# COURSE SUMMARY

GABRIELA BROWN
SYMBOLIC DYNAMICS - WOLF
SPRING 2023

## Contents

## 1. Introduction

## 2. Fundamentals of Shift Spaces

[**lind-marc**][C.1,C.2]

### 2.1. SFTs.

### 2.2. Sofic Shifts.

### 2.3. Coded Shifts. [**wolf**]

### 2.4. Topological Entropy in Shift Spaces. [**wolf**]

## 3. Computability 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. Introduction to Computability.
[**wolf**] 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 $\chi$ 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 $\pi$. 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 $\alpha$ 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 $\chi$ 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 $\chi$ such that for all $x \in S$, oracles $\varphi$ of $x$, and $n \in \mathbb{N}$ that $\chi(\varphi, n) \in \mathbb{Q}$ and $\chi(\varphi, n) - g(x)| < 2^{-n}$.

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 $\Sigma_{inf}$ 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 $\Sigma_{inf}$ being computable and we will be able to then discuss the computability of certain properties of shift spaces, like entropy.

**Proposition 3.1.6.** *Let $X, Y \in \Sigma_{inf}$ be shift spaces. Then*

$$d_S(X, Y) = \begin{cases} 0 & X = Y \\ \frac{1}{2}^l & l = \inf\{w_l \in \mathcal{L}(X) \cup \mathcal{L}(Y)\backslash\mathcal{L}(X) \cap \mathcal{L}(Y)\} \end{cases}$$

*is a metric for $\Sigma_{inf}$. That is, $l$ is smallest word length that causes $\mathcal{L}_l(X) \neq \mathcal{L}_l(Y)$, and the smaller this number the further apart the spaces are.*

**Proposition 3.1.7.** *The metric space $(\Sigma_{inf}, d_S)$ is complete.*

**Proposition 3.1.8.** *The metric space $(\Sigma_{inf}, d_S)$ is compact.*

These three propositions tell us that $\Sigma_{inf}$ is a complete, compact metric space with metric $d_S$.

### 3.2. **Computability of Sofic Shifts and SFTs.** [**wolf**]

### 3.3. **Crash Course in Ergodic Theory.** [**wolf**]

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 $\mathcal{M}_{X,f}$ 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 3.3.1.** *$\mathcal{M}_{X,f}$ is compact under the weak* topology.*

**Theorem 3.3.2.** *$\mathcal{M}_{X,f} \neq \varnothing$.*

*Proof.* // TODO: in notes $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

**Fact 3.3.3.** *$\mathcal{M}_{X,f}$ 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 $\lambda$ 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 $\nu$. In other words, the non-ergodic measures form a set of measure 0 for $\lambda$ 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 3.3.4** (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 3.3.5** (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 $\mu$ is ergodic then $\phi_f$ 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** $\mu$ $B(\mu)$ 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 3.3.6.** *If $\mu$ is ergodic, then $\mu(B(\mu)) = 1$.*

3.4. **Connecting Topological and Measure Theoretic Entropy.** [**wolf**] 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 3.4.1** (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 $\mu$-measurable, and if $\mu$ 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 3.4.2** (Variational Principle)**.**

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

If there is a measure $\mu$ such that $h_\mu(f) = h_{top}(f)$ then we call $\mu$ 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 3.4.3** (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 $n^{th}$ 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 $\mu$ 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 $\phi$ with respect to $\mu$. This leads us to our second statement of the variational principle,

**Theorem 3.4.4** (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 $\phi$. Note that if $\phi = 0$, this statement becomes the first version of the variational principle, and $\mu$ 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 $\mu$ 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 3.4.5.** $\mu(E^\infty) = \mu(E)$ *(by Poincare recurrence.)*

**Fact 3.4.6.** $\mu_E$ *is a probability measure on* $E$.

**Fact 3.4.7.** *If* $\mu$ *is ergodic then* $\mu_E$ *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 $\nu$: $l(\nu) = \frac{\tilde{\nu}}{\tilde{\nu}(X)}$, which gives us a natural bijection,

**Proposition 3.4.8.** $i(l(\nu)) = \nu$. *In other words,*

$$\mathcal{M}_{X,\sigma} \; \underset{l}{\overset{i}{\rightleftarrows}} \; \mathcal{N}_{E^\infty,\sigma_E}$$

*is a natural bijection.*

Being so convinced that we can induce the new dynamical system on $E^\infty$ 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 3.4.9.** *The entropy of the induced measure is* $h_{i(\mu)}(\sigma_E) = \frac{h_\mu(\sigma)}{\mu(E)}$.

**Fact 3.4.10.** *Given some observable* $\phi \in C(X, \mathbb{R})$, *we can define an observable for* $E^\infty$,
$$\phi_E = \Sigma_{k=0}^{\tau_E(x)-1} \phi(\sigma^k(x))$$

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

3.5. **Entropy of Coded Shifts.** [**wolf**] // TODO: wait for 05/03 lecture

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

3.7. **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)

- OR: defined by a language OR a forbidden word set

- def:

- 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

- 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.

- OR: 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

- 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,

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

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

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

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

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

### A.2. **Ergodic Theory.**

-

**measure theory basics**

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

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

- measure: 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

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

- $\mathcal{M}_X$ 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. $\int \phi d\mu_n \to \int \phi d\mu$)

## 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 |