to any measure are continuous). We will now give several statements that let us find the ergodic measures in $\mathcal{M}_{X,f}$.

Fact 4.1.1. $\mathcal{M}_{X,f}$ *is compact under the weak* topology.*

Theorem 4.1.2. $\mathcal{M}_{X,f} \neq \varnothing$.

Proof. // TODO: in notes □

Fact 4.1.3. $\mathcal{M}_{X,f}$ *is convex. (Between any two points you can draw a straight line that stays in the set).*

In a compact, convex topological space you can define extreme points. In fact, the set of ergodic measures of $\mathcal{M}_{X,f}$ is exactly the extreme points of the set (convex hull?), which we denote $\mathcal{M}_E \subseteq \mathcal{M}_{X,f}$. We write the decomposition of a general measure into ergodic measures in the following way. Let $\mu \in \mathcal{M}_{X,f}$. Then there exists a λ on $\mathcal{M}_{X,f}$ such that $\lambda(\mathcal{M}_E) = 1$ and $\int \phi d\mu = \int \int \phi d\nu d\lambda$ for all $\phi \in C(X, \mathbb{R})$ varying over ν. In other words, the non-ergodic measures form a set of measure 0 for λ and do not contribute to the integral.

We will need the following theorem in the next section to understand the technique of inducing a dynamical system onto a preferred subset.

Theorem 4.1.4 (Poincare Recurrence Theorem). *Let $\mu \in \mathcal{M}_{X,f}$, and $A \in \mathcal{B}(X)$ with $\mu(A) > 0$. Then the set of points that return to A a finite number of times has measure 0. Equivalently,*

$$\mu(\{x \in A : f^n(x) \in A \text{ infinitely many times }\}) = \mu(A).$$

The next theorem gives us a key connection between the so called time average and space average of a dynamical system. This connection is important because only the space average relies on a measure, and so it allows us to bridge topological and measure theoretic notions of entropy.

Theorem 4.1.5 (Birkhoff Ergodic Theorem). *Let $\mu \in \mathcal{M}_{X,f}$, and $\phi \in L^1(X, \mu)$. Then the time average given by*

$$\phi_f(x) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k(x)) \quad \text{exists } \mu \text{ a.e.}$$

and $\phi_f \in L^1(X, \mu)$. *Furthermore if μ is ergodic then ϕ_f is constant almost everywhere and*

$$\phi_f(x) = \int \phi d\mu \quad \text{for a.e. } x \in X.$$

TODO: not sure about this definition of the basin. This theorem gives us a natural definition of the **basin of** μ $B(\mu)$ which is the set of points in X where the time average is equal to the space average for $\phi = 1$, $\{x \in X : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \delta_{f^k(x)} = \mu\}$. Note that the basin is f-invariant.

Corollary 4.1.6. *If μ is ergodic, then $\mu(B(\mu)) = 1$.*

4.2. **Connecting Topological and Measure Theoretic Entropy.** [2, L10, L11] The following theorem gives us a measure-theoretic definition of entropy, which is defined pointwise and then integrated. Again in this section we let (X, d) be a compact metric space with $f : X \to X$ continuous and $\mathcal{M}_{X,f}$ the set of f-invariant Borel probability measures.

Theorem 4.2.1 (Katok-Brin). *Let $\mu \in \mathcal{M}_{X,f}$. Then the following limit exists*

$$h_\mu(x) = \lim_{r\to\infty} \lim_{n\to\infty} -\frac{1}{n} \log \mu(B_n(x, r)) \quad \mu\text{-a.e. } x.$$

Furthermore $x \mapsto h_\mu(x)$ is μ-measurable, and if μ is ergodic then $h_\mu(x)$ is continuous and we can define

$$h_\mu(f) = \int h_\mu(x) d\mu(x).$$

This definition requires that for measure theoretic entropy to be positive, the measure must live on points that are pushed apart, that is, the Bowen ball that travels with x is getting smaller in the limit. We are now prepared to see the connection between topological entropy and measure theoretic entropy, commonly referred to as the Variational Principle.

Theorem 4.2.2 (Variational Principle).

$$h_{top}(f) = \sup_{\mu \in \mathcal{M}_{X,f}} h_\mu(f)$$

If there is a measure μ such that $h_\mu(f) = h_{top}(f)$ then we call μ a **measure of maximal entropy (mme)**. The supremum in this equation means that for some dynamical systems there is no mme.

We've already described the computability of h_{top} for SFTs and sofic shifts, and in order to understand it for coded shifts we're going to need to use the variational principle to study the problem from the perspective of measure-theoretic entropy. But even then if we can find ergodic measures, it can still be tricky to study, so we're going to use ideas from the dynamical inverse problem and the technique of inducing to set up a space of potentials with equillibrium states that lets us study observables instead, which we really can get our hands on. After all this work we'll see that topological entropy for a coded shift is computable if it lives on the concatenations of generators, and not computable if it lives in the limit.

First we need to set up the dynamical inverse problem, which is interested in when we can reconstruct a dynamical system just from observations we make in the world. For example a scientist might measure the temperature every day at the same time, and we would like to know if it is possible to accurately reconstruct the underlying dynamical system of weather just from those observations. If $f : X \to X$ is the underlying dynamical system, the **observables** are functions

$\phi \in C(X, \mathbb{R})$. The following notion of topological pressure formalizes certain circumstances where it is possible to reconstruct from observables. It works for dynamical systems in general, but we will write it in terms of shift spaces so we can more readily apply it to the situation at hand.

Definition 4.2.3 (Topological Pressure)**.** Let $X \in \Sigma_{inv}$, with shift map $\sigma : X \to X$. The topological pressure $P_{top} : C(X, \mathbb{R}) \to \mathbb{R}$ is defined

$$P_{top}(\phi) = \begin{cases} \lim_{n \to \infty} \frac{1}{n} \log Z_n(\phi) & \text{finite alphabet}, \\ \inf_n \{ \frac{1}{n} \log Z_n(\phi) \} & \text{countable alphabet}, \end{cases}$$

where $Z_n(\phi) = \sum_{w \in \mathcal{L}_n(X)} \exp(\sup_{x \in [w]_0^{n-1}} S_n \phi(x))$ is the nth partition function, and $S_n \phi(x) = \sum_{k=0}^{n-1} \phi(\sigma^k(x))$ is the weighted orbit of x along n. This limit exists because $\log Z_n(\phi)$ is subadditive.

Note that this is a purely topological construction, but we can tie it into measure theoretic ideas. We say that a measure μ is an **equilibrium state** of the observable $\phi \in C(X, \mathbb{R})$ if $P_{top}(\phi) = F_\mu(\phi)$, where $F_\mu(\phi) = h_\mu(\sigma) + \int \phi d\mu$ is the **free energy** of ϕ with respect to μ. This leads us to our second statement of the variational principle,

Theorem 4.2.4 (Variational principle)**.**

$$P_{top}(\phi) = \sup_{\mu \in \mathcal{M}_{X,\sigma}} \left(h_\mu(\sigma) + \int \phi d\mu \right).$$

In this version we have phrased the connection between measure theoretic and topological entropy in terms of these observable functions ϕ. Note that if $\phi = 0$, this statement becomes the first version of the variational principle, and μ is an mme.

// TODO: thm about shift spaces having unique equilibrium state?

We now introduce the technique of inducing, which lets us essentially restrict a dynamical system to a preferred set in a way where we can keep track of all the important information, like the measure, entropy, and ergodicity. We will talk about this specifically in the setting of shift spaces, though the technique is broadly applicable. Let X be a shift space, with shift $\sigma : X \to X$, invariant probability measure $\mu \in \mathcal{M}_{X,\sigma}$, and a Borel measurable set $E \subseteq \mathcal{B}(X)$.

For any $x \in X$ we define the **first return time** $\tau_E(x)$ to be the first time $n \geqslant 1$ that $\sigma^n(x) \in E$. Now we consider the set $\boldsymbol{E^\infty}$ of points of E that return infinitely often to E. This lets us define the **first return map** $\sigma_E(x) : E^\infty \to E^\infty$ by $x \mapsto \sigma^{\tau_E(x)}(x)$, which sends a point to wherever it is when it first returns to E. We can use the first return map to push the measure of X onto E in the following way. Define the **induced measure** $i(\mu) = \mu_E(F) = \frac{\mu(F \cap E)}{\mu(E)}$ which only measures sets that intersect with E. This way of restricting μ to E is very similar to conditional probability, and the definition looks almost identical to Baye's theorem. Now we will establish some facts about the induced measure.

Fact 4.2.5. $\mu(E^\infty) = \mu(E)$ *(by Poincare recurrence.)*

Fact 4.2.6. μ_E *is a probability measure on* E.

Fact 4.2.7. *If* μ *is ergodic then* μ_E *is ergodic.*

Let $\mathcal{N}_{E^\infty,\sigma_E} = \{\nu \in \mathcal{M}_{E^\infty,\sigma_E} : \int \tau_E(x)d\nu(x) < \infty\}$, that is, the measures where the integral over the first return time is finite. We define

$$\tilde{\nu}(B) = \Sigma_{k=0}^{\infty}\nu(\{x : \tau_E(x) > k\} \cap f^{-k}(B)).$$

Then we define the **lift** of ν: $l(\nu) = \frac{\tilde{\nu}}{\tilde{\nu}(X)}$, which gives us a natural bijection,

Proposition 4.2.8. $i(l(\nu)) = \nu$. *In other words,*

$$\mathcal{M}_{X,\sigma} \xrightarrow{\;\;i\;\;} \mathcal{N}_{E^\infty,\sigma_E}$$
$$\xleftarrow{\;\;l\;\;}$$

is a natural bijection.

Being so convinced that we can induce the new dynamical system on E^∞ and recover information from it due to the bijection, we give several facts about the relationship between important properties of the original and induced system.

Fact 4.2.9. *The entropy of the induced measure is* $h_{i(\mu)}(\sigma_E) = \frac{h_\mu(\sigma)}{\mu(E)}$.

Fact 4.2.10. *Given some observable* $\phi \in C(X,\mathbb{R})$, *we can define an observable for* E^∞,

$$\phi_E = \Sigma_{k=0}^{\tau_E(x)-1}\phi(\sigma^k(x))$$

which gives a map from the integral with respect to the orginal measure to the integral with respect to the induced measure,

$$\int \phi d\mu \mapsto \int_E \left(\Sigma_{k=0}^{\tau_E(x)-1}\phi(\sigma^k(x))\right) d\mu = \frac{\int_E \phi d\mu}{\mu(E)} = \int \phi_E d\mu_E$$

???

Fact 4.2.11. *The free energy of the induced measure is* $F_{\mu_E,\sigma_E}(\phi_E) = \frac{F_{\mu,\sigma}(\phi)}{\mu(E)}$.

4.3. **Entropy of Coded Shifts.** [2, L12] // TODO: wait for 05/03 lecture

4.4. **Examples of Computing Entropy of Shift Spaces.** // TODO: add examples from 04/26

4.5. **Questions.**

(1) Is every shift a coded shift?

(2) I don't understand the paragraph about how to decompose a general measure into ergodic measures.

(3) definition of basin seems wrong.

(4) what's that about shift spaces having unique equilibrium state?

(5) defining the induced integral over an observable?

Appendix A. Quick Reference Sheet

A.1. **Shift Spaces.**

def: TODO

SFTs

idea: described by a transition matrix (adjacency matrix or graph with 1-1 labeling)

idea: defined by a language OR a forbidden word set

def:

thm: every SFT is topologically conjugate to a 1-step SFT (1-step on the bottom of the commutative square by the higher block representation w/ mad forbidden word len 2)

e.g. Full shift

e.g. Golden mean shift

Sofic shifts

idea: described by a finite labeled directed graph over a finite alphabet

def:

e.g. Even shift TODO

thm: Sofic Shifts are always factors of some SFT. (Sofic on the bottom of the commutative square).

Coded shifts

idea: described by an irreducible countable directed labeled graph over a finite alphabet.

idea: described by countable generating set

def: Let $\mathcal{G}$ be a TODO: define unique representation

e.g. S-gap shifts TODO

e.g. Beta shifts TODO

thm: the class of coded shifts contains both sofic shifts and SFTs.

Topological entropy of shifts

idea: count distinguishable orbits under some margin of measurement like pixels on a screen,

idea: how surprised are you by the next letter? what is the information exponential growth rate as words get longer

def: (general) exponential growth of the maximal $n - \epsilon$ separated set

def: (shifts) Let $X \in \Sigma_{\inf}$ be a shift space. Then $h_{top}(X) = \lim_{n \to \infty} \frac{1}{n} \log |\mathcal{L}_n(X)|$.

thm: the entropy of an SFT described by an irreducible transition matrix A is $\log \lambda$ where λ is the Perron eigenvalue of A.

thm: the entropy of an SFT described by a matrix which is the product of irreducible transition matrices A_i is $\log \lambda_i$ where λ_i is the max of the perron eigenvalues of the A_i.

A.2. **Ergodic Theory.**

-

Measure theory basics

def: let X a topological space, $\Sigma \subseteq P(X)$ is a sigma algebra if 1) $X, \phi \in \Sigma$, 2) arbitrary unions are in Σ, 3) complements are in Σ. (2 and 3 $\Rightarrow$ arbitrary intersections)

def: dthe Borel sigma algebra is $\mathcal{B}$ the smallest sigma algebra which contains all open sets according to the provided topology.

def: a measure is a function $\mu : \mathcal{B} \to \mathbb{R}_+$ such that 1) for a pairwise disjoint set of $A_i \in \mathcal{B}$, $\mu(\cup A_i) = \sum \mu(A_i)$. TODO what else

def: $\mathcal{M}_{X,T}$ for (X, d) compact metric space, $T : X \to X$, it is the set of all invariant Borel probability measures on X, i.e. $\mu(A) = \mu(T^{-1}(A))$

thm: $\mathcal{M}_X$ has the weak-* topology: the smallest topology where $\mu \to \int \phi d\mu$ is continuous for all $\phi \in C(X, \mathbb{R})$. (i.e. $\mu_n \to \mu$ iff. $\int \phi d\mu_n \to \int \phi d\mu$)

thm: $\mathcal{M}_X$ is compact and non-empty in the weak-* topology.

thm: the weak-* topology is metrizable by the Wasserstein-Kantorovich metric between probability distributions.

Relevant Ergodic Theory

def: a set is <u>convex</u> if drawing a line between any two points stays within the set. For $\mathcal{M}_X$, $\mu_1, \mu_2 \in \mathcal{M}_X \Rightarrow \alpha\mu_1 + (1-\alpha)\mu_2 \in \mathcal{M}_X$ for $\alpha \in [0,1]$.

def: $\mu \in M_X$ is <u>ergodic</u> if $\mu(A) = 1$ or $\mu(A) = 0$ for all T-invariant subsets A.

thm: $\mathcal{M}_E$ the extreme points of $\mathcal{M}_X$ are precisely the ergodic measures.

def: the <u>basin</u> of a measure is the set $\mathcal{B}(\mu) = \{x \in X : \lim_{n\to\infty} \sum_{k=0}^{n-1} \delta_{f^k(x)} = \mu\}$ (TODO: how is this related to the time average?)

def: <u>measure of maximal entropy</u> is a μ s.t. $\overline{h_\mu(f) = h_{top}(f)}$.

thm: (Poincare Recurrence) Let $\mu \in \mathcal{M}_X$, $A \in \mathcal{B}(X)$ with $\mu(A) > 0$. The measure on the set of points that return to A finite times is 0. i.e. $\mu(A) = \mu(\{x \in A : \text{return to } A \text{ infinitely often }\})$.

thm: (Birkhoff Ergodic) Let $\mu \in \mathcal{M}_X$, and $\phi \in L^1(X, \mu)$. The time average $\phi_f(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k(x))$ exists for x $\mu.a.e$ and if μ is ergodic then the time average equals the space average $\mu_f(x) = \int \phi d\mu$ for almost every x.

thm: If μ is ergodic then $\mu(\mathcal{B}(\mu)) = 1$.

Entropy of Coded Shifts

idea: The method of computing entropy for SFTs and Sofic shifts relies on a matrix describing the shift, and Coded shifts are countable and can't be described this way.

idea: Topological entropy can be given in terms of topological pressure, and topological pressure is equivalent for an induced system. So we induce to a system where we know how to compute measure theoretic entropy, and since measure theoretic entropy and topological pressure are related this enables us to compute the topological entropy of the original system.

def: <u>topological pressure</u>

def: <u>free energy</u>

def: <u>equillibrium state</u>

thm: Variational Principle I:

thm: Variational Principle II:

thm: Inducing

TODO

(1) finish adding computability in shift spaces content

(2) finish cheat sheet for Computability

(3) finish cheat sheet for entropy of coded shifts

(4) finish cheat sheet for shift space basics

(5) add fundamentals of shift spaces content

(6) go back and add all proofs

(7) read review paper

(1) subadditivity??

(2) how is the result of the computability discussion different than the computing the entropy of a coded shift

Appendix B. Lecture index

L1	01/25
L2	02/01
L3	02/15
L4	02/22
L5	03/02
L6	03/08
L7	03/15
L8	03/22
L9	04/05
L10	04/19
L11	04/26
L12	05/03

References

[1] Brian Marcus Douglas Lind. *An Introduction to Symbolic Dynamics and Coding.* Ed. by Cambridge University Press. Second. Cambridge Mathematical Library, 2021.

[2] Christian Wolf. *Symbolic Dynamics and Computability.* CUNY Topics Course. Spring 2023.