Notes for Junior Seminar

Princeton Math Members

Spring 2014

Abstract

These are a compilation of lecture notes from *MAT983 - Ergodic and Dynamic Properties of Shift Spaces*. They are based on Chapters 1, 5 and 6 from *Substitutions in Dynamics, Arithemtics and Combinatorics* by N. Pytheas Fogg [1]. The seminar's goals are to introduce shift spaces and related dynamical systems and to explore topological and ergodic properties in the context of shift spaces. The individuals responsible for each lecture are listed by initial at the beginning of each title[1].

[1] JC = Jay Jojo Cheng, KD = Kubrat Danailov, RJ = Russell Jack Jenkins, ZQ = Zhaonan Henry Qu, PR = Paul Rapoport, AY = Alexander Yu, AZ = Albert Zhou.

Contents

1 Chapter 1

1.1 JF: Definitions up to Shift Spaces

Lecture date: 2014/02/17.

$\mathcal{A}$ is an **alphabet**, a finite set of symbols, with $\#\mathcal{A} > 1$. A **word** $W = w_1 w_2 \dots w_n$ is any string of letters in $\mathcal{A}$, and ε is the empty word. Let $\mathcal{A}^*$ be the collection of all words with letters in $\mathcal{A}$, including ε. We endow $\mathcal{A}^*$ with **concatenation** as multiplication, i.e.

$$UV = u_1 \dots u_n v_1 \dots v_m \tag{1}$$

where $U = u_1 \dots u_n$ and $V = v_1 \dots v_m$. We let $|U|$ denote the **length** (or **size**) of U and $|U|_a$ denote the number of occurrences of a in U.

Definition 1.1. For $\mathbb{N} = \{0, 1, 2, 3 \dots\} \subseteq \mathbb{Z}$, the set

$$\mathcal{A}^{\mathbb{F}} = \{u : u = (u_n)_{n \in \mathbb{F}} \text{ and } \forall n \in \mathbb{F}, u_n \in \mathcal{A}\} \tag{2}$$

where $\mathbb{F} \in \{\mathbb{N}, \mathbb{Z}\}$, is the set of **one-sided sequences** when $\mathbb{F} = \mathbb{N}$ and **two-sided sequences** if $\mathbb{F} = \mathbb{Z}$.

For a sequence u, $W = w_0 \dots w_n$ is a **factor** of u if there exists m so that

$$u_m u_{m+1} \dots u_{m+n} = w_0 w_1 \dots w_n. \tag{3}$$

In this case, we say that W occurs in position m in u. W is a **factor** of word U if there exist (possibly empty) words V and X so that

$$U = VWX. \tag{4}$$

The **language** of u is

$$\mathcal{L}(u) = \{W \in \mathcal{A}^* : W \text{ is a factor of } u\} \tag{5}$$

and is the set of all factors of u. The set of all factors of length n of u is $\mathcal{L}_n(u)$.

Definition 1.2. u is **recurrent** if for each $W \in \mathcal{L}(u)$, the set of occurrences of W in u is not bounded (from either above or below if $u \in \mathcal{A}^{\mathbb{Z}}$).

Example 1.3. If $\mathcal{A} = \{0, 1\}$, define $u, v \in \mathcal{A}^{\mathbb{N}}$ by

$$u_n = \begin{cases} 0, & n \text{ is even,} \\ 1, & n \text{ is odd,} \end{cases} \quad \text{and } v_n = \begin{cases} 0, & n \leq 10, \\ 1, & n > 10. \end{cases} \tag{6}$$

We see that $\mathcal{L}_n(u) = 2$ for all n and u is recurrent. However,

$$\mathcal{L}_n(v) = \min\{n + 1, 12\} \tag{7}$$

and v is not recurrent.

Definition 1.4. A sequence u is **ultimately periodic** if there exists t and n_0 so that

$$\forall |n| \geq n_0,\ u_{n+t} = u_n, \tag{8}$$

and u is **periodic** if $n_0 = 0$. We call u **shift-periodic** in either case usually use ultimately periodic to mean $n_0 > 0$.

Example 1.5. From Example 1.3, u is periodic with $t = 2$ and v is ultimately periodic with $t = 1$ and $n_0 = 11$.

Definition 1.6. For sequence u, the **complexity function** $p_u : \mathbb{N} \to \mathbb{N}$ is given by

$$p_u(n) = \#\mathcal{L}_n(u) \tag{9}$$

or $p_u(n)$ is the number of factors of length n in u.

There are some fundamental facts about the complexity function:

- If $d = \#\mathcal{A}$, then for each $n \in \mathbb{N}$, $1 \leq p_u(n) \leq d^n$.

  Because u is an infinite sequence, there mus exist a factor of length n, namely $u_1 \ldots u_n$. The total number of words of length n is d^n.

- For each n, $p_u(n) \leq p_u(n+1)$.

  If U is a factor of length n at position m, then the factor of length $n+1$ at position m is of the form Ux for $x \in \mathcal{A}$. So for every factor of U length n, there exists at least one factor of length $n+1$ that begins with U.

We will now define $\mathcal{A}^{\mathbb{N}}$ (or $\mathcal{A}^{\mathbb{Z}}$) as a **topological space**.

Definition 1.7. For $\mathcal{A}^{\mathbb{N}}$ we use the **product topology** $\mathcal{T}$ given by the discrete topology on $\mathcal{A}$.

$\mathcal{T}$ may be defined by many equivalent means. Define a metric on $\mathcal{A}^{\mathbb{N}}$(or $\mathcal{A}^{\mathbb{Z}}$) by

$$d(u, v) = 2^{-n} \tag{10}$$

where $n = \min\{|n| : u_n \neq v_n\}$. $\mathcal{T}$ is the topology generated by this metric. Alternately, a cylinder for word $W = w_0 w_1 \ldots w_n$, $[W]$ is the cylinder generated by W in $\mathcal{A}^{\mathbb{N}}$,

$$[W] = \{u \in \mathcal{A}^{\mathbb{N}} : u_0 = w_0,\ u_1 = w_1,\ \ \ldots u_n = w_n\}. \tag{11}$$

$\mathcal{T}$ is the topology generated by basis $\{[W] : W \in \mathcal{A}^*\}$. We may do the same for $\mathcal{A}^{\mathbb{Z}}$ by defining a cylinder $[V.W]$ by $u \in [V.W]$ if and only if $u_{-m} \ldots u_n = VW$ where $m = |V|$ and $n + 1 = |W|$.

The topological space $(\mathcal{A}^{\mathbb{N}}, \mathcal{T})$ is a **complete**, **metric**, **compact** and **totally disconnected** space. To have a **topological dynamical system**, we need a continuous map on this space.

Definition 1.8. For $\mathcal{A}^{\mathbb{F}}$, $\mathbb{F} \in \{\mathbb{N}, \mathbb{Z}\}$, the (left) shift $S : \mathcal{A}^{\mathbb{F}} \to \mathcal{A}^{\mathbb{F}}$ is given by

$$\big(S(u)\big)_n = u_{n+1},\ \forall u \in \mathcal{A}^{\mathbb{F}}, n \in \mathbb{F}. \tag{12}$$

Example 1.9. From Example 1.3, $S(u) = 10101010\ldots$ and $S(S(u)) = u$. For $k \geq 0$,

$$S^k(v)_n = \begin{cases} 1, & n \leq 10-k \\ 0, & n > 10-k. \end{cases} \tag{13}$$

In particular, $S^k(v) = S^{11}(v) = 0000\ldots$ for $k \geq 11$.

By verifying that $d(S(u), S(v)) \leq 2d(u, v)$, we see that $S : \mathcal{A}^{\mathbb{F}} \to \mathcal{A}^{\mathbb{F}}$ is continuous. Furthermore, if $\mathbb{F} = \mathbb{Z}$, S is a **homeomorphism**. If $\mathbb{F} = \mathbb{N}$, S is surjective, but not injective.

In this way, $(\mathcal{A}^{\mathbb{F}}, \mathcal{T}, S)$ is a **topological dynamical system**. We will typically look at subspaces, $(X, \mathcal{T}, S)$ for $X \subseteq \mathcal{A}^{\mathbb{F}}$ such that $S(X) \subseteq X$. There will be some rule or method to generate X. Our first method will be to use an element u.

Definition 1.10. If $u \in \mathcal{A}^{\mathbb{N}}$, then the **orbit** of u under S is

$$\mathcal{O}_S(u) = \{S^k(u) : k \in \mathbb{N}\}. \tag{14}$$

The **(symbolic) dynamical system associated to** u is $(X_u, \mathcal{T}, S)$ where $X_u = \overline{\mathcal{O}_S(u)}$.

What does it mean to belong to X_u?

Lemma 1.11. *(Lemma 1.1.2 in the text) If $w \in \mathcal{A}^{\mathbb{N}}$, the following are equivalent:*

1. *$w \in X_u$,*
2. *for each $n \in \mathbb{N}$, there exists $k_n \in \mathbb{N}$ so that $w_0 \ldots w_n$ appears in position k_n in u,*
3. *$\mathcal{L}(w) \subseteq \mathcal{L}(u)$.*

1.2 JF: Measure Theory

Lecture date: 2014/02/23.

Definition 1.12. A dynamical system $(X, \mathcal{T}, T)$ is **minimal** if and only if for every $x \in X$, $\overline{\mathcal{O}_T(x)} = X$, or every orbit is **dense** in X.

Lemma 1.13. *Let $u \in \mathcal{A}^{\mathbb{N}}$ and $(X_u, \mathcal{T}, S)$ its related shift space. Then the following are equivalent:*

1. *$(X_u, \mathcal{T}, S)$ is minimal.*
2. *u is **uniformly recurrent (minimal)**, or for every $n \in \mathbb{N}$ there exists $N = N(n)$ so that*
$$\forall U \in \mathcal{L}_n(u), W \in \mathcal{L}_N(u),\ U \text{ is a factor of } W. \tag{15}$$
   *In other words, every word appears with with **bounded gaps**.*
3. *For every $w \in X_u$, $\mathcal{L}(w) = \mathcal{L}(u)$.*

Example 1.14. Let $\mathcal{A} = \{0,1\}^{\mathbb{N}}$ and $u \in \mathcal{A}^{\mathbb{N}}$ be

$$u = 0100110001110000111100000011111\dots \tag{16}$$

or given by

$$u_n = \begin{cases} 0, & \exists k \in \mathbb{N},\ k(k+1) \le n < (k+1)^2 \\ 1, & \exists k \in \mathbb{N},\ (k+1)^2 \le n < (k+1)(k+2). \end{cases} \tag{17}$$

(X_u, S) is not a minimal shift space. The constant sequence $\mathbf{0} = 000\cdots \in X_u$ and $X_{\mathbf{0}} = \{\mathbf{0}\} \neq X_u$. $1 \in \mathcal{L}(u)$, but there exists longer and longer words of the form $0^n = 0\dots0$, where 1 does not appear in u.

Definition 1.15. If (X, d) is a compact metric space, let $\mathcal{C}(X)$ be the vector space of continuous functions from X to $\mathbb{C}$. Then a **finite (positive) Borel measure** on X is any map $\mu : \mathcal{C}(X) \to \mathbb{C}$ such that

1. μ is linear.
2. $0 < \mu(\mathbf{1}) < \infty$, where $\mathbf{1}$ is the constant function with value 1.
3. If $f \ge 0$, then $\mu(f) \ge 0$.

μ is a probability measure on X if also $\mu(\mathbf{1}) = 1$.

We define the notation

$$\int_X f(x) d\mu(x) := \mu(f). \tag{18}$$

Example 1.16. If $x \in X$, then the **Dirac Mass** at x, $\delta_x : \mathcal{C}(X) \to \mathbb{C}$, defined by

$$\delta_x(f) = f(x) \tag{19}$$

defines a probability measure on X. So the space of all measures is non-empty.

Definition 1.17. Let $\mathcal{M}_X$ denote the set of probability measures of X. We endow this with the **weak-* topology**, which is defined by

$$\mu_n \to \mu \iff \forall f \in \mathcal{C}(X),\ \mu_n(f) \to \mu(f). \tag{20}$$

$\mathcal{M}_X$ is compact and metrizable, so it is also sequentially compact.

Definition 1.18. A measure $\mu \in \mathcal{M}_X$ is **T-invariant** for continuous $T : X \to X$ if

$$\forall f \in \mathcal{C}(X), \mu(f \circ T) = \mu(f). \tag{21}$$

Equivalently, if we define the measure $T_*\mu$ by $T_*\mu(f) = \mu(f \circ T)$ for all $f \in \mathcal{C}(X)$, then we say μ is T-invariant if $T_*\mu = \mu$.

$\mathcal{M}_X$ is **convex**, meaning for any $\mu, \nu \in \mathcal{M}(X)$ and $t \in [0, 1]$, $t\mu + (1-t)\nu$ defined by

$$(t\mu + (1-t)\nu)(f) = t\mu(f) + (1-t)\nu(f) \tag{22}$$

is an element of $\mathcal{M}_X$. In fact, the set of all finite measures has the structure of a **cone**: for μ, ν finite measures and $a > 0$, $\mu + \nu$ and $a\mu$ are measures.

We now show the following useful fact for our space X, $\mathcal{M}_X$ contains at least one T invariant measure.

Lemma 1.19. *(1.4.2 in Text) If (X, T) is a topological dynamical system, with X metrizable and compact, then there exists a T-invariant probability measure μ.*

Proof. Let $x \in X$ and μ_N be given by

$$\mu_N := \frac{1}{N} \sum_{j=0}^{N-1} \delta_{T^j x}. \tag{23}$$

We check that each $\mu_n \in \mathcal{M}_X$. By compactness, we may choose a convergent sub-sequence $K_N \nearrow \infty$ and define

$$\mu = \lim_{N \to \infty} \mu_{K_N}. \tag{24}$$

We may check that for each $f \in \mathcal{C}(X)$, $T_*\mu(f) = \mu(f)$ and so $T_*\mu = \mu$. $\square$

We have another notion of measure that is given in analysis courses. We will talk about its definition, and how it relates to the above definition for our case.

Definition 1.20. For a set X, let $\mathcal{P}(X)$ be the power set of X. A set $\mathcal{B} \subseteq \mathcal{P}(X)$ is a **σ-algebra** if:

1. $\emptyset \in \mathcal{B}$.
2. $\forall A \in \mathcal{B}$, $X \setminus A \in \mathcal{B}$.
3. $\forall A_1, \ldots, A_n, \cdots \in \mathcal{B}$, $\displaystyle\bigcup_{n=1}^{\infty} A_n \in \mathcal{B}$.

Generally, σ-algebras are difficult to write down. We would like to develop such an algebra from a set that is easier to describe:

Definition 1.21. If $\mathcal{B}' \subset \mathcal{P}(X)$, then we define the σ-algebra **generated** by $\mathcal{B}'$ to be

$$\mathcal{B}(\mathcal{B}') = \bigcap_{\mathcal{B}'' \in \Sigma} \mathcal{B}'' \tag{25}$$

where $\Sigma = \{\mathcal{B}'' \subseteq \mathcal{P}(X) : \mathcal{B}'' \supseteq \mathcal{B}' \text{ and } \mathcal{B}'' \text{ is a } \sigma\text{-algebra}\}$.

Example 1.22. If $X = [0,1]$ and $\mathcal{C} = \{[0, 1/n]\}_{n=1}^{\infty}$, then $\mathcal{B}(\mathcal{C})$ is the collection of all countable disjoint unions of sets

$$\emptyset, \{0\}, (1/(n+1), 1/n] \text{ for } n \geq 1. \tag{26}$$

When we already have a topological space $(X, \mathcal{T})$, we want to define a σ-algebra on X that works well with the topology.

Definition 1.23. If $(X, \mathcal{T})$ is a topological space, then the **Borel σ-algebra** is the one generated by $\mathcal{T}$, or $\mathcal{B} = \mathcal{B}(\mathcal{T})$.

This shall be our σ-algebra: the **Borel σ-algebra** generated by the topology generated by cylinders. We now will relate some ideas from topology to σ-algebras.

Definition 1.24. A pair $(X, \mathcal{B})$, X is a set with $\mathcal{B}$ a σ-algebra on X, is called a **measure space**. A function $T : X \to X$, is called **measurable** iff

$$\forall A \in \mathcal{B},\ T^{-1}(A) \in \mathcal{B}. \tag{27}$$

Also, $A \subseteq X$ is **measurable** iff $A \in \mathcal{B}$.

Likewise, $f : X \to Y$ is measurable if $f^{-1}(B)$is measurable for any measurable $B \subseteq Y$. Furthermore, if X and Y are both topological spaces and are endowed with their respective σ-algebras, then any continuous function $f : X \to Y$ is measurable.

Much like a topological space, a measure space is defined by a set X and a collection $\mathcal{B}$ of "good" subsets. However, once we have a measure space, we may now define more.

Definition 1.25. If $(X, \mathcal{B})$ is a measure space, then $\mu : \mathcal{B} \to [0, \infty)$ is a **finite (positive) measure** on $(X, \mathcal{B})$ if and only if

1. $\forall A \in \mathcal{B},\ \mu(A) \geq 0$
2. For disjoint $A_1 \ldots, A_n, \cdots \in \mathcal{B}$,

$$\mu\left(\bigcup_{n=1}^{\infty} A_1\right) = \sum_{n=1}^{\infty} \mu(A_n) \tag{28}$$

3. $\mu(X) < \infty$.

μ is a **probability measure** if $\mu(X)$.

Example 1.26. If $X = [0,1)$ with the Borel σ-algebra, then we have the Lebesgue measure λ, which is defined by

$$\lambda((a,b)) = b - a.$$

This is a probability measure on X.

We consider measurable maps that preserve a given measure, also given a measurable map we consider measures that do not change under this measure.

Definition 1.27. Let $(X, \mathcal{B})$ be a measure space, $T : X \to X$ measurable and μ a probability measure. Then we say that μ is **T-invariant** or T is **μ-invariant** if

$$\forall A \in \mathcal{B},\ \mu(T^{-1}A) = \mu(A). \tag{29}$$

Example 1.28. If (X, λ) are from Example 1.26, let T be the times two map mod 1, or

$$T(x) = \begin{cases} 2x, & x \in [0, 1/2), \\ 2x - 1, & x \in [1/2, 1). \end{cases}$$

We may confirm that T is λ-invariant by checking intervals.

From a measure μ, we may develop the idea of an integral with respect to μ. We begin with defining integrals on $f : X \to \mathbb{C}$ that are simple. We then say that a measurable function is integrable if we may express its integral as a reasonable limit of simple functions.

For our compact metric space, the **Reisz Represntation Theorem** says that our two notions of Borel measures are equivalent. Specifically, if μ is a Borel probability measure, then

$$f \mapsto \int_X f d\mu \tag{30}$$

defined a linear functional from $\mathcal{C}(X)$ to $\mathbb{C}$ as given in Definition 1.15. Conversely, if there is a measure from that definition, there must exist a corresponding $\mu : \mathcal{B} \to [0,1]$ from Definition 1.25.

1.3 JF: Ergodicity and Generic Points

Lecture date: 2014/03/03.

Definition 1.29. If $(X, \mathcal{B}, \mu, T)$ is a probability measure preserving space, it is **ergodic** if

$$\forall A \in \mathcal{B},\ T^{-1}A = A \Rightarrow \mu(A) \in \{0, 1\}. \tag{31}$$

Remark 1.30. Here are some equivalent notions to ergodic:

1. $\forall A \in \mathcal{B}$, $\mu(T^{-1}A \Delta A) = 0 \Rightarrow \mu(A) \in \{0, 1\}$.
2. $\forall f : X \to X$ measurable and $f \circ T = f$, then f is constant μ-almost everywhere.

There a number of beneficial properties follow from ergodicity. We will present one here, the Birkhoff Ergodic Theorem.

Theorem 1.31. *(Birkhoff's Ergodic Theorem) If (X, μ, T) is an ergodic probability system, then for every integrable $f : X \to \mathbb{C}$ and μ-almost every $x \in X$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{n-1} f(T^j x) = \int_X f \ d\mu. \tag{32}$$

So in particular, if $f = \chi_B$ for some measurable B, then for μ-almost every $x \in X$,

$$\lim_{N\to\infty} \frac{\#\{0 \le j \le N-1 : T^j x \in B\}}{N} = \mu(B). \tag{33}$$

Example 1.32. In terms of our shift space $(X_u, \mathcal{B}, S)$, if $\mu \in \mathcal{M}_X$, this says that for μ almost every $w \in X_u$,

$$\lim_{M\to\infty} \frac{1}{M} N(W, w|_0^{M-1}) = \mu([W]) \tag{34}$$

for every $W \in \mathcal{A}^*$, where $w|_0^{M-1} = w_0 w_1 \ldots w_{M-1}$ and $N(W, V)$ is the number of occurrences of W in V.

So if we define the **frequency** of W in V as

$$\delta(W, V) = \frac{N(W, V)}{|V|}, \tag{35}$$

then we say that w is **generic** for μ if for every W, $\delta(W, w|_0^M) \to \mu([W])$. By Birkhoff's Ergodic Theorem, if μ is an ergodic measure, generic points exist.

Remember that $\mathcal{M}_X$ is convex, so we may talk about **extremal points** of $\mathcal{M}_X$. These are points $\mu \in \mathcal{M}_X$ so that for any $\nu_1, \nu_2 \in \mathcal{M}_X$ and $t \in (0, 1)$

$$\mu = t\nu_1 + (1-t)\nu_2 \Rightarrow \nu_1 = \nu_2 = \mu. \tag{36}$$

Theorem 1.33. *The ergodic measures in $\mathcal{M}_X$, $\mathcal{E}_X$, are the extremal points.*

Proof. Suppose $\mu \in \mathcal{E}_X$ and $\mu = t\nu_1 + (1-t)\nu_2$ for $t \in (0,1)$ and $\nu_1, \nu_2 \in \mathcal{M}_X$. Then for every measurable A, $\mu(A) = 0 \Rightarrow \nu_1(A) = \nu_2(A) = 0$. Therefore each measure ν_j has a **Radon-Nikodym** derivative f_j with respect to μ, or

$$\nu_j(B) = \int_B f_j d\mu \tag{37}$$

for f_j integrable (and therefore measurable). We may also show that $f_j(T \circ x) = f_j(x)$ for μ almost every x, and so f_j is constant μ-almost everywhere by Remark 1.30.

So there is a μ full measure set $E \subseteq X$ so that

$$f_1(x) = c_1 \text{ and } f_2(x) = c_2 \text{ for all } x \in E. \tag{38}$$

Then because

$$\mu(X \setminus E) \Rightarrow \nu_j(X \setminus E) = 0 \Rightarrow \nu_j(E) = 1, \tag{39}$$

we see that

$$\nu_j(X) = \int_E f_j d\mu = \int_E c_j d\mu = c_j \mu(E) = c_j. \tag{40}$$

And so $c_1 = c_2 = 1$, which implies that $\nu_1 = \nu_2$ and so $\mu = \nu_1 = \nu_2$. Therefore μ is extremal.

If $\mu \notin \mathcal{E}_X$, then there exists T-invariant set A so that $0 < \mu(A) < 1$. Then let ν_1, ν_2 be given by

$$\nu_1(B) = \frac{\mu(A \cap B)}{\mu(A)} \text{ and } \nu_2(B) = \frac{\mu((X \setminus A) \cap B)}{\mu(X \setminus A)}. \tag{41}$$

$\nu_1, \nu_2 \in \mathcal{M}_X$, so μ is not extremal as

$$\mu = t\nu_1 + (1-t)\nu_2 \tag{42}$$

where $t = \mu(A)$.

□

We will not formally state it here, but the **Ergodic Decomposition Theorem** says that every $\mu \in \mathcal{M}_X$ is a "convex combination" or "barycenter" of the extremal points $\mathcal{E}_X$. In light of these theorems, we may give the following definition:

Definition 1.34. A measure space $(X, \mathcal{B}, T)$ is **uniquely ergodic** if one of the following equivalent conditions hold:

1. $\mathcal{M}_X$ consists of one element.
2. $\mathcal{E}_X$ consists of one element.

Back to sequences and shifts. This will be useful when proving unique ergodicity of shift spaces.

Proposition 1.35. *If $u \in \mathcal{A}^{\mathbb{N}}$ and X_u is the associated shift, then X_u is uniquely ergodic if and only if for each $W \in \mathcal{A}^*$,*

$$f_k(W) := \lim_{N\to\infty} \delta(W, u|_k^{k+N-1}) \tag{43}$$

exists for each k, agrees for each k and converges uniformly in k.

The unique invariant measure μ satisfies $\mu([W]) = f(W)$, where $f(W)$ is the common limit above.

Proof. First assume $f(W)$ is well defined and is the uniform convergent limit of Equation (43). Then for any $\mu \in \mathcal{E}_X$, there exists $w \in X_u$ so that w is generic for μ. For each n, there exists k_n so that $w|_0^n = u|_{k_n}^{k_n+n}$, therefore for each $W \in \mathcal{A}^*$,

$$\mu([W]) = \lim_{n\to\infty} \delta(W, w|_0^n) = \lim_{n\to\infty} \delta(W, u|_{k_n}^{k_n+n}) = f(W). \tag{44}$$

Therefore any ergodic measure agrees with f on all cylinders. Therefore they are all equal, or $\#\mathcal{E}_X = 1$.

Now assume that Equation (43) does not converge uniformly, then there are choices $W \in \mathcal{A}^*$, $\ell_n \nearrow \infty$, k_n and j_n so that

$$\lim_{n\to\infty} \delta(W, u|_{k_n}^{k_n+\ell_n}) > \lim_{n\to\infty} \delta(W, u|_{j_n}^{j_n+\ell_n}). \tag{45}$$

Choose a subsequence $m_n \nearrow \infty$ so that $\mu_1 \in \mathcal{M}_X$ is given by

$$\mu_1([V]) = \lim_{n\to\infty} \delta(V, u|_{k_{m_n}}^{k_{m_n}+\ell_{m_n}}) \tag{46}$$

for each $V \in \mathcal{A}^*$. Likewise we choose a subsequence $m'_n \nearrow \infty$ so that

$$\mu_2([V]) = \lim_{n\to\infty} \delta(V, u|_{j_{m'_n}}^{j_{m'_n}+\ell_{m'_n}}) \tag{47}$$

for each $V \in \mathcal{A}^*$ defines $\mu_2 \in \mathcal{M}_X$. Because $\mu_1(W) > \mu_2(W)$, $\mu_1 \neq \mu_2$ and therefore $\mathcal{M}_X$ contains at least two elements.

$\square$

1.4 RJ: Substitutions

Lecture date: 2014/03/03.

A **substitution** σ is an application from an alphabet $\mathcal{A}$ into $\mathcal{A}^\star - \{\varepsilon\}$, the set of non-empty finite words on $\mathcal{A}$. We can extend substitutions to maps from $\mathcal{A}^\star \to \mathcal{A}^\star$ by concatenation. In general, if W, W' are words of $\mathcal{A}^\star$, then

$$\sigma(WW') = \sigma(W)\sigma(W'). \tag{48}$$

Similar extensions can be made to $\mathcal{A}^\mathbb{N}$ and $\mathcal{A}^\mathbb{Z}$.

Example 1.36. Let $\mathcal{A} = \{a, b\}$ and let σ be given by $\sigma(a) = aba$ and $\sigma(b) = baa$. Here, we have

$$\sigma(a)\sigma(b) = ababaa = \sigma(ab). \tag{49}$$

Remark 1.37. Substitutions are continuous maps from $\mathcal{A}^\mathbb{N} \to \mathcal{A}^\mathbb{N}$ (the same is true for $\mathcal{A}^\mathbb{Z} \to \mathcal{A}^\mathbb{Z}$). Suppose for two sequences u and v we have $d(u, v) \leq 2^{-n}$. Then we know u and v agree on at least the first n letters. An application of σ will map these n letters to the same word W which has length at least n, and so $\sigma(u)$ and $\sigma(v)$ will agree on at least the first n letters. Thus, we have $d(\sigma(u), \sigma(v)) \leq 2^{-n}$.

Definition 1.38. An (infinite) sequence u of a substitution σ is called a **fixed point** if we have $\sigma(u) = u$.

Proposition 1.39. *Let $\mathcal{A}$ be an alphabet, σ a substitution over $\mathcal{A}$ and $a \in \mathcal{A}$ a letter such that $|\sigma(a)| \geq 2$ and $\sigma(a) = a*$ — that is, $\sigma(a)$ begins with a. Then there exists a fixed point u of σ that begins with a.*

Proof. Let $W(0) = a$ and $W(n+1) = \sigma(W(n))$. Then we have

$$W(n+1) = \sigma(W(n)) = \sigma^{n+1}(a) = \sigma^n(aba) = W(n)\sigma^n(b)\sigma^n(a). \tag{50}$$

This shows that the sequence $W(n+1)$ begins with the sequence $W(n)$. If we define a sequence of sets by

$$\mathcal{C}_n = [W(n)],$$

the cylinder defined by the sequence $W(n)$, then we have $\mathcal{C}_{n+1} \subset \mathcal{C}_n$, because if $v \in \mathcal{C}_{n+1}$, then v must begin with the sequence $W(n)$, and so $v \in \mathcal{C}_n$. Note that cylinders are clopen sets in our topology. Because our topological space is compact, closed subsets are also compact. Thus, we have that $\mathcal{C}_n$ is a compact set, and so

$$\cdots \subset \mathcal{C}_{n+1} \subset \mathcal{C}_n \subset \mathcal{C}_{n-1} \subset \cdots \mathcal{C}_0$$

is a sequence of nested, nonempty compact sets. By Cantor's intersection theorem, we have

$$\mathcal{I} = \bigcap_{n=0}^{\infty} \mathcal{C}_n \neq \emptyset.$$

Furthermore, because we have $|\sigma(a)| \geq 2$, the length of the sequence $W(n)$ is growing large as $n \to \infty$. Thus, the diameter $\operatorname{diam}([W(n)]) \to 0$ $n \to \infty$. Thus, because $\mathcal{I}$ is nonempty, we have $|\mathcal{I}| = 1$. To conclude, note that if we have $u \in \mathcal{C}_n$, then $\sigma(u) \in \mathcal{C}_{n+1}$, and thus $\sigma(u) \in \mathcal{C}_n$ because $\mathcal{C}_{n+1} \subset \mathcal{C}_n$. Therefore, because $\mathcal{I}$ contains one element, it must be a sequence such that $\sigma(u) = u$. $\square$

We now turn our discussion to the incidence matrix of a substitution.

Definition 1.40. Let $\mathcal{A} = \{a_1, \ldots, a_d\}$ and let σ be a substitution over $\mathcal{A}$. The **incidence matrix** M_σ of σ is the $d \times d$ matrix with entries

$$M_{\sigma_{(i,j)}} = |\sigma(a_j)|_{a_i},$$

where $|\sigma(a_j)|_{a_i}$ is the number of occurrences of a_i in $\sigma(a_j)$.

Example 1.41. Let $\mathcal{A} = \{a, b\}$ and let $\sigma(a) = aba$, $\sigma(b) = baa$ as before. Then the incidence matrix of σ is

$$M_\sigma = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}.$$

Let $\mathbf{1} : \mathcal{A}^\star \to \mathbb{Z}^d$ denote the **canonical homomorphism**, defined for all $W \in \mathcal{A}^\star$ by

$$\mathbf{1}(W) = (|w|_{a_i})_{1 \leq i \leq d} \in \mathbb{N}^d. \tag{51}$$

The canonical homomorphism is useful because we have

$$M_\sigma = \left(\mathbf{1}\sigma(a_1)|\cdots|\mathbf{1}\sigma(a_d)\right), \tag{52}$$

which gives the following commutative relation for all $W \in \mathcal{A}^\star$:

$$\mathbf{1}(\sigma(W)) = M_\sigma \mathbf{1}(W). \tag{53}$$

1.5 RJ: Primitivity

Lecture date: 2014/03/03.

Definition 1.42. A substitution σ is **primitive** if for all $a, b \in \mathcal{A}$, there exists some $k > 0$ such that a occurs in $\sigma^k(b)$.

Primitivity is a nice property because of its close relationship to minimality.

Proposition 1.43. *If u is a fixed point of a primitive substitution σ, then u is minimal (or uniformly recurrent as given in Lemma 1.13).*

Proof. Let $u = au_1u_2\cdots$ be the fixed point beginning with a. For any $k > 0$, we have

$$u = \sigma^k(u) = \sigma^k(a)\sigma^k(u_1)\cdots. \tag{54}$$

Because σ is primitive, there exists some $K > 0$ such that a occurs in $\sigma^K(b)$. Thus, we know that a occurs in $\sigma(u_1)$ for each i and, obviously, $\sigma(a)$. If we let

$$N = 2 \cdot \max_{b \in \mathcal{A}} |\sigma^K(b)|,$$

we know that a appears in every factor of size N in u. But then, so does every $\sigma^n(a)$ in $u = \sigma^n(u)$, hence so does every word in u. More precisely, if we let

$$N' = 2 \cdot \max_{b \in \mathcal{A}} |\sigma^{K+n}(b)|,$$

then any factor of size N' contains $\sigma^n(a)$ as a factor. Because u begins with $\sigma^M(a)$ for any M, we can choose n large enough such that $\sigma^n(a)$ contains any factor of u. Thus, u is minimal because every factor is recurrent with bounded gaps. □

A characterization of primitivity. Let us now add the following three assumptions on substitutions.

(1) There exists a letter $a \in \mathcal{A}$ such that $\sigma(a) = a*$.

(2) $\lim_{n\to\infty} |\sigma^n(b)| = \infty$ for all $b \in \mathcal{A}$.
Note: As we saw earlier, assumptions (1) and (2) guarantee the existence of an infinite fixed point beginning with a.

(3) All letters in $\mathcal{A}$ occur in the fixed point beginning with a.

Proposition 1.44. *Let σ be a substitution satisfying* (1), (2), (3). *The substitution is primitive if and only if the fixed point u of σ beginning with a is minimal.*

Proof. Proposition 1.43 proves the forward direction. Suppose u is minimal. Every factor of u appears in u with bounded gaps. By assumption (3), every letter b occurs in $u = \sigma(u)$ so that $\sigma^k(b)$ is a factor of u for every $b \in \mathcal{A}$ and $k \geq 1$. For K large, because of assumption (2) we know that $\sigma^K(b)$ must contain a. Because $\sigma^K(b)$ contains a and, by assumption (3) $\sigma^N(a)$ contains every letter of $\mathcal{A}$, we have that $\sigma^{K+N}(b)$ contains $\sigma^N(a)$ as a factor and, thus, every letter c of $\mathcal{A}$ is contained in $\sigma^{K+N}(b)$ for each $b, c \in \mathcal{A}$. Thus σ is primitive. □

2 Chapter 5

2.1 PR: The Morse Sequence and its properties

Lecture date: 2014/03/10.

The Morse sequence, as described in Chapter 2 of the text, is motivated by the desire to create a minimal word on {0, 1} which is neither periodic nor ultimately periodic. We recall that a minimal sequence is one in which every word of finite length appears with gaps of bounded size. We can define it in three ways.

Definition 2.1. The Morse sequence can be defined using an **arithmetic definition**: for the binary digits of the binary representation of the natural number x, let x_j be the jth digit, which ranges from 0 to n. Let

$$f(x) = \sum_{j=0}^{n} x_j,$$

and let M be the infinite word on $\{0,1\}^{\mathbb{N}}$ such that

$$M_n = \begin{cases} 0 & , f(x) \text{ is even} \\ 1 & , f(x) \text{ is odd.} \end{cases}$$

Definition 2.2. The Morse sequence can also be defined using an **iterative limit process**: we start with 0, and at each step, append the bitwise negation of the current sequence to its end. We define the **bitwise negation**, or complement, of a sequence s_n on $\{0,1\}^{\mathbb{N}}$ as the sequence $t_n = 1 - s_n$. The first few steps of this process are shown below.

$$0(1) \to 01(10) \to 0110(1001) \to 01101001(10010110) \to \ldots$$

Definition 2.3. Finally, the Morse sequence can be defined using a **set of substitution rules**:

$$\sigma : \{0,1\} \to \{0,1\}^2 : \tag{55}$$
$$\sigma(0) = 01, \tag{56}$$
$$\sigma(1) = 10. \tag{57}$$

We call any word of the form $\sigma^n(0)$ or $\sigma^n(1)$ an **n-word**. It is important to note two things: first, that $\sigma^n(0) = \sigma^{n-1}(0)\sigma^{n-1}(1)$, and second, that $\sigma(u) = u$. See Proposition 1.39 for a discussion about fixed points.

Proposition 2.4. *u, the Morse Sequence, is minimal, and is neither periodic nor ultimately periodic.*

Proof. By Proposition 1.43, u is minimal as σ is primitive.

Referring to the binary digit sum definition of the Morse sequence, we know that $u_{2n+1} \neq u_n$, so $n + 1$ can't be a period, for any n, so u can't be periodic.

Further, for any p, n_0, we can always choose $n \geq n_0$ so that $n = mp - 1$ for some m, so that n_0 represents the length of the initial, nonperiodic portion of the supposedly ultimately periodic sequence, so u is not ultimately periodic either. $\square$

Definition 2.5. We define a notion of difference called the **Hamming distance**. The Hamming distance between two words of the same length, U and V, is given by

$$\overline{d}((u_1...u_n),(v_1...v_n)) = \frac{1}{n}|1 \leq i \leq n; u_i \neq v_i|, \tag{58}$$

that is, the proportion of letters at which the two sequences of equal length disagree.

Proposition 2.6. *If U, V are two n-words, and if W is a word of length 2^{n+1} which occurs in u at position i, and $\overline{d}(UV,W) < \frac{1}{4}$, then $W = UV$, and i is a multiple of 2^n.*

Proof. The proof is immediate for $n = 0$. Suppose the proposition is true for $n-1$. We check the case for n. For U,V,W,i as above, $W = W_1W_2$, with either $\overline{d}(U,W_1) < \frac{1}{4}$, or $\overline{d}(V,W_2) < \frac{1}{4}$. Because both U and V are each comprised of two $(n-1)$-words, the induction hypothesis implies that there exists some integer k such that $i = k * 2^{n-1}$. We can define $W_1 = W_{11}W_{12}$, $W_2 = W_{21}W_{22}$, $U = U_1U_2$, $V = V_1V_2$. Then $W_{11}, W_{12}, W_{21}, W_{22}$ are all $(n-1)$-words, and must be at Hamming distance at most $\frac{1}{2}$ from U_1, U_2, V_1, V_2 respectively; since the distance between any two m-words must be either 0 or 1, we must have $W_1 = U$, $W_2 = V$.

It remains to be proven that k is even. Suppose k is odd. For a word S be the **complement** of S, defined as the word obtained by replacing every instance of 0 with 1 and 1 with 0 in S, be denoted by S'. Then let $W = abb'c$, $U = U_1U_1{}'$, $V = V_1V_1{}'$. Then we must have that $a' = b = c$, so $a'abb'cc' = bb'bb'bb'$ would have to occur in u, which is impossible, as neither 000 nor 111 occurs, by inspection, so neither can 010101 or 101010. Thus k must be even. $\square$

We call a word **squarefree** if no word in the language of that word is of the form WW for any word W. We call a word **cubefree** if no word in the language of that word is of the form WWW for any word W.

It seems like we should be able to prove that the Morse sequence is cubefree. We can in fact do better! We say that a word is **free of powers** $2+\epsilon$ if for any word V, with initial letter v, the sequence VVv never appears in the word.

Proposition 2.7. *The Morse Sequence is free of powers $2+\epsilon$.*

Proof. Let $l = |V|$. First, we consider the possibility that l is odd and that VVv occurs starting at an odd position in u. Let $V = va_1b_1...a_pb_p$. We can put bars to indicate the separation of the 1-words composing V like so: $V = v|a_1b_1|...|a_pb_p|$. Then we know that the sequence $VVv = v|a_1b_1|...|a_pb_p|va_1|b_1a_2|...|b_pv|$ must occur in the Morse sequence, with bars as indicated. Thus exactly one of each pair can be 0, and the other must be 1: that is, $v = a_1 + b_1 = ... = a_p + b_p = v + a_1 = ... = b_p + v = 1$, and $v + a_1 = b_p + v$, which is impossible, because then every a_n and b_n would have to be 0. The same reasoning applies for an occurrence of an odd-length V at an even position, with the case that for V of length 1, 000 and 111 are both impossible.

For even l and occurence of VVv at some position $2k$, then we have $V = |vb_1|...|a_pb_p|$. For $W = va_2..a_p$, $\sigma(WWv)$ appears in the Morse sequence at $2k$, so WWv must appear at position k. Then v is the first letter of W, and $|W| < |V|$. But this is impossible: we can follow the same chain of reasoning to find that another, smaller sequence of the form UUv must also appear at position $\frac{k}{2}$, and so on, infinitely, which is impossible. The same reasoning applies for the occurrence of VVv in an odd position. □

Remark 2.8. Interestingly, we can connect the cubefreeness of the Morse sequence to the first attempt at determining a draw in chess, which was draw by threefold repetition of move sequence. If we used this definition, the fact that the Morse sequence is cubefree implies an infinite game of chess without threefold repetition would be possible from the very beginning of the game. To see this, let black and white both make moves according to the Morse sequence such that a 0 corresponds to a queenside knight move and a 1 corresponds to a kingside knight move. Because the Morse sequence is cubefree, this sequence of moves will never have a threefold repetition, and the game can continue forever. This is why chess now uses draw by threefold repetition of only position, rather than move sequence: because there are only a finite number of possible boardstates in chess, no game can last for a number of moves longer than twice the number of boardstates.

Proposition 2.9. *The sequence defined by the difference of successive terms of the Morse Sequence is an infinite squarefree word on three letters.*

Proof. First, if the word $AA = a_1...a_ka_1...a_k$ appears at position i in v, then since $u_{2k+i} - u_i = 2 * (a_1 + ... + a_k)$ is in $\{-1, 0, 1\}$, $(a_1 + ... + a_k) = 0$. But then for e in $\{0, 1\}$, we would have to have the sequence in v consisting of $e(e + a_1)...(e + a_1 + ... + a_{k-1})$ in v, which is impossible. □

Proposition 2.10. *Any factor of the Morse Sequence of size at least five has a unique decomposition into 1-words, possibly with another letter before or after it.*

Proof. In essence, what this means is that there is only one way to draw the bars used in the last proof. If our word contains a 00 or 11, we have a natural place to draw a bar, and the proposition is true; otherwise, the word would have to contain 01010 or 10101, and since the Morse sequence is free of powers $2 + \varepsilon$, neither of these sequences appear. □

Recall that we define a **complexity function** $p_u(n)$ of a sequence u as the function which maps each natural number to $|L_n(u)|$, that is, the number of subwords of size n in u.

Proposition 2.11. *For u, the Morse Sequence, $p_u(n)$ is given by $p_u(1) = 2$, $p_u(2) = 4$, and for $n \geq 3$, if $n = 2^r + q + 1$, for natural r, and $0 < q \leq 2^r$, then $p_u(n) = 3 * 2^r + 4q$ if $0 < q \leq 2^{r-1}$, and $p_u(n) = 2^{r+2} + 2q$ if $2^{r-1} < q \leq 2^r$.*

Proof. We start by noticing that the previous proof implies that $p_u(n) = p_0(n) + p_1(n)$, where we define $p_0(n)$ as the number of words appear immediately after a bar, and $p_1(n)$ is the number of words appearing immediately before a bar. Let $n = 2k+1$ for a word W such that $|W| = n$. If W is in the first category, then $W = |a_1b_1|...|a_kb_k|a_{k+1}$, with $a_i + b_i = 1$, and there are $k+1$ such words, so $p_0(2k+1) = p_u(k+1)$. Similarly, $p_1(2k+1) = p_u(k+1)$, $p_0(2k) = p_u(k)$, and $p_1(2k) = p_u(k+1)$, and thus $p_u(2k+1) = 2 * p_u(k+1)$, $p_u(2k) = p_u(k) = p_u(k+1)$.

We can then do induction on q: assume the formulae as given in the proposition are true for some $q = 2m$. Then for $q = 2m+1$, $p_u(n) = 2*p_u(2^{r-1}+m+2)$. If $0 < q \leq 2^{r-1}$, then $0 < m \leq 2^{r-2}$, and we have

$$2 * p_u(2^{r-1} + m + 2) = 12 * 2^{r-2} + 8m + 4 = 6 * 2^{r-1} + 4 * (2q + 1).$$

Similarly, for the induction hypothesis $q = 2m - 1$, $0 < q \leq 2^{r-1}$, for $q = 2m$ we have

$$\begin{aligned} p_u(n) &= p_u(2^{r-1} + (m+1) - 1) + p_u(2^{r-1} + m + 1) \\ &= 6 * 2^{r-2} + 4m + 4 - 4 + 6 * r^{r-2} + 4m \\ &= 6 * 2^{r-1} + 4q. \end{aligned}$$

Now suppose instead that $2^{r-1} < q \leq 2^r$. Then going through the induction hypothesis as before for odd q, we will have $2^{r-2} < q \leq 2^{r-1}$, and thus

$$2 * p_u(2^{r-1} + m + 2) = 16 * 2^{r-2} + 4m + 2 = 8 * 2^{r-1} + 2q,$$

and similarly, for even q, we will have

$$\begin{aligned} p_u(n) &= p_u(2^{r-1} + (m+1) - 1) + p_u(2^{r-1} + m + 1) \\ &= 8 * 2^{r-2} + 2m + 2 - 2 + 8 * 2^{r-2} + 2m = 8 * 2^{r-1} + 2q. \end{aligned}$$

□

In all cases, we note that $p_u(n) < 4n$, and since by the induction, $p_u(n+1) - p_u(n)$ will only ever be 4 or 2, that expression indeed takes only two values.

2.2 KD: Measure-theoretic Look at the Morse Sequence and Topological Congujacy

Lecture date: 2014/03/10.

These lecture notes will be divided in two parts: those relating to the main result of section 5.1 in the book, and results which will be used by people who speak later.

2.2.1 KD: A theorem in measure theory regarding the Morse Sequence

As before, unless otherwise noted we will denote the Morse sequence by u and its orbit closure as X_u.

Theorem 2.12. *(Main result) There exists a measure μ such that the measure-theoretic dynamical system (X_u, S, μ) is ergodic. Furthermore, with regards to that measure, there exists a $P \subset X_u$ such that $\mu(P) = 1$ and the shift S takes P to P bijectively.*

We will establish the proof of this result to the end of this section; we will now prepare the ground with a few propositions. Our first job is to construct the set P, and we will then establish the measure μ and show $\mu(P) = 1$.

Proposition 2.13. *If v is a recurrent sequence, then the shift S is surjective onto X_v.*

Proof. Suppose w is a sequence in X_v where v is a recurrent sequence and $w = w_0w_1w_2...$ Put $W_i = w_0w_1w_2...w_i$ for $i \geq 0$. As $w \in X_v$ by a lemma from previous lectures, we know that $\Lambda(w) \subset \Lambda(v)$ (here, $\Lambda(v)$ is the language of v). Hence as v is recurrent, for each i, W_i occurs infinitely many times in v. We can then find indices $j_{0,i} < j_{1,i} < j_{2,i}...$ such that W_i occurs in v at positions $j_{k,i}$. Then, we can easily see that for $k \geq 1$, $j_{k,i} \geq 1$. Now, consider the sequence $v_{j_{1,i}-1}, v_{j_{2,i}-1}, v_{j_{3,i}-1}...$ (these are the letters which precede W_i in v). As all of the elements of that sequence come from a finite alphabet, at least one element from that sequence occurs infinitely many times; label one such element as a_i.

To sum up so far, we have found a sequence a_i such that for each i, a_iW_i occurs infinitely many times in v. Again, the elements of the sequence $a_0, a_1, ...$ come from a finite alphabet, so there exists an element of that alphabet which occurs infinitely many times in that sequence. Label that element as a. Hence, for arbitrarily big i, aW_i occurs in v; this means that for all i, aW_i occurs in v. Now observe that $\Lambda(aw) = \Lambda(w) \cup \{aW_0, aW_1, aW_2, ...\}$. We already know that $\Lambda(w) \subset \Lambda(v)$, and we just proved that for all i, aW_i is in $\Lambda(v)$. Hence, $\Lambda(aw) \subset \Lambda(v)$, which establishes $aw \in X_v$. To conclude the proof of the lemma, remember that $S(aw) = w$, and as w was a random element of X_v we have proven that the shift is surjective onto X_v. □

Proposition 2.14. *Let v be a recurrent sequence, such that for all $n \geq 0$ $p_u(n) \leq Cn$ for some constant C. Denote by F the subset of all elements of X_v*

which have more than one sequence in their preimage under S i.e. $F = \{w \in X_u : |S^{-1}(w)| \geq 2\}$. Then, F is a finite set and $|F| \leq C+1$.

Proof. First, we will prove a helpful fact. Suppose the set $\{n \in \mathbb{N} : p_u(n+1) - p_u(n) \leq C+1\}$ is bounded i.e. $\exists M : \forall n \geq M, p_u(n+1) - p_u(n) > C+1$. Then observe that for all $k \geq 0$,

$$p_u(M+k) - p_u(M) = \sum_{t=0}^{k-1} p_u(M+t+1) - p_u(M+t) >$$

$$> \sum_{t=0}^{k-1}(C+1) = k(C+1).$$

Dividing the leftmost and rightmost sides of that inequality by $M+k$, we see that

$$\frac{p_u(M+K) - p_u(M)}{M+K} > \frac{k(C+1)}{M+K}.$$

Now, we let k go to infinity; as the $p_u(M)$ term is constant and $\lim_{k\to\infty} \frac{k(C+1)}{M+k} = C+1$ we see that

$$\liminf_{k\to\infty} \frac{p_u(M+K)}{M+K} \geq C+1.$$

This clearly contradicts our assumption that $p_u(n) \leq Cn$ for all n. Hence, the set $\{n \in \mathbb{N} : p_u(n+1) - p_u(n) \leq C+1\}$ is unbounded.

Now, suppose that n is such that $p_u(n+1) - p_u(n) \leq C+1$. As v is recurrent, each word from $\Lambda_n(v)$ occurs in v at least one in a position different than the leftmost (0-th position). Hence, for each word $W \in \Lambda_n(v)$ we can find $a_W : a_W W \in \Lambda_{n+1}(v)$; denote by F_{n+1} the set of all $W \in \Lambda_n(v)$ such that the choice of a_W is not unique i.e

$$F_{n+1} = \{W \in \Lambda_n(v) : \exists a, b \in \mathbb{A}, a \neq b, aW \in \Lambda_{n+1}(v), bW \in \Lambda_{n+1}(v)\}.$$

Clearly, as $p_u(n+1) - p_u(n) \leq C+1$, F_{n+1} is finite and $|F_{n+1}| \leq C+1$ (else we would have more than $p_u(n) + C + 1$ words in $p_u(n+1)$).

Finally, let F be as defined in the statement of our proposition, and suppose that $|F| > C+1$. Hence, we can extract more than $C+1$ sequences in X_v with more than 1 preimage under S; those sequences will have to all have different "beginnings" after the first T letters for some big T such that $p_u(T+1) - p_u(T) \leq C+1$ (by beginnings, we mean the words formed by the first T letters of the sequence). Denote a subset of those beginning words of length T by $W_1, W_2, ...W_N$ where $N > C+1$. Yet, as for each $1 \leq i \leq T$, W_i is the beginning of a sequence in F, W_i will have to belong to F_{T+1}. As $|F_{T+1}| \leq C+1$, and the W_i are $N > C+1$ distinct words in F_{T+1}, we get a contradiction with our assumption for the size of F. Hence, F is finite and $|F| \leq C+1$. □

X_u is a recurrent sequence, which by Proposition 5.1.9 (or the last proposition from Paul's talk) satisfies the conditions of all of our propositions so far (for

$C = 4$). Hence, we now proceed to define the set P, as previously mentioned in Theorem 2.12, and to prove that P differs from X_u by a countable set and the shift S defines a bijection from P to P.

Definition 2.15. Define $F \in X_u$ as in Proposition 2.14 and define

$$P = X_u \setminus \cup_{n\in\mathbb{Z}} S^n F.$$

Lemma 2.16. *P differs from X_u by a countable set, and S defines a bijection from P to P.*

Proof. As F is finite, clearly for $n \geq 0$, $|F| \geq |S^n F|$. Also, as we are operating on a finite alphabet, for $n < 0$, $|F||\mathbb{A}|^{|n|} \geq |S^n F|$ (the LHS is simply the number of words of length $|n|$ over the alphabet $\mathbb{A}$). Hence, $X_u \setminus P = \cup_{n\in\mathbb{Z}} S^n F$ is the countable union of sets with finitely many elements, and is hence countable.

Next, we will show that $S(P) \subset P$. Indeed, suppose $\exists \alpha \in X_u, \alpha \notin P$: $\exists w \in P, S(w) = \alpha$. Hence, $\alpha \in \cup_{n\in\mathbb{Z}} S^n F$ and then $w \in S^{-1}(\cup_{n\in\mathbb{Z}} S^n F) = \cup_{n\in\mathbb{Z}} S^n F$, which contradicts $w \in P$. Hence, $S(P) =\subset P$. It is clear that that $S^{-1}(P) \cap X_u \subset P$; if $\alpha \in \cup_{n\in\mathbb{Z}} S^n F = X_u \setminus P$, $S(\alpha) \in \cup_{n\in\mathbb{Z}} S^n F$. As $S(P) \subset P$ and $S^{-1}(P) \subset P$, by Proposition 2.13, $S(P) = P$ surjectively.

Finally, we will establish that S restricted to P is injective. Indeed, suppose $w_1, w_2 \in P, w_1 \neq w_2 : S(w_1) = S(w_2) = \alpha$. Hence, by the definition of F from before, $\alpha \in F \subset \cup_{n\in\mathbb{Z}} S^n F$ i.e. $\alpha \notin P$. Still, we just proved that if $w \in P$, $S(w) \in P$; we reach a contradiction about the existence of such w_1, w_2 and α. Then, the shift is injective on P, which concludes the proof of our lemma. ☐

We now reach the measure-theoretic content of the talk. The point of the next few technical propositions is to prove that (X_u, S) is uniquely ergodic with sole ergodic measure μ.

Proposition 2.17. *Let σ be the Morse substitution and $N(W, V)$ be the number of occurrences of the word W in the word V. Then, for any factor W, the following limits exist and are equal:*

$$\lim_{n\to\infty} \frac{N(W, \sigma^n(0))}{2^n} = \lim_{n\to\infty} \frac{N(W, \sigma^n(1))}{2^n}.$$

Proof. Define the sequence

$$a_n = \frac{N(W, \sigma^n(0)) + N(W, \sigma^n(1))}{2^{n+1}}.$$

We now know that $\sigma^{n+1}(0) = \sigma^n(0)\sigma^n(1)$ and $\sigma^{n+1}(1) = \sigma^n(1)\sigma^n(0)$. Hence for any factor W and $e = \{0, 1\}$,

$$2^{n+1}a_n + |W| = N(W, \sigma^n(0)) + N(W, \sigma^n(1)) + |W| \geq$$

$$\geq N(W, \sigma^{n+1}(e)) \geq$$

$$\geq N(W, \sigma^n(0)) + N(W, \sigma^n(1)) = 2^{n+1}a_n$$

(the lower inequality is obvious from the decomposition of $\sigma^{n+1}(e)$ into $\sigma^n(0)$ and $\sigma^n(1)$; the upper inequality comes from the same decomposition and the fact that there may be no more than $|W|$ occurrences of W in $\sigma^{n+1}(e)$ which start in the first half of the string and end in the second half).

Summing the corresponding lower inequalities for $e = 0$ and $e = 1$ and substituting by the sequence a_n where we can, we see that for all n,

$$2^{n+2}a_{n+1} = N(W, \sigma^{n+1}(0)) + N(W, \sigma^{n+1}(1)) \geq 2^{n+2}a_n.$$

It remains for us to note that $N(W, \sigma^n(e)) \leq |\sigma^n(e)| = 2^n$, which means that $a_n \leq \frac{2^n+2^n}{2^{n+1}} = 1$. Then, the sequence a_n is non-decreasing and bounded, and hence for each W, that sequence converges to some limit; denote it by f_W. Now, we reiterate from above we have that for $e = \{0, 1\}$,

$$a_n + \frac{|W|}{2^{n+1}} \geq \frac{N(W, \sigma^{n+1}(e))}{2^{n+1}} \geq a_n;$$

since W is fixed, $|W|$ is constant and hence as we let n go to infinity, by the sandwich theorem we get that $\lim_{n\to\infty} \frac{N(W,\sigma^n(e))}{2^n} = f_W$, which concludes the proof of the proposition. $\square$

Proposition 2.18. *For any factor W, $\lim_{n\to\infty} \frac{N(W,u_k u_{k+1}...u_{k+n})}{n+1} = f_W$ uniformly in k.*

Proof. Fix k and write $V = u_k u_{k+1}...u_{k+n}$. Hence, by the recursive properties of the Morse sequence, for each p we can write

$$V = A\sigma^p(u_j)\sigma^p(u_{j+1})...\sigma^p(u_{j+l-1})B$$

where $|A| < 2^p$ and $|B| < 2^p$. We then have $n + 1 = l2^p + |A| + |B|$ and $l2^p \leq n + 1 \leq (l + 2)2^p$. By the decomposition of V into at most $(l + 2)$ strings of length at most 2^p, we can employ considerations similar to those in our previous proposition to obtain the following sequence of inequalities:

$$\sum_{i=j}^{j+l-1} N(W, \sigma^p(u_i)) \leq N(W, V) \leq \sum_{i=j}^{j+l-1} N(W, \sigma^p(u_i)) + |A| + |B| + (l + 1)|W|.$$

These are true for all n and all p, with l defined accordingly. We will consider the left inequality and the right inequality separately; for both cases, fix $\epsilon > 0$ and pick p big enough so

$$|N(W, \sigma^p(e)) - f_W| < \epsilon.$$

For the left inequality, observe that

$$\frac{N(W, V)}{n + 1} \geq \frac{N(W, V}{(l + 2)2^p} \geq \frac{1}{l + 2} \sum_{i=j}^{j+l-1} \frac{N(W, \sigma^p(u_i))}{2^p} \geq \frac{l}{l + 2} f_w - \epsilon.$$

For the right inequality, see that

$$\frac{N(W,V)}{n+1} \leq \frac{N(W,V)}{l2^p} \leq \frac{1}{l}\left(\sum_{i=j}^{j+l-1} \frac{N(W,\sigma^p(u_i))}{2^p} + \frac{|A|+|B|}{2^p} + \frac{(l+1)|W|}{2^p}\right) \leq$$

$$\leq \frac{1}{l}\left(l(f_W+\epsilon) + 2 + \frac{(l+1)|W|}{2^p}\right) \leq f_W + \epsilon + \frac{2}{l} + \frac{(l+1)|W|}{l2^p}$$

$$\leq f_W + \epsilon + \frac{2}{l} + \frac{|W|}{2^{p-1}}.$$

Combining both results, we see that for p big enough so that $\frac{|W|}{2^{p-1}} < \epsilon$,

$$\frac{l}{l+2} f_w - \epsilon \leq \frac{N(W,V)}{n+1} \leq f_W + 2\epsilon + \frac{2}{l}.$$

Clearly with ϵ (and p big enough) fixed, l is monotonically increasing in n and goes to infinity as n goes to infinity; pick n big enough so that $\frac{2}{l} < \epsilon$ (here, we will use that $(l+2)2^p \geq n+1$ or $l \geq \frac{n+1-2^{p+1}}{2^p}$; thus, $\frac{2}{l} \leq \frac{2^{p+1}}{n+1-2^{p+1}}$). Hence, as $f_W \leq 1$, for n big enough,

$$\frac{l}{l+2} f_W = f_W - \frac{2}{l+2} f_W \geq f_W - \frac{2}{l+2} \geq f_W - \frac{2}{l} \geq f_W - \epsilon.$$

Hence, for k, W fixed and n big enough (choice depending only on ϵ), we get

$$f_W - 2\epsilon \leq \frac{N(W,V)}{n+1} \leq f_W + 3\epsilon;$$

this clearly establishes the convergence of $\frac{N(W,V)}{n+1}$ uniformly in k to f_W. □

Corollary 2.19. *The system (X_u, S) is uniquely ergodic with measure μ which satisfies $\mu([W]) = f_W$ for each factor W. As a consequence, the measure preserving system (X_u, S, μ) is ergodic.*

Proof. The first point of the corollary is a direct consequence of Proposition 2.18 and a theorem from Jon's talks (Theorem 5.1.21 from the book). The second point is also a direct consequence of the theorem that if there is a unique S-invariant measure, the system defined by it is ergodic. □

We will now establish that the measure μ is non-atomic.

Proposition 2.20. *If W is a word of length l occurring in the Morse sequence, $\mu([W]) < \frac{6}{l}$.*

Proof. Fix $p, n \in \mathbb{N}, p > n$, $e = 0, 1$ and $f = 0, 1$ and put $H = \sigma^p(0)$, $U = \sigma^n(e)$ and $V = \sigma^n(f)$. By a proposition in Paul's talk (Proposition 5.1.4 in the book), UV can occur in H only in positions divisible by 2^n; in H there are exactly 2^{p-n} such positions and hence we get

$$N(UV, H) \leq 2^{p-n}.$$

Now, remember that how we defined μ:

$$\mu([UV]) = f_{UV} = \lim_{p\to\infty} \frac{N(UV,H)}{|H|} = \lim_{p\to\infty} \frac{N(UV,H)}{2^p} \leq 2^{-n}.$$

Fix W, put $l = |W|$. If $l \leq 5$ we are trivially done; if $l \geq 6$, find $n \geq 1$ such that $3.2^n \leq l < 3.2^{n+1} = 6.2^n$. Hence, as $l \geq 3.2^n$, W will contain at least one block which looks like $\sigma^n(e)\sigma^n(f)$ (to see that, find an occurrence of W in u and then divide u into blocks of size 2^n starting from the start; at least 2 consecutive blocks will fall within W, and those will correspond to $\sigma^n(e)\sigma^n(f)$ for some $e, f \in 0, 1$). Hence,

$$\mu(W) \leq \mu(\sigma^n(e)\sigma^n(f)) \leq 2^{-n} < \frac{6}{l}$$

(we use the upper bound $l < 6.2^n$) as desired. □

Corollary 2.21. *If $w \in X_u$, $\mu(w) = 0$.*

Proof. If $W_n = w_0w_1...w_{n-1}$ is the word composed of the first n letters of $w \in X_u$, then by the properties of the measure and Proposition 2.20 we see

$$\mu(w) \leq \mu([W_n]) < \frac{6}{n};$$

letting n go to infinity yields the desired result. □

We are now ready to prove Theorem 2.12.

Proof. The existence of $\mu : (X_u, S, \mu)$ is ergodic is the statement of Corollary 2.19. With P as constructed for the purposes of Lemma 2.16, all we have left to prove is that $\mu(P) = 1$. By Lemma 2.16, P differs from X_u by a countable set. As the measure is non-atomic(Corollary 2.21), the measure of that countable set is 0, which implies $\mu(P) = 1$, concluding the proof of our theorem. □

2.2.2 KD: Topological conjugacy

We now turn our attention to material which will be useful for later.

Definition 2.22. Two topological dynamical systems (X, T) and (Y, S) are topologically conjugate if there is a homeomorphism $\phi : X \to Y$ such that $\phi \circ T = S \circ \phi$.

Proposition 2.23. *Topological conjugacy preserves unique ergodicity.*

Proof. First, suppose that μ is an ergodic measure on (X, T). In this proof, we will utilize the definition of the measure on the Borel algebra; the proof using the definition as a linear functional on the set of functions is analogous. Define the measure ν on (Y, S) as $\nu(I) = \mu(\phi^{-1}(I))$ for each I in the Borel algebra of Y (by the continuity and bijectivity of ϕ and ϕ^{-1}, $\phi^{-1}(I)$ is an element of the Borel algebra of X). We then have that $\nu(Y) = \mu(\phi^{-1}(Y)) = \mu(X) = 1$ by the ergodicity of μ. The monotonicity, sub-additivity and disjoint-sum properties of ν are immediate conseruenceses of the corresponding properties of μ; the only thing we have to prove is that ν is S-invariant. Let I be an element of the Borel algebra of Y. Then, by topological conjugacy and the T-invariance of μ we have

$$\nu(S(I)) = \mu(\phi^{-1}S(I)) = \mu(T\phi^{-1}(I)) = \mu(\phi^{-1}(I)) = \nu(I),$$

as we desired. Hence, if X has a T-invariant measure, Y has a S-invariant measure.

The construction above, by the bijectivity and continuity of ϕ^{-1}, easily reveals that if μ_1 and μ_2 are two distinct ergodic measures on (X, T), then the measures $\nu_1 = \mu_1 \circ \phi^{-1}$ and $\nu_2 = \mu_2 \circ \phi^{-1}$ will be two distinct ergodic measures on (Y, S). This concludes the proof of the proposition. $\square$

We conclude this lecture with two properties of topological conjugacy which relate to our discussion of recurrent sequences.

Proposition 2.24. *Suppose u and v are two sequences on finite alphabets and (X_u, S_u) and (X_v, S_v) are topologically conjugate through the homeomorphism $\phi : X_u \to X_v$. Then, there is an integer q such that $(\phi(w))_i$ depends only on the word $w_i w_{i+1} ... w_{i+q}$.*

Proof. Define $\phi'_p : X_v \to A_v$ as the projection onto the p-th letter in the sequence; clearly these functions are continuous. Put $\phi_p : X_u \to A_v$, $\phi_p = \phi'_p \circ \phi$; clearly those maps are continuous as well.

We will first show that $\phi_0(w)$ depends only on the first "few" letters in w. Indeed, enumerate the elements of A_v by $a_1, a_2, ... a_d$. By the continuity of ϕ_0 and the fact that each a_i is both a closed and an open subset of A_v with respect to any proper metric defined on the finite set A_v, we know that for each i, $\phi_0^{-1}(a_i)$ is a set which is both open and closed. Since the space X_u is compact, the set $\phi_0^{-1}(a_i)$ can then be covered by a finite union of cylinders; denote those as $[W_{i,1}], [W_{i,2}]...[W_{i,t_i}]$. Put $q + 1 = \max_{1 \leq i \leq d} (\max_{1 \leq j \leq t_i} |W_{i,j}|)$; by the way we select the cylinders $W_{i,j}$ it is clear that ϕ_0 will depend only on the first $(q+1)$ letters of the input.

Now observe that by the definition of ϕ_p and topological conjugacy

$$\phi_p = \phi'_p \circ \phi = \phi'_0 \circ S_v^p \circ \phi = \phi'_0 \circ \phi \circ S_u^p = \phi_0 \circ S_u^p.$$

Now, as $\phi_0(w)$ depends only on the first $q+1$ letters of w, $\phi_p(w) = \phi_0 \circ S_u^p(w)$ will depend only on the first $q + 1$ letters of $S_u^p(w)$, or equivalently only on $w_p w_{p+1} ... w_{p+q}$ as desired. $\square$

Corollary 2.25. *Under the same hypothesis as in Proposition 2.24, if $p_u(n + 1) - p_u(n)$ is bounded then so is $p_v(n + 1) - p_v(n)$; if $p_u(n) \leq Cn + C'$, then $p_u(n) \leq Cn + C''$ for some constant C''.*

Proof. Without loss of generality, assume that $v = \phi(u)$. Then, by Proposition 2.24, for each k, n, the word $v_k v_{k+1} ... v_{k+n-1}$ depends only on the word

$u_k u_{k+1} ... u_{k+n+q-1}$. Hence, we obtain that there is a surjection from $\Lambda_u(n+q)$ to $\Lambda_v(n)$ and hence $p_u(n+q) \geq p_v(n)$. Analogously, as (X_v, S_v) and (X_u, S_u) are topologically conjugate through ϕ^{-1}, we can find q' such that $p_u(n) \leq p_v(n+q')$.

Now, suppose $p_u(n+1) - p_u(n) \leq M$ for all n. Hence, by the above, we get that for $n \geq q'+1$, $p_v(n+1) - p_v(n) \leq p_u(n+1+q) - p_u(n-q') \leq (q+q'+1)M$. Then, $p_v(n+1) - p_v(n)$ is bounded by

$$\max\{(q+q'+1)M, \max_{1 \leq i \leq q'} (p_v(i+1) - p_v(i))\}.$$

Suppose that $p_u(n) \leq Cn + C'$ for some constants C, C'. Then, by the above, $p_v(n) \leq p_u(n+q) \leq C(n+q) + C' = Cn + C''$, where $C'' = C' + Cq$; this point concludes the proof of our corollary. $\square$

2.3 ZQ: Preliminaries for Rokhlin Stacks

Lecture date: 2014/03/24.

This note will pave the way for the definition of rank 1 systems and the Rokhlin Stack that generates the Morse system. We will first review a few simple facts covered in previous lectures. They seem trivial, but are in fact all the ingredients that go into the results in this note. All sequences and words will be in the space X_u where u is the Morse sequence, unless otherwise noted.

Lemma 2.26. *$\forall n \in \mathbb{N}$, $u_{2n} = u_n$ and $u_{2n+1} = 1 - u_n$.*
At position $k2^n$ of the sequence occurs $\sigma^n(0)$ if $u_k = 0$, and $\sigma^n(1)$ if $u_k = 1$.

We recall Proposition 2.6. In particular, if a word starts with two consecutive n-words, then it must occur at a position that is a multiple of 2^n. Henceforth, the above proposition will be referred to as the **recognizability property**.

Recall also that the system (X_u, S) is uniquely ergodic by Theorem 2.12, and the shift S is bijective on a subset of X_u of measure 1.

Note that a cylinder in X_u, denoted by $[W]$, denotes all sequences in X_u that start with the word W. Given a cylinder $[e]$ where e is 0 or 1, $\sigma^n[e]$ represents sequences of the form $\sigma^n(e)\sigma^n(w')$ where ew' is some sequence in X_u. On the other hand, $[\sigma^n(e)\sigma^n(f)]$ where $e, f \in \{0, 1\}$, denotes words in X_u that start with $\sigma^n(e)\sigma^n(f)$.

We first give a very simple lemma concerning $[\sigma^n(e)\sigma^n(e')]$ and $\sigma^{n+1}[e]$.

Lemma 2.27. *The cylinder $\sigma^{n+1}[e]$ is strictly contained in $[\sigma^n(e)\sigma^n(e')]$. The elements in $[\sigma^n(e)\sigma^n(e')]$ that are not in $\sigma^{n+1}[e]$ are exactly the closure of $\bigcup_{k\in\mathbb{N}} S^{(2k-1)2^n}(u)$.*

Proof. It is obvious that $\sigma^{n+1}[e]$ is in $[\sigma^n(e)\sigma^n(e')]$. However, *a priori*, the latter could begin with the form $\sigma^n(e)\sigma^n(e')\sigma^n(e)\sigma^n(e)\dots$ whereas the former, being σ^{n+1} of a sequence, cannot.

The recognizability property tells us that if a sequence v in X_u starts with two n-words then it must occur at a position $i = j2^n$ for some j, or that v is the limit $\lim_{m\to\infty} S^{j_m 2^n}u$ for some sequence j_m that goes to infinity. In other words, v is in the closure of the space $\bigcup_{k\in\mathbb{N}} S^{k2^n}(u)$.

We write $v_m = S^{j_m 2^n}u = \sigma^n(e)\sigma^n(e')ab\dots.$ then a occurs in u at a even position, $a \neq b$, $u_{j_m} = e$, and $v_m = \sigma^n(S^{j_m}u) \in \sigma^n[e]$. Since $v = \lim_{m\to\infty} v_m$, v will be in $\sigma^{n+1}[e]$ if we can find a subsequence of even numbers j_{m_p} that diverges, since then $v_{m_p} \in \sigma^{n+1}[e]$, and $\lim_{p\to\infty} v_{m_p} = v \in \sigma^{n+1}[e]$.

However, if all of j_m are odd, then $S^{j_m}u$ cannot be equal to $\sigma(ew)$ for any w in X_u, so $v_m \notin \sigma^{n+1}[e]$, and so $v \notin \sigma^{n+1}[e]$. In conclusion, the elements in $[\sigma^n(e)\sigma^n(e')]$ that are not in $\sigma^{n+1}[e]$ are exactly those that start at position $j2^n$ in u for odd j, or the limit of such sequences, that is, the closure of $\bigcup_{k\in\mathbb{N}} S^{(2k-1)2^n}(u)$. □

We now give a lemma that concerns the partition of X_u into disjoint cylinders.

Lemma 2.28. *We have*

$$X_u = \cup_{e=0,1} \cup_{k=0}^{2^n-1} S^k \sigma^n[e], \tag{59}$$

and this union is disjoint.

Proof. First, note that the set on the right is closed, being a finite union of shifted cylinders. Therefore, to show the inclusion, we need to show that $S^m u \in \cup_{e=0,1} \cup_{k=0}^{2^n-1} S^k \sigma^n[e]$ for all $m \in \mathbb{N}$. We can write $m = 2^n a + b$ where $0 \leq b < 2^n$. So $S^m u = S^b S^{a2^n} u$. Moreover, $S^{a2^n} u = u_{a2^n} u_{a2^n+1} \cdots = \sigma^n(u_a u_{a+1} \ldots) \in \sigma^n([u_a])$, so $S^m u \in S^b \sigma^n([u_a])$.

We now prove that $S^p \sigma^n[a] \cap S^q \sigma^n[b] = \emptyset$ for any letters a, b, if $p \neq q$. Without loss of generality, we may assume $p > q$. Suppose that on the contrary, there is some element in the intersection. This element can be written in two forms: $S^p \sigma^n(w)$ for some $w \in [a]$, and $S^q \sigma^n(w')$ for some $w' \in [b]$. As a consequence of bijectivity, $\sigma^n(w) = S^{q-p} \sigma^n(w')$. By definition, $w' = \lim S^{m_i} u$ and $w = \lim S^{p_i} u$. For i large enough, $S^{q-p}(\sigma^n(u_{m_i} \ldots u_{m_i+i}))$ will be a prefix of $\sigma^n(u_{p_i} \ldots u_{p_i+i})$. Hence $\sigma^n(u_{m_i} \ldots u_{m_i+i})$ will occur at position $j2^n + q - p$ for some j. By the recognizability property, $q - p$ must be a multiple of 2^n, and since $p, q < 2^n$, $p = q$, a contradiction. □

The next lemma further decomposes each disjoint $S^k \sigma^n[e]$ into smaller sets.

Lemma 2.29. *For every $n \geq 0$, we have*

$$\sigma^n[e] = [\sigma^n(e)\sigma^n(0)] \cup [\sigma^n(e)\sigma^n(1)] \tag{60}$$

for $e = 0$ or 1. If $e + e' = 1$, then

$$\sigma^n[e] = \sigma^{n+1}[e] \cup S^{2^n} \sigma^{n+1}[e'] \tag{61}$$

Moreover, the measure of the cylinder $[0]$ is $\mu([0]) = 1/2$, and $\mu(\sigma^n[e']) = \mu(\sigma^n[e])$.

Proof. If $w \in \sigma^n[e]$, then $w = \sigma^n(w')$, with $w'_0 = e$. So $w \in [\sigma^n(e)\sigma^n(w'_1)]$.

Conversely, if $w \in [\sigma^n(e)\sigma^n(f)]$, and $w = \lim_{p \to +\infty} S^{k_p} u$, then for large enough k_p, the first p letters of w will cover the word $\sigma^n(e)\sigma^n(f)$, i.e.

$$w_0 \ldots w_{2^{n+1}-1} = \sigma^n(e)\sigma^n(f) = u_{k_p} \ldots u_{k_p+2^{n+1}-1},$$

and, because of the recognizability property, k_p is a multiple of 2^n. So $k_p = l_p 2^n$. Define $w' = \lim S^{l_p}(u)$. Note that the limit exists, because $\lim_{p \to +\infty} S^{k_p} u$ exists. Moreover, $\sigma^n(w') = \sigma^n(\lim S^{l_p}(u)) = \lim(\sigma^n(S^{l_p}(u)) = \lim S^{k_p}(u) = w$. Since $w' \in [e]$, the converse is proved.

We will first show the second decomposition is included in the first one. That $\sigma^{n+1}[e] \subset [\sigma^n(e)\sigma^n(e')]$ was established in our simple lemma at the beginning. It is also obvious that $S^{2^n} \sigma^{n+1}[e']$, whose elements start with $\sigma^n(e)$, is contained in $[\sigma^n(e)\sigma^n(e')] \cup [\sigma^n(e)\sigma^n(e')]$.

Now, we show the first decomposition is included in the second. It is obvious that $[\sigma^n(e)\sigma^n(e)]$ is included in $S^{2^n}\sigma^{n+1}[e']$, since the latter consists of sequences that begin with $\sigma^n(e)$.

From Lemma 2.27, we know that the elements in $[\sigma^n(e)\sigma^n(e')]$ that are not in $\sigma^{n+1}[e]$ is the closure of $\bigcup_{k\in\mathbb{N}} S^{(2k-1)2^n}(u)$. The key of this proof is the realization that, such an element w can be written as limits of sequences of the form $\sigma^n(e \mid e'e\dots) = \sigma^n(v_j)$ where the bar is to indicate that the letter preceding it is at an odd position, and the letter after it is at an even position. Because v_js start at an odd position, we know that the sequences $e'v_j$ occur at an even position p_j in u. So the sequences s_j starting from positions $p_j/2$ satisfy $S^{2^n}\sigma^{n+1}(s_j) = S^{2^n}\sigma^n(e'v_j) = \sigma^n(v_j)$ which converges to w, so that $w \in S^{2^n}\sigma^{n+1}[e']$, hence $[\sigma^n(e)\sigma^n(e')]$ is included in the second decomposition.

We now proceed to show the claims about the measure. Since the Morse sequence is defined up to a change of 0 and 1, the symmetry forces $\mu([1]) = \mu([0])$ and so $\mu([0]) = 1/2$. Similarly, if we exchange the role of e' and e, we get $\mu(\sigma^n[e']) = \mu(\sigma^n[e])$. The proof is complete. $\square$

The lemmas introduced here will be useful in our definition of rank 1 systems and Rokhlin stacks.

2.4 AY: Spectral Theory

Lecture date: 2014/03/24.

2.4.1 AY: Hilbert Spaces and $\mathcal{L}^2(X, \mu)$

Definition 2.30. A **Hilbert Space** is a set $\mathcal{H}$ satisfying the following:
(i) $\mathcal{H}$ is a vector space over $\mathbb{C}$ or $\mathbb{R}$
(ii) $\mathcal{H}$ is equipped with inner product $(\cdot, \cdot)$ such that
• the map $f \mapsto (f, g)$ is linear on $\mathcal{H}$ for fixed $g \in \mathcal{H}$
•$(f, g) = \overline{(g, f)}$
•$(f, f) \geq 0$ for all $f \in \mathcal{H}$
and define $\|f\| = (f, f)^{1/2}$
(iii) $\|f\| = 0$ if and only if $f = 0$
(iv) Cauchy-Schwarz inequality and triangle inequality hold for all $f, g \in \mathcal{H}$:

$$|(f, g)| \leq \|f\|\|g\| \text{ and } \|f + g\| \leq \|f\| + \|g\|$$

(v) $\mathcal{H}$ is complete in the metric $d(f, g) = \|f - g\|$

Note that any separable Hilbert space has an **orthonormal basis**, a subset $\{e_i\}_{i=1}^{\infty}$ of $\mathcal{H}$ such that every two elements in the set are orthogonal and finite linear combinations of elements in the set are dense in $\mathcal{H}$. This follows from the separability assumption and using induction with the Gram-Schmidt algorithm. If the basis is finite, then the Hilbert space is said to be **finite-dimensional** and otherwise it is said to be **infinite-dimensional**

Definition 2.31. Given a measure space $(X, \mathcal{M}, \mu)$, we can define a space $\mathcal{L}^2(X, \mu)$ as the set of all equivalence classes of measureable functions (where $f \sim g$ if and only if $f = g$ a.e.) for which

$$\int_X |f(x)|^2 d\mu(x) < \infty$$

with norm given by

$$\|f\|_{\mathcal{L}^2(X,\mu)} = \left(\int_X |f(x)|^2 d\mu(x) \right)^{1/2}$$

and inner product given by

$$(f, g) = \int_X f(x)\overline{g(x)} d\mu(x)$$

and it can be shown that $\mathcal{L}^2(X, \mu)$ is in fact a Hilbert Space.

2.4.2 AY: Unitary Operators and Spectrum

Definition 2.32. A bounded linear operator $U : H \to H$ on Hilbert Space H is called a **Unitary Operator** if U is surjective and $\|Uf\| = \|f\|$ for all $f \in H$

Definition 2.33. For measure-theoretic dynamical system (X, T, μ) where both T and T^{-1} are measureable and measure-preserving, we can define a special unitary operator U_T on the Hilbert space $\mathcal{L}^2(X, \mu)$ by $f \mapsto f \circ T$. To show that U_T is indeed a unitary operator, note that U_T is surjective because for any $f \in \mathcal{L}^2(X, \mu)$, we have $U_T(f \circ T^{-1}) = f$. Moreover, $\|U_T f\| = \|f\|$ for all $f \in \mathcal{L}^2(X, \mu)$ because T is measure-preserving. To be more explicit, recall that the integral $\int_X f d\mu(x)$ is defined by a limit of integrals of simple functions of the form $\Sigma_{i=1}^k a_k \chi_{E_i}$ for which

$$\int_X (\Sigma_{i=1}^k a_k \chi_{E_i} d\mu(x)) = \Sigma_{i=1}^k a_k \mu(E_i)$$

Notice that $U_T(\chi_{E_i}) = \chi_{T^{-1}E_i}$ so that

$$\int_X (\Sigma_{i=1}^k a_k U_T(\chi_{E_i}) d\mu(x)) = \Sigma_{i=1}^k a_k \mu(T^{-1}E_i) = \Sigma_{i=1}^k a_k \mu(E_i)$$

We see that $\|U_T f\| = \|f\|$ for simple functions $f \in \mathcal{L}^2(X, \mu)$. For a general $f \in \mathcal{L}^2(X, \mu)$, we may assume that f is real valued and non-negative (by considering the positive and negative part of f separately). Then there exists a sequence $\{f_n\}$ of simple functions increasing to f so that

$$\int_X U_T f d\mu(x) = \lim_{n\to\infty} \int_X U_T f_n d\mu(x) = \lim_{n\to\infty} \int_X f_n d\mu(x) = \int_X f d\mu(x)$$

Definition 2.34. The **eigenvalues** of (X, T, μ) are defined to be the eigenvalues of U_T and similarly the **eigenfunctions** of (X, T, μ) are defined to be the eigenvectors of U_T. The **spectrum** of (X, T, μ) is the set of eigenvalues of U_T. An eigenvalue is said to be a **simple eigenvalue** if the dimension of its eigenspace is 1.

Proposition 2.35. *Given $S = (X, T, \mu)$, the following hold:*
(i) If $U_T f = \lambda f, f \in \mathcal{L}^2(X, \mu), f \neq 0$ then $|\lambda| = 1$
(ii) S is ergodic if and only if the eigenvalue 1 is a simple eigenvalue for U_T
(iii) If S is ergodic, then $|f|$ is constant a.e. for any eigenfunction f. If f, g are eigenfunctions corresponding to eigenvalue λ then $f = cg$ a.e. for some constant c.
(iv) Eigenfunctions corresponding to different eigenvalues are orthogonal

Proof. (i) We have $\|f\| = \|U_T f\| = |\lambda|\|f\|$ and since $f \neq 0$ then $|\lambda| = 1$.

(ii) Suppose S is not ergodic. Then there exists an invariant set E such that $\mu(E) \neq 0$ and $\mu(E) \neq 1$. We show that χ_E is a eigenfunction with eigenvalue 1. Indeed, $U_T \chi_E = \chi_E T = \chi_{T^{-1}E} = \chi_E$. Moreover, χ_E is not constant a.e. so the eigenvalue 1 is not simple. Suppose the eigenvalue 1 is not simple. Then there exists and eigenfunction f with eigenvalue 1 such that f is not constant a.e. We show that f is T-invariant. Indeed, $f \circ T = U_T f = f$ so by Remark 1.30 S is not ergodic.

(iii) If λ is the eigenvalue of f then $|\lambda| = 1$ by (i). We have $|U_T f| = |\lambda||f| = |f|$ so $|f|$ is T-invariant and by Remark 1.30 $|f|$ is constant a.e. If g is another eigenfunction with eigenvalue λ then $|g|$ is constant a.e. and since $g \neq 0$ then f/g is T-invariant and hence constant a.e. by Remark 1.30.

(iv) Take distinct eigenvalues λ, μ with corresponding eigenfunctions f, g such that $U_T f = \lambda f, U_T g = \mu g$. Then

$$(f, g) = (U_T f, U_T g) = (\lambda f, \mu g) = \lambda\overline{\mu}(f, g)$$

and since $\lambda\overline{\mu} \neq 1$ then $(f, g) = 0$ □

Definition 2.36. The spectrum of (X, T, μ) is said to be **discrete** if $\mathcal{L}^2(X, \mu)$ admits a Hilbert basis of eigenfunctions. Thus, if $\mathcal{L}^2(X, \mu)$ is separable, then there are at most a countable number of eigenvalues.

An important contribution of the spectral theory is that it provides useful tools for distinguishing between isomorphism classes of topological dynamical systems or measure-theoretic dynamical systems.

Definition 2.37. A property of a topological (measure-theoretic) dynamical system is said to be an **invariant** if both (X_1, T_1) and (X_2, T_2) have the property if and only if the two topological dynamic systems are topologically conjugate (measure-theoretically isomorphic).

Before continuing the discussion of invariants, we describe the analogue of spectrum for topological dynamical systems.

Definition 2.38. Let (X, T) be a topological dynamical system, where T is a homeomorphism. Let $\mathcal{C}(X)$ be the set of non-zero complex-valued continuous functions on X. Then $f \in \mathcal{C}(X)$ is said to be an **eigenfunction** of T if there exists $\lambda \in \mathbb{C}$ called the eigenvalue of f such that $\forall x \in X, f(Tx) = \lambda f(x)$. The set of all eigenvalues corresponding to these eigenfunctions is called the **topological spectrum** of T. If the eigenfunctions of T span $\mathcal{C}(X)$ then we say that T has **topological discrete spectrum**.

Theorem 2.39. *(i) Any invertible and minimal topological dynamical system with topological discrete spectrum is topologically conjugate to a minimal rotation on a compact abelian metric group.*

(ii) Two minimal topological dynamical systems (X_1, T_1) and (X_2, T_2) with topological discrete spectrum, where T_1, T_2 are homeomorphisms, are topologically conjugate if and only if they have the same eigenvalues.

To prove Theorem 2.39(i), we first need to define the notion of a transitive topological dynamical system, which is similar to but weaker than minimality.

Definition 2.40. A topological dynamical system (X, T) is **topologically transitive** if there exists $x \in X$ such that $O_T(x) = \{T^n(x) | n \in \mathbb{Z}\}$.

Proof of Theorem 2.39(i) (Theorem 5.18 in [4]). Suppose (X, T) is a minimal topological dynamical system with discrete spectrum. Clearly, (X, T) is transitive. Let d be the metric on X. We show that T is an isometry for other metric ρ on X. Let K denote the unit circle in $\mathbb{C}$. Since (X, T) has discrete spectrum, we may choose a sequence of functions $\{f_n\}$ where each $f_n : X \to K$ has $f_n(T) = \lambda_n f_n$, and f_n is a basis for $\mathcal{C}(X)$. By Stone-Weierstrass, f_n separates points of X. Thus, we may define the metric ρ by:

$$\rho(x, y) = \Sigma_{n=1}^{\infty} \frac{|f_n(x) - f_n(y)|}{2^n}$$

Moreover,

$$\rho(Tx, Ty) = \Sigma_{n=1}^{\infty} \frac{|\lambda_n f_n(x) - \lambda_n f_n(y)|}{2^n} = \rho(x, y)$$

since $|\lambda_n| = 1$ as shown before. Thus, T is an isometry.

We show that ρ generates the same topology d on X. It suffices to show that the identity map $(X, d) \to (X, \rho)$ is continuous, since a bijective continuous map from a compact space onto a Hausdorff space is a homeomorphism. Fix $\epsilon > 0$ and take N large enough so that $\Sigma_{n=N+1}^{\infty} \frac{2}{2^n} < \frac{\epsilon}{2}$. Since the functions $f_n, 1 \leq n \leq N$ are continuous there exists $\delta > 0$ such that $d(x, y) < \delta$ implies $|f_n(x) - f_n(y)| < \frac{\epsilon}{2}$, for all $1 \leq n \leq N$. Then $d(x, y) < \delta$ implies $\rho(x, y) < \Sigma_{n=1}^{N}(\frac{\epsilon}{2^{n+1}}) + \frac{\epsilon}{2} \leq \epsilon$.

Next, we show that T is topologically conjugate to a minimal rotation on a compact abelian metric group. Since (X, T) is transitive, there exists some $x_0 \in X$ such that $\overline{O_T(x_0)} = X$. Now, we define a group structure by defining multiplication $*$ on $O_T(x_0)$ by $T^n(x_0) * T^m(x_0) = T^{n+m}(x_0)$. Then

$$\begin{aligned}
&\rho(T^n(x_0) * T^m(x_0), T^p(x_0) * T^q(x_0)) \\
&= \rho(T^{n+m}(x_0), T^{p+q}(x_0)) \\
&\leq \rho(T^{n+m}(x_0), T^{p+m}(x_0)) + \rho(T^{p+m}(x_0), T^{p+q}(x_0)) \\
&= \rho(T^n(x_0), T^p(x_0)) + (T^m(x_0), T^q(x_0))
\end{aligned}$$

which implies that the map $* : X_T(x_0) \times X_T(x_0) \to X_T(x_0)$ is uniformly continuous, and since $\overline{O_T(x_0)} = X$ then $*$ may be extended uniquely to a continuous map $* : X \times X \to X$. Moreover, the inverting map $T^n(x_0) \mapsto T^{-n}(x_0)$ has

$$\begin{aligned}
&\rho(T^{-n}(x_0), T^{-m}(x_0)) \\
&= \rho(T^{-n+m+n}(x_0), T^{-m+m+n}(x_0)) \\
&= \rho(T^m(x_0), T^n(x_0))
\end{aligned}$$

and hence the inverting map is uniformly continuous and can be extended to a continuous map on X. Thus, X is a topological group and is abelian since the subgroup $\{T^n(x_0) | n \in \mathbb{Z}\}$ is abelian and dense in X. Finally, since $T(T^n(x_0)) = T^{n+1}(x_0) = T(x_0) * T^n(x_0)$ then $T(x) = T(x_0) * x$ so we see that T is the rotation by $T(x_0)$. Then if we define $S : X \to X$ by $S(x) = T * x$ then it is easy to see that $\phi : X \to X$, the identity map $(X, d) \to (X, \rho)$, has $\phi \circ T = S \circ \phi$. $\square$

Before proving part (ii), we need a lemma first which is proved in proposition 2.54 in Kurka:

Lemma 2.41. *If (X, R_a) is a minimal rotation on a compact group, where $R_a(x) = a * x$ then λ is an eigenvalue of R_a if and only if λ is the eigenvalue corresponding to an eigenfunction which is a multiple of a character f of X, in which case $\lambda = f(a)$.*

Proof of Theorem 2.39(ii) (Theorem 2.56 [3]). Suppose (X_1, T_1) and (X_2, T_2) have the same spectrum. By part (i), we may assume that the two systems are minimal rotations on compact Abelian groups $(X, R_a), (Y, R_b)$. By the lemma, we have $\{f(a) | f \in \hat{X}\} = \{g(b) | g \in \hat{Y}\}$. Define the map $\psi : \hat{Y} \to \hat{X}$ by $(\psi(g))(a) = g(b)$. It is easy to see that ψ is a bijective group homomorphism. By Pontrjagin duality theorem, there exists a bijective gorup homomorphism $\varphi : X \to Y$ such that $\hat{\varphi} = \psi$, or $\psi(g) = g \circ \varphi$. Now, for any $g \in \hat{Y}$, we have $g(\varphi(a)) = (\psi(g))(a) = g(b)$. Since $\hat{Y}$ separates points, and g is arbitrary, then we must have $\varphi(a) = b$. Now, it is easy to see that $\varphi : (X, R_a) \to (Y, R_b)$ is a conjugacy. For any $x \in X$,

$$\varphi R_a(x) = \varphi(a * x) = \varphi(a) * \varphi(x) = b * \varphi(x) = R_b \varphi(x)$$

Suppose (X_1, T_1) and (X_2, T_2) are topologically conjugate. Then there exists $\varphi : X \to Y$ such that $\varphi \circ T_1 = T_2 \circ \varphi$. Suppose we have an eigenfunction $f \in \mathcal{C}(Y)$ such that $f \circ T_2 = \lambda f$. Then $f \circ \varphi \circ T_1 = \lambda f \circ \varphi$ so that $f \circ \varphi$ is an eigenfunction in $\mathcal{C}(X)$ with eigenvalue λ. Thus, (X_1, T_1) and (X_2, T_2) have the same spectrum. ☐

It turns out that this theorem has an analogue for measure-theoretic dynamical systems. Before we can describe this analogue, we must define a new notion of isomorphism.

Definition 2.42. Measure-theoretic dynamical systems (X, T, μ) & (Z, R, ρ) are **measure-theoretically isomorphic** if there exists $X_1 \subseteq X, Z_1 \subseteq Z$ and a bimeasureable bijection $\phi : X_1 \to Z_1$ such that $\mu(X_1) = \rho(Z_1) = 1, \phi\mu = \rho$ and $R\phi = \phi T$

The proof of the following analogue theorem appears in [2]:

Theorem 2.43. *(Von Neumann)*

(i) Any invertible and ergodic system with discrete spectrum is measure-theoretically isomorphic to a rotation on a compact abelian group, equipped with Haar measure.

(ii) Two invertible and ergodic transformations with identical discrete spectrum are measure-theoretically isomorphic.

Definition 2.44. Given a unitary operator U on Hilbert space H, the **cyclic space generated by** $f \in H$ is the closure of the subspace generated by the set $\{U^n(f) | n \in \mathbb{Z}\}$. Let $H(f)$ denote the cyclic space generated by f.

Theorem 2.45. *If L is a separable Hilbert space approximated by an increasing sequence of cyclic spaces, then L itself is a cyclic space.*

Proof. By assumption, $L = \overline{\cup(H(f_n))}$. We show the existence of $g \in L$ such that $L = H(g)$. Since L is separable, there exists a dense sequence $\{g_n\}$ in L. To show $L = H(g)$, it suffices to show that $\cap_{n,p}\{g|d(g_n, H(g) < 1/p\}$ is non-empty. By Baire's theorem, we can show this by showing that given fixed $f \in L$ and $\epsilon > 0$ that $\{g|d(f, H(g)) < \epsilon\}$ is dense in L. Take any $h \in L$, and without loss of generality we may assume that $\|h\| = 1$. Since $f, h \in L$ then there exists an integer m such that $d(f, H(f_m)) < \epsilon$ and $d(h, H(f_m)) < \epsilon$. The second inequality implies that there exists a polynomial P such that $d(h, P(U)f_m) < \epsilon$. We show that there exists a polynomial Q such that $d(h, Q(U)f_m) < 2\epsilon$. Write

$$P(U)f_m = \Pi_{i=1}^{k}(U - a_i)(f_m)$$

Consider ϵ_0 such that $0 < \epsilon_0 < 1$ and define

$$Q(U)f_m = \Pi_{i=1}^{k}(U - a_i + \epsilon_0)(f_m) = P(U)f_m + \Sigma_{i=0}^{k}(\epsilon_0)^i(p_i(U))(f_m)$$

where the $p_i(U)$ are polynomials in U. Notice that

$$\begin{aligned} |Q(U)f_m| &\leq |P(U)f_m| + \Sigma_{i=0}^{k}|(\epsilon_0)^i(p_i(U)(f_m))| \\ &\leq |P(U)f_m| + (\epsilon_0)\Sigma_{i=0}^{k}|(p_i(U)(f_m))| \end{aligned}$$

On the unit circle K,

$$|Q(U)f_m - P(U)f_m| \leq |Q(U)f_m| - |P(U)f_m| \leq \epsilon_0\Sigma_{i=0}^{k}|(p_i(U)(f_m)| \leq \epsilon_0 M$$

where we have used the fact that $|(p_i(U)(f_m)|$ is bounded by some constant M, being a real function on a compact set. Thus, there exists Q such that $d(P(U)f_m, Q(U)f_m) < \epsilon$. By triangle inequality, we have $d(h, Q(U)f_m) < d(h, P(U)f_m) + d(P(U)f_m, Q(U)f_m) = 2\epsilon$. Since $Q(U)$ is non-zero on K, then $Q(U)$ is invertible on K, and we may approximate this inverse by polynomials in U. Let $\{S_i\}$ be a sequence that approximates $Q^{-1}(U)$. Let $g = Q(U)f_m$. By the above, $d(h, g) < 2\epsilon$. We show that $d(f, H(g)) < 2\epsilon$. To do so, we first show $d(H(f_m), H(g)) < \epsilon$. We have

$$\lim_{n\to\infty}(S_j(U)g) = f_m$$

For some j, $d(H(f_m), H(S_j(U)g)) < \epsilon$ and $d(H(S_j(U)g), H(Q^{-1}(U)g)) < \epsilon$, and clearly $d(H(Q^{-1}(U)g), H(g)) = 0$ so that

$$\begin{aligned} d(H(f_m), H(g)) &< d(H(f_m), H(S_j(U)g)) + d(H(S_j(U)g), H(Q^{-1}(U)g)) \\ &\quad + d(H(Q^{-1}(U)g), H(g)) \\ &< 2\epsilon \end{aligned}$$

□

2.5 AZ: The Morse Sequence and Rokhlin Stacks

Lecture date: 2014/03/31.

Motivation: To provide a geometric interpretation of the Morse dynamical system (X_u, S). The interpretation covered today will break the Morse system into a series of stacking "towers." It turns out that, at each step of the series, the Morse system comes out to be two disjoint towers that are twice the height of each of the towers in the previous step.

Definition 2.46. A sequence of partitions $P^n = \{P_1^n, ..., P_{k_n}^n\}$ of X **generates** a measure-theoretic dynamical system (X, T, μ) if there exists a set E with $\mu(E) = 0$ such that, for every pair $(x, x') \in (X \setminus E)^2$, if x and x' are in the same set of the partition P^n for every $n \geq 0$, then $x = x'$.

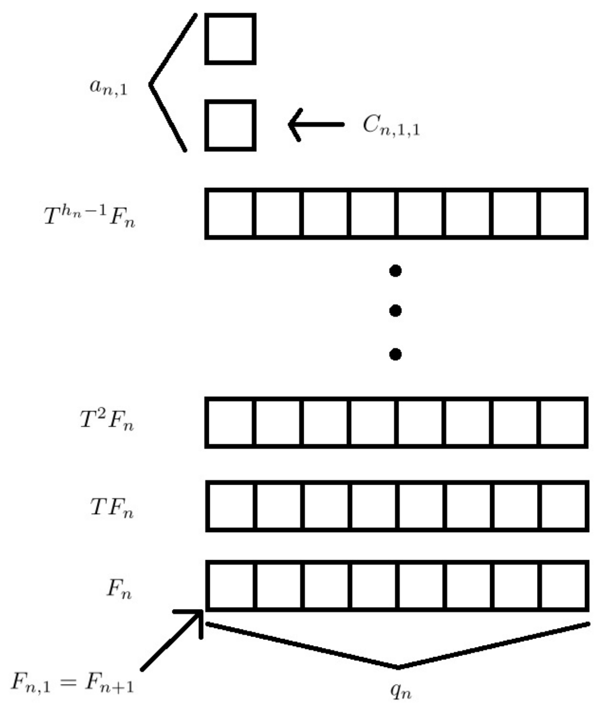

Figure 1: A diagram of a rank one transformation.

Next, we define what it means for a system to be of rank one. Refer to Figure 1.

Definition 2.47. We say that a system (X, T, μ) is of **rank one** if there exist sequences of positive integers $(q_n)_{n\in\mathbb{N}}$, $(a_{n,i})_{n\in\mathbb{N},1\leq i<q_n}$, and $(h_n)_{n\in\mathbb{N}}$ and subsets of X, denoted by $(F_n)_{n\in\mathbb{N}}$, $(F_{n,i})_{n\in\mathbb{N},1\leq i\leq q_n}$ and $(C_{n,i,j})_{n\in\mathbb{N},1\leq i<q_n,1\leq j\leq a_{n,i}}$ such that, for every fixed n,

(a) $(F_{n,i})_{1\leq i\leq q_n}$ is a partition of F_n,

(b) the sets $(T^k F_n)_{1\leq k\leq h_n-1}$ are disjoint

(c) $T^{h_n}F_{n,i} = C_{n,i,1}$ if $a_{n,i}\neq 0$ and $i<q_n$,

(d) $T^{h_n}F_{n,i} = F_{n,i+1}$ if $a_{n,i}=0$ and $i<q_n$,

(e) $TC_{n,i,j} = C_{n,i,j+1}$ if $j<a_{n,i}$,

(f) $TC_{n,i,a_{n,i}} = F_{n,i+1}$ if $i<q_n$,

(g) $F_{n+1} = F_{n,1}$.

Additionally, h_n must satisfy both the recurrence relation $h_0 = 1$ and $h_{n+1} = q_n h_n + \sum_{j=1}^{q_n-1} a_{n,i}$ and the inequality

$$\sum_{n=0}^{+\infty} \frac{h_{n+1}-q_n h_n}{h_{n+1}} < +\infty, \tag{62}$$

and the sequence of partitions $\{F_n, TF_n, ..., T^{h_n-1}F_n, X\setminus \cup_{k=0}^{h_n-1} T^k F_n\}$ generates the system (X,T,μ).

Definition 2.48. The union of the disjoint $(T^k F_n)_{1\leq k\leq h_n-1}$ from the last definition is called a **Rokhlin stack** of base F_n. We say also that the system is **generated** by sequence of Rokhlin stacks with bases F_n.

The following lemma is given without proof.

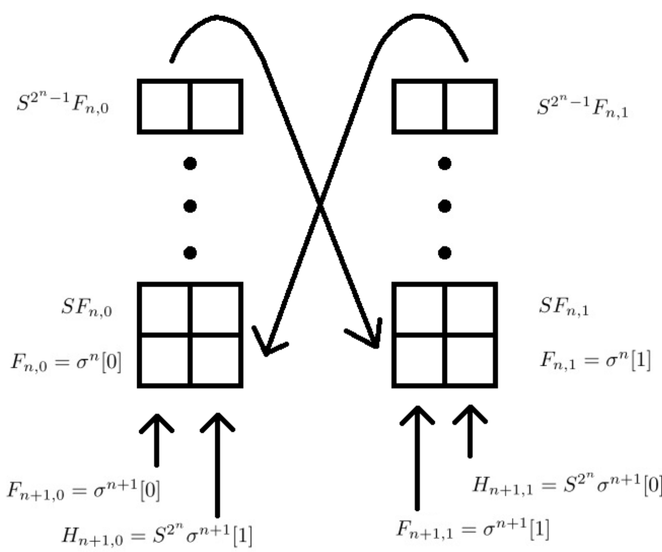

Figure 2: n^{th} stage

Lemma 2.49. *5.2.4 in the text Every substitution σ defines a continuous map from X_u to X_u.*

We now proceed to the main result. We may build the sets $F_{n,e} = \sigma^n[e]$ and their images $(S^k F_{n,e})_{0\leq k\leq 2^n-1}$ in the following way:

(1) Conventionally, for fixed e and n, we denote the set $S^k F_{n,e}$ by dots drawn one above the other as k increases.

(2) At stage n, we have two Rokhlin stacks (each n-stack being $\cup_{k=0}^{2^n-1} S^k F_{n,e}$, with $e = 0$ or 1) with bases $F_{n,0}$ and $F_{n,1}$, with height 2^n, whose levels $S^k F_{n,e}$ are disjoint sets of measure 2^{-n-1}.

(3) The shift map S sends each level of each stack, except the top ones, onto the level immediately above; S is not explicit on the top levels.

(4) In the beginning, $F_{0,0}$ and $F_{0,1}$ are two disjoint sets of measure $1/2$.

(5) At stage n, we cut $F_{n,e}$ into two subsets of equal measure $F_{n+1,e}$ and $H_{n+1,e}$. The shift map S becomes explicit on part of the levels where it was not yet so, as it sends $S^{2^n-1}F_{n+1,0}$ onto $H_{n+1,1}$ and $S^{2^n-1}F_{n+1,1}$ onto $H_{n+1,0}$. This defines the $(n+1)$-stacks, which will have height 2^{n+1}.

The process is illustrated in Figure 2. In the diagram, $F_{n,0}$ denotes $\sigma^n[0]$, and $F_{n,1}$ denotes $\sigma^n[1]$. From Lemma 2.29, we have $\sigma^n[e] = \sigma^{n+1}[e] \cup S^{2^n}\sigma^{n+1}[e']$, where $e + e' = 1$. Note that the top stack associated with $F_{n,0}$ is thus

$$S^{2^n-1}\sigma^n[0] = S^{2^n-1}\sigma^{n+1}[0] \cup S^{2^{n+1}-1}\sigma^{n+1}[1]. \tag{63}$$

Applying S to both sides gives us

$$S^{2^n}\sigma^n[0] = S^{2^n}\sigma^{n+1}[0] \cup S^{2^{n+1}}\sigma^{n+1}[1]. \tag{64}$$

Likewise, we have

$$S^{2^n}\sigma^n[1] = S^{2^n}\sigma^{n+1}[1] \cup S^{2^{n+1}}\sigma^{n+1}[0], \tag{65}$$

and combining these two statements shows how the stacks are rearranged from the nth step to the $(n+1)$th step, see Figure 3.

We must prove a proposition to conclude that the stacks indeed generate (X_u, T, μ).

Proposition 2.50. *There exists a countable set E such that for every pair $(w, w') \in (X_u \setminus E)^2$, if w and w' are in the same stack and the same level $S^k F_{n,e}$, for all $n \geq 0$, $0 \leq k \leq 2^n - 1$ and $e = 0, 1$, then $w = w'$.*

Proof. If w and w' are always in the same level of the same stack, they are in the same $S^{k_n}\sigma^n[e_n]$ for all n; and hence the sequences w and w' coincide between the indices 0 and $w^n - 1 - k_n$; this implies $w = w'$ if $2^n - 1 - k_n \to +\infty$.

The proof of the case where $2^n - 1 - k_n$ remains bounded is omitted for now because the solution provided in the book is unclear. ∎

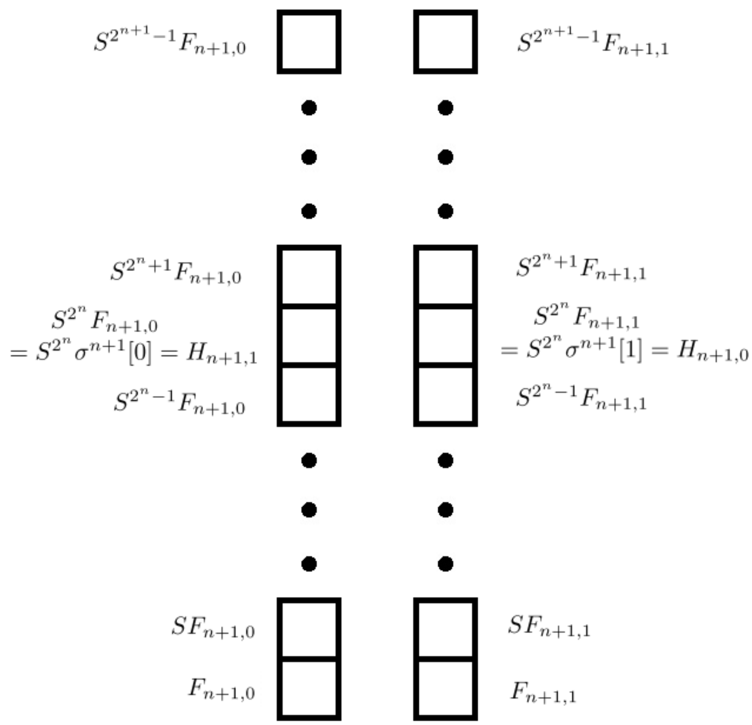

Figure 3: $(n+1)^{th}$ stage

2.6 JC: The dyadic rotation

Lecture date: 2014/03/31

The goal of this lecture is to compare the Morse system to geometric systems. We will focus on the most elementary case of this, systems on intervals. Kubrat mentioned topological conjugacy in his lecture as well as some interesting properties of it. In particular, topological conjugacy preserves unique ergodicity. We will define new notions of isomorphism, since topological conjugacy turns out to be too strong.

We recall from Definition 2.42 that a **measure theoretic isomorphism** ϕ between systems (X, T, μ) and (Z, R, ρ) is a bimeasureable bijection from $X_1 \subseteq X$ to $Z_1 \subseteq Z$ such that $\mu(X_1) = \rho(Z_1) = 1$, $\phi\mu = \rho$ and $R\phi = \phi T$.

We note that topological conjugacy implies that two systems are measure-theoretically isomorphic. This notion of measure-theoretically isomorphic is weaker than topological conjugacy since X_1 and Z_1 are more general. For these isomorphisms, we remove sets of measure zero and then conjugate the spaces through an invertible measureable transformation. We now define a more moderate notion of isomorphism:

Definition 2.51. Two systems (X, T) and (Z, R) are **semi-topologically conjugate** if there exist $X_1 \subset X$, $Z_1 \subset Z$, and ϕ abicontinuous bijection from X_1 to Z_1 such that $X \setminus X_1$ and $Z \setminus Z_1$ are countable, and $R \circ \phi = \phi \circ T$. In that case, if (X, T) is a symbolic system, we say that it is a **coding** of (Z, R).

Proposition 2.52. *Semi-topological conjugacy preserves unique ergodicity and, for uniquely ergodic systems, the semi-topological conjugacy of* (X, T) *and* (Z, R) *implies the measure-theoretic isomorphism of* (X, T, μ) *and* (Z, R, ρ).

Proof. The proof that semi-topological conjugacy preserves unique ergodicity is similar to the proof of Proposition 2.23 and is omitted here. ☐

We now introduce rotations.

Definition 2.53. A **rotation** is the dynamical system (G, R, λ) where G is a compact group, R is a translation of G, and λ is the Haar measure on G.

Example 2.54. Let G be the torus $\mathbb{T}^1$, and $Rx = x + \alpha$ modulo 1 for irrational α. Here the point x is getting translated by α, taking the decimal part if necessary. A common result is to show that this is ergodic.

Now we will introduce the **Van der Corput map**. It is a map from $Y_0 = [0, 1]$ to itself defined by

$$R_0(1 - \frac{1}{2^n} + x) = \frac{1}{2^{n+1}} + x, \text{ for } 0 \leq x < \frac{1}{2^{n+1}}, n \in \mathbb{N} \text{ and } R_0(1) = 0. \quad (66)$$

The Van der Corput map preserves the Lebesgue measure λ_0.

Every number in $[0, 1]$ has a binary decimal representation $0.\omega_0\omega_1....$ We note that the sequence after the decimal is an element of $Y = \{0, 1\}^{\mathbb{N}}$. More formally, we define the mapping

$$\chi(x) = (\omega_0\omega_1...) \text{ whenever } x = \sum \omega_j 2^{-j-1} \quad (67)$$

Definition 2.55. The **dyadic rotation** is a mapping R from $Y = \{0, 1\}^{\mathbb{N}}$ to $Y = \{0, 1\}^{\mathbb{N}}$ defined as

$$R = \chi R_0 \chi^{-1} \quad (68)$$

Then we see that R takes a binary sequence, sends it to the decimal fraction, Van der Corput maps it, then sends the result back to a binary representation. If we apply R to a point of the form $0abc$, it gets sent to a point of the form $1abc$.

The set Y with addition in base two is a compact group. It contains a copy of $\mathbb{N}$ by $\phi(1) = 10...0...$, $\phi(2) = \phi(1) + \phi(1)$. It also contains a copy of $\mathbb{Z}$ since we see that $10...0... + 11111... = 00000...$ so we can define $\phi(-1) = 111....$ We call $(Y, +)$ the group of **2-adic integers**. Then R is the rotation $\omega \mapsto \omega + 1$. Using the terms we defined before, we see that (Y_0, R_0) and (Y, R) are semi-topologically conjugate.

Definition 2.56. A **character** on a group G is a homomorphism $\gamma : G \to \mathbb{T}^1$. If G is a topological group, we also require that γ is continuous.

Proposition 2.57. *The system (Y, R) is uniquely ergodic, and hence the system (Y, R, λ) is ergodic.*

Proof. For each measure ν on Y, we can define a Fourier transform $\hat{\nu}(\gamma) = \int_{g\in Y} \gamma(g)d\nu$. If μ is invariant under R, we have $\hat{\mu(\gamma)} = \int_{g\in Y} \gamma(Rg)d\mu = \gamma(1)\hat{\mu}(\gamma)$. Thus, if $\hat{\mu}(\gamma) \neq 0$,$\gamma(1) = 1$. Then we have that $\gamma(n) = 1$ for all integers since γ is a homomorphism, so γ must be identically 1. Then we have that $\hat{\mu}$ is zero on characters which are not the identity and one on the character 1. Thus, μ is the Haar measure λ. Then by Proposition 5.1.22 in the book, the system (Y, R, λ) is ergodic.

Next, we consider the action of the operator $U : \mathcal{L}^2(Y, \lambda) \to \mathcal{L}^2(Y, \lambda)$ defined by $Uf(w) = f(Rw)$.

Proposition 2.58. *The system (Y, R) has a discrete spectrum. Its eigenvalues are $e^{2i\pi\alpha}$ for all the dyadic rationals $\alpha = p2^{-k}$, for $p, l \in \mathbb{Z}$.*

Proof. It is a classical result that for a compact group, the characters generate a dense subspace of $\mathcal{L}^2(Y, \lambda)$. Since γ is a character, $\gamma(Rg) = \gamma(1)\gamma(g)$. Thus all the characters are eigenfunctions and all the eigenvalues are the $\gamma(1)$. Then it is sufficient to find all the characters of Y. The integers are dense in Y, so a character of Y must also be a character of $\mathbb{Z}$, hence $\gamma(n) = e^{2i\pi n\alpha}$ for an $\alpha \in [0, 1]$. Then if we can extend this γ continuously to Y, it will remain what we want. Next, we have that $\omega = \omega_0\omega_1... = \lim_{k\to+\infty} \omega_0...\omega_k 00...0... = \lim_{k\to+\infty} \sum_{j=0}^{k} \omega_j 2^j$. Then we need to find all the α such that for all ω, $e^{2\pi i\alpha\beta_k}$ converges as $k \to +\infty$ where $\beta_k = \sum_{j=0}^{k} \omega_j 2^j$. We then write $\alpha = \sum_{k=0}^{+\infty} \alpha_k 2^{-k-1}$. The α_i are ultimately equal to 0 or for infinitely many k, $\alpha_k = 1$, $\alpha_{k+1} = 0$. The sequence cannot converge if it is the latter, so α must have a dyadic expansion that terminates. This means it is a dyadic rational number. Each α yields an eigenvalue $e^{2\pi i\alpha}$. □

2.7 JF: Morse Shift as a Two-Point Extension

We recall the uniquely ergodic rotation (Y, R, λ) on the 2-adic integers from Definition 2.55. If (Y_0, R_0) is the Van der Corput map, we recall that (Y_0, R_0, λ_0) (where λ_0 is Lebesgue measure) is semi-topologically conjugate to (Y, R, λ), and therefore these maps are also measure theoretically isomorphic.

Proposition 2.59. *The system (Y, R) is rank one.*

Proof. Let $G_0 = Y$ and for $n \geq 1$, let

$$G_n = [\underbrace{0 \dots 0}_{n}].$$

We see that for $0 \leq k < 2^n$ the images $R^k G_n$ are disjoint and

$$\{G_n, RG_n, \dots, R^{2^n-1}G_n\}$$

partition Y. We also have that $G_n = G_{n+1} + R^{2n}G_{n+1}$, which shows us how to achieve the stack for $n+1$ from the stack for n. To confirm then that this system is rank one by these stacks, we need that these partitions generate the system.

Suppose for each $n \in \mathbb{N}$,

$$w, w' \in R^{k_n}G_n \text{ for some } 0 \leq k_n < 2^n.$$

Then as $R^{k_n}G_n$ is a cylinder on a word of length n,

$$w_0 w_1 \dots w_{n-1} = w'_0 w'_1 \dots w'_{n-1}$$

for all n. Therefore $w = w'$. □

We will now show a relationship between the uniquely ergodic Morse Shift (X_u, S, μ) and the system (Y, R, λ). On X_u, define the map $\tau : X_u \to X_u$ by $\tau(w) = w'$, where $w_n + w'_n = 1$ for all $n \in \mathbb{N}$. τ is a homeomorphism on X_u and $s \circ \tau = \tau \circ S$. We define an equivalence on X_u by

$$w \sim \tilde{w} \iff w = \tilde{w} \text{ or } w = \tau(\tilde{w})$$

and let $\overline{X}_u$ be quotient of X_u under this equivalence. This covering $\pi : X_u \to \overline{X}_u$ is two-fold, and we call the induced system $(\overline{X}_u, \overline{S}, \overline{\mu})$ a **factor with fiber two of** (X_u, S, μ).

Furthermore, we may also define a map $\phi : \overline{X}_u \times \mathbb{Z}_2 \to X_u$, by

$$\phi(\overline{w}, e) = w$$

where $\pi(w) = \overline{w}$ and $w_0 = e$. ϕ defines a topological conjugacy, and the measure on $\overline{X}_u \times \mathbb{Z}_2$ is $\overline{\mu}$ times uniform measure m on $\mathbb{Z}_2$. The action of S in X_u induces an action on $\overline{X}_u \times \mathbb{Z}_2$ by

$$S(\overline{w}, e) = (\overline{S}\overline{w}, z(e, \overline{w}))$$

where $z(e, \overline{w}) = (\overline{w}, e)_1$. This map is measurable and continuous. We say that (X_u, S) is a **skew product** of $(\overline{X}_u, \overline{S})$ by $\mathbb{Z}_2$.

Consider the sequence of two stacks that generate the system (X_u, S) from Lecture 2.5. These were of the form $S^k F_{e,n}$, where $0 \leq k < 2^n$ and $F_{e,n} = \sigma^n[e]$. Because $\pi(F_{0,n}) = \pi(F_{1,n})$, call this mutual image $\overline{F}_n$ with stack $\overline{S}^k \overline{F}_n$ for $0 \leq k < 2^n$. Under these stacks, we see that $(\overline{X}_u, \overline{S}, \overline{\mu})$ is rank one.

We now come to the relationship between (X_u, S) with (Y, R). We call (X_u, S) a **two-point extension** of $(\overline{X}_u, \overline{S})$, as it is homeomorphic to the second system times a two point space. If we show that $(\overline{X_u}, \overline{S})$ is topologically semi-conjugate to (Y, R), then we have also that (X_u, S) is a coding of a two-point extension of (Y, R) as follows. Recall from Definition 2.51 that a topological semi-conjuacy is called a coding if one system is a shift.

Proposition 2.60. *$(\overline{X}_u, \overline{S})$ is topologically semi-conjugate to (Y, R)*

Proof. Given $\overline{w} \in \overline{X}_u$, let $(k_n)_{n \in \mathbb{N}}$ be given by $\overline{w} \in \overline{S}^k \overline{F}_n$. Note that either

$$k_{n+1} = k_n \text{ or } k_{n+1} = k_n + 2^n \tag{69}$$

For each n, let $U_{k,n} = R^k[\underbrace{0 \ldots 0}_{n}] \subseteq Y$ for $0 \leq k < 2^n$. By (69),

$$U_{k_{n+1},n+1} = [z_1 z_2 \ldots z_n e] \text{ where } U_{k_n,n} = [z_1 z_2 \ldots z_n] \text{ and } e \in \{0,1\}. \tag{70}$$

or $U_{k_{n+1},n+1} \subseteq U_{k_n,n}$. Therefore, $\phi : \overline{X}_u \to Y$ given by

$$\psi(w) = \bigcap_{n=1}^{\infty} U_{k_n,n}$$

is well-defined and onto. The function ψ is continuous as

$$\psi^{-1}([m]) = \overline{S}^m \overline{F}_n$$

where $2^{n-1} \leq m < 2^n$ and m is the representation of integer m in Y. Also, $\psi(\overline{S}^k \overline{F}_n) = U_{k,n}$, and so the image of an open set is open as well.

Let $X_1 \subseteq \overline{X}_u$ be the image under π of the set of w in X_u such that the position of w in each n-stack pair uniquely determines w and $Y_1 = \psi(X_1)$. Then $\overline{X}_u \setminus X_1$ and $Y \setminus Y_1$ are countable and ψ defines a topological semi-conjugacy. $\square$

Because ψ is onto, we may define ψ^{-1} by making a choice for each $y \in Y$ such that $\psi^{-1}(y)$. This will not be an issue, as we will only use the effective composition $\psi \circ \psi^{-1}$, which will be the identity. As we see in Figure 4, we may define a skew-product $Y \times \mathbb{Z}_2$ over (Y, R, λ) by

$$(y, e) \mapsto (R(y), \tilde{z}(e, y)) \text{ for } \tilde{z} = z \circ \tilde{\psi}^{-1}. \tag{71}$$

The system (X_u, S, μ) is topologically semi-conjugate to $(Y \times \mathbb{Z}_2, R \rtimes \tilde{z}, \lambda \times m)$, a two point extension of (Y, R, λ), by $\tilde{\psi} \circ \phi$.

$$\begin{tikzcd}
(\overline{X}_u \times \mathbb{Z}_2, \overline{S} \rtimes z, \overline{\mu} \times m) \arrow[d, "\tilde{\psi}=\psi \times id_2"'] \arrow[rr, bend left, "id_1"] & (X_u, S, \mu) \arrow[l, "\phi"'] \arrow[r, "\pi"] & (\overline{X}_u, \overline{S}, \overline{\mu}) \arrow[d, "\psi"] \\
(Y \times \mathbb{Z}_2, R \rtimes \tilde{z}, \lambda \times m) \arrow[rr, "id_1"'] & & (Y, R, \lambda)
\end{tikzcd}$$

Figure 4: The diagram of maps. Here ψ is an onto topological semi-conjugacy, ϕ is a topological conjugacy, π is a factor map of fiber two, and id_j represents projection onto coordinate j.

We end this section with a discussion about the spectrum of the Morse Shift. To do this, we use the association between the factor $(\overline{X}_u, \overline{S})$ and the rotation (Y, R). Recall from Proposition 2.58 that the spectrum of (Y, R) has discrete spectrum with the set of eigenvalues to be all values of the form $e^{2\pi\imath\alpha}$ for dyadic rationals α.

Proposition 2.61. *The Morse Shift has non-discrete simple spectrum, and the eigenvalues are the values $e^{2\pi\imath\alpha}$ for each dyadic rational α.*

Proof. Consider vector space $L = \mathcal{L}^2(X_u, \mu)$ with unitary operator $U : L \to L$ given by $Uf = f \circ S$. Consider another linear operator $V : L \to L$ by $Vf = f \circ \tau$.

Let $H_1 = \{f \in L : Vf = f\}$ and $H_{-1} = \{f \in L : Vf = -f\}$. It follows that

$$H = H_1 + H_2 \text{ and } H_1 \perp H_{-1}. \tag{72}$$

Furthermore, $H_1 \simeq \mathcal{L}^2(\overline{X}_u, \overline{\mu})$. Because S and τ commute, $H_{\pm 1}$ is U invariant.

Because $(\overline{X}_u, \overline{S}, \overline{\mu})$ and $(Y, R.\lambda)$ are measure theoretically isomorphic, the eigenvalues in the spectrum H with eigenfunctions in H_1 are all values $e^{2\pi\imath\alpha}$ where α is a dyadic rational. Because (X_u, S, μ) is ergodic, all these eigenvalues are simple.

Suppose by contradiction that there are eigenfunctions $f \in H \setminus H_1$ of U. Because such a function may be expressed as $f = f_1 + f_{-1}$ where $f_{\pm 1} \in H_{\pm 1}$ and $f_{-1} \neq 0$, then $Uf = e^{2\pi\imath\beta} f$ would imply that $Uf_{\pm 1} = e^{2\pi\imath\beta} f_{\pm 1}$, as $H_{\pm 1}$ are U invariant. Therefore, our contradictory assumption may without loss of generality consider an eigenfunction $f \in H_{-1}$. But then $g = f \cdot f \in H_1$, and so $Ug = e^{2\pi\imath 2\beta} g$ implies that 2β is a dyadic rational. And so β is a dyadic rational. Because there exists a $e^{2\pi\imath\beta}$-eigenfunction $g' \in H_1$, the dimension of the $e^{2\pi\imath\beta}$-eigenspace must be at least two, as $g' \perp g$. This however is a contradiction to ergodicity.

Therefore, the eigenvalues are all simple and of the desired form, but the span of the eigenfunctions are not dense in H (just H_1). $\square$

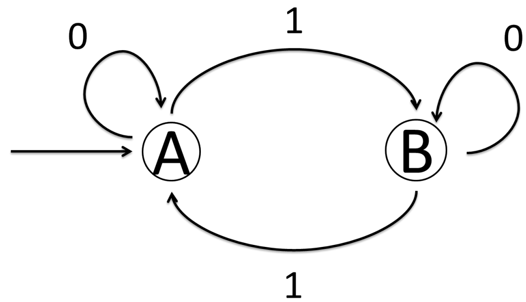

Figure 5: The automaton that generates the Morse sequence.

2.8 JC: Automata and the Rudin-Shapiro sequence

Lecture date: 2014-04-07.

The goal of this lecture is to define what automata and automatic sequences are and to introduce some properties of the Rudin-Shapiro sequence.

First we will define what a k-automaton is.

Definition 2.62. A k**-automaton** is represented by a directed graph (digraph) with a finite set $\mathcal{S}$ of vertices called **states**, k oriented edges labeled from 0 to $k-1$ called **transition maps** from $\mathcal{S}$ to itself, and a set Y and a map ϕ from $\mathcal{S} \to Y$ called the **output function** or **exit map**. One of the states in $\mathcal{S}$ is called the **initial state**.

Definition 2.63. We call a sequence (u_n) k**-automatic** if it is generated by a k-automaton such that when taking the base k expansion of the integer $n = \sum_{i=0}^{j} r_i k^i$, we can follow the digits $r_0, r_1 ... r_j$ along the oriented edges starting from the initial state to get to the state $a(n)$. Then the term u_n of the sequence is $u(n) := \phi(a(n))$.

We say that the automaton generates this k-automatic sequence in **reverse reading** because the digits are read in increasing order of powers. We can give an equivalent definition of k-automaticity by reading the digits in a decreasing order of powers. This is proved in Proposition 1.3.4 in the textbook.

Example 2.64. Here we show how a 2-automaton generates the Morse sequence in direct reading. The state A exits as 0 and the state B exits as 1. See Figure 5

Now we will introduce the Rudin-Shapiro sequence.

Definition 2.65. The **Rudin-Shapiro sequence** (ϵ_n) over alphabet $\{-1, +1\}$ is defined by $\epsilon_0 = 1$ and for $n \in \mathbb{N}$, $\epsilon_{2n} = \epsilon_n$ and $\epsilon_{2n+1} = (-1)^n \epsilon_n$.

We discussed how the Morse sequence gives the parity of the sum of digits of n, or the parity of the number of occurences of 1 in the dyadic development of n.

Similarly, the Rudin-Shapiro sequence gives the parity of the number of words 11 in the dyadic development of n. We shall see this in the next proposition.

Proposition 2.66. *For integer r with dyadic development $r = \sum_{i=0}^{k} r_i 2^i$ with $r_i \in \{0, 1\}$, we have that*

$$\epsilon_r = (-1)^{\sum_{i\geq 0} r_i r_{i+1}} . \tag{73}$$

Proof. We will proceed by induction on the number of digits. It is clear when $k = 0$. Then assuming the proposition is true for k, we will prove it is true for $k+1$. We define $e(r)$ to be the sum $\sum_{i\geq 0} r_i r_{i+1}$ where $r = \sum_{i=0}^{k} r_i 2^i$. First, we note that $(-1)^{e(2r)} = \epsilon_{2n} = \epsilon_n = (-1)^{e(r)}$ since we are just shifting the digits. We note that this is not in mod 2. Similarly, $e(2r + 1) = r + e(r)$ simply by the properties of exponents.

Then we know that the formula to be proven works for $r = \sum_{i=1}^{k} r_i 2^{i-1}$ because r has k digits by our inductive hypothesis. Then we will show it works for a number $2r + r_0$ where $r_0 = 0$ or 1.

We get the equation

$$e(2r + r_0) = r_0 r + e(r) = r_0 r_1 + e(r) = \sum_{i=0}^{k-1} r_i r_{i+1}. \tag{74}$$

□

Then we can say that in the Rudin-Shapiro sequence counts the number of times the word $'11'$ appears in the binary expansion.

Example 2.67. The first few terms of the Rudin-Shapiro sequence are

$$+1, +1, +1, -1, +1, +1, -1, +1... \tag{75}$$

Example 2.68. This is the Rudin-Shapiro sequence automaton. The states A and B map to $+1$ and the states C and D map to -1. See Figure 6.

Then again thinking of integers as words over the alphabet $\{0, 1\}$, this 2-automaton in reverse or direct reading with exit map $\phi(a) = \phi(b) = 1$ and $\phi(c) = \phi(d) = -1$ generates the Rudin-Shapiro sequence.

Remark 2.69. Over the four-letter alphabet $\{a, b, c, d\}$, the Rudin-Shapiro sequence is the fixed point u (starting with the letter a of the substition

$$a \mapsto ab \; b \mapsto ac \; c \mapsto db \; d \mapsto dc \tag{76}$$

If we let v be the sequence obtained by applying ϕ to u, we can say that v is the limit of the words $\phi(\sigma^n(a))$ We see quickly that $\phi(\sigma^n(c)) = -\phi(\sigma^n(b))$ where the $-$ sign flips the bits. Then we can say that v is given by the limit of the recursion formula $A_{n+1} = A_n B_n$ and $B_{n+1} = A_n(-B_n)$.

Remark 2.70. We can say that v is a 2-automatic sequence, but it is not the fixed point of a substitution.

Using the notion of primitivity from Jack's first lecture, we can quickly verify that for any two letters l, m there is some k such that l occurs in $\sigma^k(m)$. Then since u is defined as the fixed point of σ and σ is primitive, by Prop. 1.43, u is minimal.

Remark 2.71. The complexity of u is $8n - 8$.

We can also calculate its complexity in a similar way as the Morse sequence. We will not do the calculation during the lecture. To connect to notions that Kubrat mentioned, we have the following remark.

Remark 2.72. X_u and X_v are topologically conjugate.

To connect the Rudin-Shapiro sequence to Jay's first lecture, we have the following remark.

Remark 2.73. The dynamical system generated by the Rudin-Shapiro sequence (u or v) is uniquely ergodic. We can prove this following the way this was proved for the Morse sequence last week.

Now we will study some spectral properties of the system (X_u, S, μ) generated by the Rudin-Shapiro sequence.

We want to answer the question "Could a function with a spectral type equivalent to the Lebesgue measure and with simple spectrum exist?"

We answer this question positively with the next proposition.

Definition 2.74. The **spectral type** of f or of the generated cyclic space is the finite positive measure ρ on $\mathbb{T}^1$ defined by $\hat{\rho}(n) = (U^n f, f)$. Its total mass is $\|f\|_H^2$

Proposition 2.75. *Let f be a function defined over X_v such that $f(w) = w_0$ and let $f_N = \sum_{n=0}^{N-1} z_n e^{2\pi i n x}$ where $z_n = 1$ when $v_n = 1$, 0 otherwise. Its spectral type is the weak-* limit of the measures with density (Radon-Nikodym derivative from lecture 3) $\frac{1}{N} |f_N|^2$ with respect to the Lebesgue measure λ.*

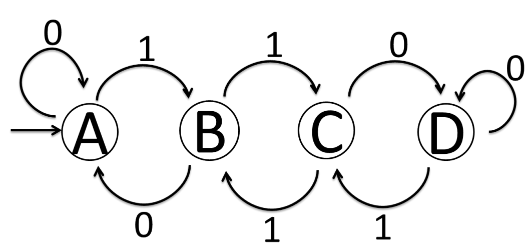

Figure 6: The automaton that generates the Rudin-Shapiro sequence

Proof. The spectral type ρ_f is defined by

$$\hat{\rho}_f(p) = (f, U^p f) = \mu(\{w; w_0 = 1, w_p = 1\}) \tag{77}$$

There is a large section in chapter 1 of the book which discusses "correlation" measures. Suffice it to say that the frequency of the words $a_0...a_p$ where $a_0 = a_p = 1$ is exactly

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{n=0}^{N-1} 1_{v_n=v_{n+p}=1} \tag{78}$$

We check that $\hat{\rho}_N(p) = \frac{1}{N}\sum_{n=0}^{N-p-1} z_n z_{n+p}$ where $z_n z_{n+p} = 1$ only if $v_n = v_{n+p} = 1$. Lastly, from a result from real analysis, the convergence of Fourier coefficients is equivalent to the weak-star topology convergence of measures. Then we can conclude that ρ_f is the weak-* topology limit of ρ_N. □

Proposition 2.76. *The associated operator U has a spectrum with multiplicity at most four; all the dyadic rationals are eigenvalues; there exist functions the spectral type of which is equivalent to the Lebesgue measure.*

Proof. (X, S) is a four-point extension of a system with a simple spectrum, so $\mathcal{L}^2(X,\mu)$ is a limit of $H_n^1 + H_n^2 + H_n^3 + H_n^4$ where H_n^i are cyclic space which can be empty and not orthogonal. One can generalize Theorem 2.45 from Alex's first lecture to show that $\mathcal{L}^2(X,\mu)$ is generated by at most four cyclic spaces. One of these contains eigenvectors of the dyadic rotation, so the dyadic rationals are the eigenvalues by Proposition 2.58 of Jay's first lecture.

If we normalize f from the previous proposition to be orthogonal to constants, then $\int f d\mu = 1/2$. Then letting $g = 2f - 1$, we use the previous proposition to get that ρ_g is the vague limit of $\rho_n' = 2^{-n}|g_n|^2\lambda$, where $g_n(x) = \sum_{j=0}^{2^n-1} v_k e^{2\pi ijx}$ (recall that v is the sequence deduced from the Rudin-Shapiro sequence by letter-to-letter projection). If $B_n = w_{n,0}...w_{n,2^n-1}$, then we let $h_n(x) = \sum_{j=0}^{2^n-1} w_{n,k} e^{2\pi ijx}$. Then based on the discussion before, we have that $g_{n+1}(x) = g_n(x) + e^{2\pi i 2^n x} h_n(x)$ and $h_{n+1}(x) = g_n(x) - e^{2\pi i 2^n x} h_n(x)$, so we can say $|g_{n+1}^2(x)| + |h_{n+1}^2(x)| = 2(|g_n^2(x)| + |h_n^2(x)|)$ and for $n = 0$, it is equal to 2^{n+1} for all x. Then we get the inequality $\frac{|g_n|^2}{2^{n+1}} \leq 1$ for all n and the densities of ρ_n' must converge in $\mathcal{L}^2(X,\mu)$. Lastly, the density of ρ_g must be strictly positive, since $\frac{|g_n|^2}{2^{n+1}} \to 0$ implies $\frac{|h_n|^2}{2^{n+1}} \to 1$, which contradicts the recursion formula. □

2.9 AY: The Unique Ergodicity of Primitive Substitutions

Lecture date: 2014-4-14.

Definition 2.77. The substitution σ over alphabet $\mathcal{A}$ is primitive is there exists k such that for each pair $a, b \in \mathcal{A}$ the letter a occurs in $\sigma^k(b)$

Our first result tells us that if we take repeatedly apply a primitive substitution to a fixed letter in $\mathcal{A}$, the frequencies of each letter in the resulting images $\mathcal{A}$ stabilize to a limit.

Lemma 2.78. *For each $a \in \mathcal{A}$ and $e \in \mathcal{A}$, if $N(e, \sigma^n(a)) = |\sigma^n(a)|_e$ denotes the number of occurrences of the letter e in the word $\sigma^n(a)$, then $\frac{N(e,\sigma^n(a))}{|\sigma^n(a)|}$ tends to a positive limit f_e independent of a in the limit as $n \to \infty$.*

Proof. For a given word V, let $\mathbf{1}(V)$ denote the vector $(|V|_e, e \in \mathcal{A})$. Let M_σ denote the incidence matrix of σ defined by $M_\sigma = (m_{i,j})_{i,j\in\mathcal{A}}$ where $m_{i,j} = N(i, \sigma(j))$. Recall that for any word V, we have $\mathbf{1}(\sigma(V)) = M_\sigma \mathbf{1}(V)$. Since σ is primitive, then there exists integer k such that M_σ^k has only positive integer entries. By Perron-Frobenius theorem, M_σ has a positive eigenvalue α corresponding to eigenvector u such that the following hold: $\alpha > 1$ since M_σ has all integer entries, α is simple, $\alpha > \lambda$ for any other eigenvalue λ By the Power-iteration algorithm, we know that $\frac{M_\sigma^n \mathbf{1}(a)}{\alpha^n}$ converges to a positive multiple u_a of u. To be explicit, we can write $\mathbf{1}(a) = c_1 v_1 + c_2 v_2 + ... + c_N v_N$ where $v_1 = (1, 0, ..., 0)^t$ is the vector u expressed under the Jordan basis, and $v_i, 2 \le i \le N$ are generalized eigenvectors. By Perron-Frobenius, c_1 is real and positive. Write $M_\sigma = VJV^{-1}$ where J is the Jordan form. Then

$$\frac{M_\sigma^n \mathbf{1}(a)}{\alpha^n} \tag{79}$$

$$= \frac{VJ^n}{\alpha^n}(c_1 v_1 + c_2 v_2 + ... + c_N v_N) \tag{80}$$

$$= c_1 u + \frac{VJ^n}{\alpha^n}(c_2 v_2 + ... + c_N v_N) \tag{81}$$

Now, $\frac{J^n}{\alpha^n} \to E_{1,1}$ where $E_{1,1}$ is the matrix with entry 1 at $(1, 1)$ and 0 everywhere else, since α dominates all other eigenvalues.

Then

$$c_1 u + \frac{VJ^n}{\alpha^n}(c_2 v_2 + ... + c_N v_N) \to c_1 u \text{ as } n \to \infty \tag{82}$$

Note that we can write $|\sigma^n(a)| = \mathbf{1}(\sigma^n(a)) \cdot e$ where $e = (1, ..., 1)$. Then we have the limit $\frac{N(\sigma^n(a))}{|\sigma^n(a)|} \to \frac{u_a}{\langle u^a, e\rangle}$ where $N(\sigma^n(a)) = (N(e, \sigma^n(a))_{e\in\mathcal{A}}$. Since this is a multiple of u whose coordinates sum to 1, then this limit is independent of a.

□

Our next result will extend the previous result to show that the frequencies of words occurring in the fixed point of the primitive substitution σ also stabilize in a similar fashion.

Lemma 2.79. *Let u be a fixed point of the primitive substitution σ. Let $\mathcal{A}_l$ be the alphabet whose letters are the words of length l that occur in u. For each letter $a \in \mathcal{A}$, and any word $W \in \mathcal{A}_l$, the $\frac{N(W,\sigma^n(a))}{|\sigma^n(a)|}$ tends to a positive limit f_W independent of a as $n \to \infty$.*

Proof. Define a new substitution ζ_l by the following: given $W \in \mathcal{A}_l$ where $W = w_1...w_l$, $\sigma(W) = w'_1...w'_m$, and $q = |\sigma(w_1)|$, then define

$$\zeta_l(W) = (w'_1...w'_l)(w'_2...w'_{l+1})...(w'_q...w'_{q+l-1}) \tag{83}$$

which is well-defined since $q + l - 1 \leq m$ since $|\sigma(e)| > 0, \forall e \in \mathcal{A}$ We show that $U_l = (u_0...u_{l-1})(u_1...u_l)...(u_n...u_{n+l-1})...$ is a fixed point for ζ_l. Let $q_0 = |\sigma(u_0)|$. We can write $\sigma(u_0) = u_0...u_{q_0-1}$ since u is a fixed point of σ. Let $f(i)$ denote the factor of u of length l beginning at index i. Now,

$$\begin{aligned} \zeta_l(u_0...u_{l-1}) &= (f(0))(f(1))...(f(q_0-1)) \\ &= (u_0...u_{l-1})(u_1...u_l)...(u_{q_0-1}...u_{q_0+l-2}) \end{aligned} \tag{84}$$

Continuing in this fashion, we can define $q_1 = |\sigma(u_1)|$ and repeat the argument above to obtain

$$\begin{aligned} \zeta_l(u_1...u_l) &= (f(q_0))(f(q_0+1))...(f(q_0+q_1-1)) \\ &= (u_{q_0}...u_{q_0+l-1})...(u_{q_0+q_1-1}...u_{q_0+q_1+l-2}) \end{aligned} \tag{85}$$

And repeating this argument indefinitely shows that U_l is fixed by ζ_l Next we show that ζ_l is primitive. Recall the assumption that all letters in $\mathcal{A}$ occur at least once in the fixed point u. Since u is a fixed point of σ, then for any $a \in \mathcal{A}$ and integer $p \in \mathbb{Z}$ the word $\sigma^p(a)$ is a factor of u. Since u is a fixed point of a primitive substitution, then u is minimal. Since $\mathcal{A}_l$ is finite ($|\mathcal{A}_l|$ is clearly bounded above by $|\mathcal{A}|^l$, and since $|\sigma^p(a)| \to \infty$ as $p \to \infty$ then we can pick p sufficiently large so that each $W \in \mathcal{A}_l$ is a factor of $\sigma^p(a)$. Since σ is primitive, there exists an integer $m > 0$ such that for any $b \in \mathcal{A}$ we have that a occurs in $\sigma^m(b)$. Take any $V, W \in \mathcal{A}_l$. By primitivity, we know a occurs as the first letter (in the alphabet $\mathcal{A}$) of one of the letters (in the alphabet $\mathcal{A}_l$) of $\zeta_l^m(V)$. By the choice of p, we know that $\zeta_l^{m+p}(V)$ contains the letter $W \in \mathcal{A}_l$. Write $V = v_1...v_l$. It is clear by the definition of ζ_l that $|\sigma^n(v_1)| = |\zeta_l^n(V)|$. As $n \to \infty$ we see that $N(W, \zeta_l^n(V))$ differs by at most l from $N(W, \sigma^n(v_1))$ because we can write $\sigma^n(v_1) = u_Q u_{Q+1}...u_{Q+R}$ as some factor of u, and $\zeta_l^n(V) = (u_Q...u_{Q+l-1})(u_{Q+1}...u_{Q+l})...(u_{Q+R}...u_{Q+R+l-1})$. Thus, by the previous lemma, there exists some positive limit f_W such that

$$\frac{N(W, \sigma^n(a))}{|\sigma^n(a)|} \to \frac{N(W, \zeta_l^n(V))}{|\zeta_l^n(V)|} \to f_W \tag{86}$$

□

Third, we prove a general result regarding symbolic dynamical systems induced by fixed points.

Proposition 2.80. *If v is a fixed point of a primitive substitution σ then the symbolic dynamical system (X_v, T) is uniquely ergodic*

Proof. Let u be the eigenvector associated to the dominating eigenvalue α of M_σ. Recall from the proof of Lemma 2.78 that for each $a \in \mathcal{A}$ there exists some positive constant c_a

$$\frac{M_\sigma^n \mathbf{1}(a)}{\alpha^n} \to u_a = c_a u \text{ as } n \to \infty \tag{87}$$

Thus, for p sufficiently large,

$$|\sigma^p(a)| = M_\sigma^p \mathbf{1}(a) \cdot e = \alpha^p c_a u \cdot e \tag{88}$$

Thus, there exists constants c, d such that

$$c\alpha^p < \ \inf_{a\in\mathcal{A}} |\sigma^p(a)| < \ \sup_{a\in\mathcal{A}} |\sigma^p(a)| < d\alpha^p \tag{89}$$

Recall that if s is the Morse sequence then for any factor W of s,

$$\frac{N(W, s_k ... s_{k+n})}{n+1} \to f_W, \text{uniformly in } k \tag{90}$$

We can prove analogously that for our fixed point v of σ,

$$\frac{N(W, v_k ... v_{k+n})}{n+1} \to f_W, \text{uniformly in } k \tag{91}$$

where we use the inequalities $c\alpha^p < \ \inf_{a\in\mathcal{A}} |\sigma^p(a)| < \ \sup_{a\in\mathcal{A}} |\sigma^p(a)| < d\alpha^p$ instead of $|\sigma^p(s_i)| = 2^p$. By Proposition 1.35, the symbolic dynamical system (X_v, T) is uniquely ergodic with measure μ defined by $\mu([W]) = f_W$ for all cylinders associated with each factor W of v. $\square$

Proposition 2.81. *If u is a non-periodic fixed point of a primitive substitution, then the unique invariant probability measure of the symbolic dynamical system (X_u, T) is non-atomic*

Proof. Suppose there exists $w \in X_u$ with $\mu(w) = \eta > 0$. Then for all cylinders W_n, $n \in \mathbb{Z}$, we have $\mu([W_n]) \geq \eta$. Write $W_n = w_0...w_n$. Since $\mu([W_n]) = f_{W_n}$ then for each n, there exist integers j_n, i_n sufficiently large such that $j_n > i_n$ and $j_n - i_n < \frac{1}{\eta}$ such that W_n occurs in u at indices j_n and i_n. If $n > \frac{2}{\eta} > j_n - i_n$ then $w_0...w_{\lfloor n/2 \rfloor} = w_{j_n - i_n}...w_{j_n - i_n + \lfloor n/2 \rfloor}$. Since $0 < j_n - i_n < \frac{1}{\eta}$ then $j_n - i_n$ takes on finitely many values and hence there exists an integer k such that $w_0...w_{\lfloor n/2 \rfloor} = w_k...w_{k+\lfloor n/2 \rfloor}$ for infinitely many n. Since this is true for arbitrarily large n, then actually $w_0...w_{\lfloor n/2 \rfloor} = w_k...w_{k+\lfloor n/2 \rfloor}$ holds for all n, which implies w is periodic. By Lemma 1.13, for any initial word $u_0...u_n$ in u, there exists N such that $u_0...u_n = w_N...w_{N+n}$. Thus, if w is periodic with period k then so is u, a contradiction since u is non-periodic.

□

Proposition 2.82. *If σ is primitive, then its fixed points have an at most linear complexity.*

Proof. Let u be a fixed point of σ. For all $a \in \mathcal{A}$, we know that $|\sigma^n(a)| \to \infty$ as $n \to \infty$, so we may find p such that $\inf_{a \in \mathcal{A}} |\sigma^{p-1}(a)| \leq n \leq \inf_{a \in \mathcal{A}} |\sigma^p(a)|$. These inequalities imply that any factor of length n in u appears in either in some $\sigma^p(a)$ or some $\sigma^p(ab)$. Given fixed $a, b \in \mathcal{A}$, there are at most $|\sigma^p(ab)|$ factors of length n in $\sigma^p(a)$ or $\sigma^p(ab)$, counting according to the position of the first letter of the factor. Clearly, there are at most $K = |\mathcal{A}|^2$ possible pairs $a, b \in \mathcal{A}$. Recall from Proposition 2.80 above, that we may choose p large enough so that

$$c\alpha^p < \inf_{a \in \mathcal{A}} |\sigma^p(a)| < \sup_{a \in \mathcal{A}} |\sigma^p(a)| < d\alpha^p \tag{92}$$

Then $|\sigma^p(ab)| < 2d\alpha^p$ so that for n sufficiently large,

$$p_u(n) \leq K|\sigma^p(ab)| < 2Kd\alpha^p \tag{93}$$

By Proposition 2.80 and the inequality we established earlier, we have the inequalities

$$c\alpha^{p-1} < \inf_{a \in \mathcal{A}} |\sigma^{p-1}(a)| < n \tag{94}$$

Thus, we have

$$p_u(n) \leq K|\sigma^p(ab)| < 2Kd\alpha^p < 2K\frac{d}{c}\alpha n \tag{95}$$

□

2.10 ZQ: Rauzy Induction and irrational rotation

Lecture date: 2014-04-14.

To put things in perspective, we have already seen that the Morse system is a coding of a two-point extension of the dyadic rotation. The two main results presented in this talk are that the Fibonacci system is a coding of a particular irrational rotation, and that the complexity function of the Fibonacci sequence is $n+1$, the smallest possible for a non-periodic sequence.

We begin with simple observations and definitions. Denote by $\mathbb{T}^1$ the unit circle centered around the origin in the complex plane. Given a real number $r \in [0,1)$, we can associate this number with the point $e^{2\pi i r}$ on the torus. With this association, and given an irrational number α, we define the **irrational rotation of** $\mathbb{T}^1$ **of angle** α by $\overline{R}(x) = xe^{2\pi i\alpha}$. Similarly, given the fundamental domain $[0,1)$, we define the rotation by $R(x) = x + \alpha$ if $x \in [0, 1-\alpha)$, and $R(x) = x+\alpha-1$ if $x \in [1-\alpha, 1)$. Note that $(\mathbb{T}^1, \overline{R})$ and $([0,1), R)$ are naturally semi-topologically conjugate, with the bicontinuous map ϕ from $[0,1)$ to $\mathbb{T}^1$ defined as $\phi(r) = e^{2\pi i r}$.

Definition 2.83. Let Q be a partition of X into two sets P_1 and P_2. For every point $w \in X$, its P-name is the sequence $P(w)$ such that $P(w)_n = i$ if $T^n w \in P_i$. $P(w)$ is also sometimes called the itinerary of w. We shall use these two interchangeably.

Definition 2.84. For a transformation T on X, a set $A \subset X$, and a point $x \in A$, we call first return time of x in A and denote by $n_A(x)$ the (possibly infinite) smallest integer $m > 0$ such that $T^m x \in A$. The induced map of T on A is the map $T^{n_A(x)}x$ defined on $A \cap \{x; n_A(x) < +\infty\}$.

The particular irrational rotation we are going to look at is $\alpha = \frac{1}{2}(\sqrt{5}-1)$. Let P_1 be the set $[0, 1-\alpha)$ and P_0 the set $[1-\alpha, 1)$. Let v be the P-name of the point α under R. We are interested in the trajectory of the rotation. Note that $\alpha \approx 0.62$, so that we have the ordering $0 < 1-\alpha < 0.5 < \alpha < 1$. Also, $\frac{1}{\alpha} = \frac{2}{\sqrt{5}-1} = \frac{2(\sqrt{5}+1)}{4} = \frac{\sqrt{5}+1}{2} = \alpha + 1$.

The following proposition gives a representation of the P-name as the image of a particular sequence under substitution and shift. It will be useful later when we show that v is in fact the Fibonacci sequence. Furthermore, the Rauzy induction used is interesting itself.

Proposition 2.85. *The sequence v is the image by $S\tau_0$ of the fixed point of the substitution τ, where S is the shift, $\tau_0(0) = 10$, $\tau_0(1) = 0$, $\tau(0) = 001$, $\tau(1) = 01$.*

Proof. Note that $\alpha = R(0)$, so that by definition, $v_n = 1$ if $R^{n+1}(0) \in [0, 1-\alpha)$ and $v_n = 0$ if $R^{n+1}(0) \in [1-\alpha, 1)$. So we can write $v = Sv''$ where $v''_n = 0$ if $R^n(0) \in [1-\alpha, 1)$, and $v''_n = 1$ if $R^n(0) \in [0, 1-\alpha)$.

Let I be the interval $[0, \alpha)$ and R' the induced map of R on I, $n(x)$ being the first return time of x in I.

We further partition $[0, \alpha)$ and look at the itinerary of 0 in that partition.

If $x \in [0, 1-\alpha)$, then $x+\alpha \in [\alpha, 1)$, which is the complement of I. $x+2\alpha \in [2\alpha-1, \alpha) \subset I$, so that $n(x) = 2$, and $R'(x) = x+2\alpha-1$. On the other hand, if $x \in [1-\alpha, \alpha)$, then $x+\alpha \in [0, 2\alpha-1)$, so that $n(x) = 1$, and $R'(x) = x+\alpha-1$.

Let v' be the sequence defined by $v'_n = 0$ if $R'^n(0) \in [0, 1-\alpha)$, $v'_n = 1$ if $R'^n(0) \in [1-\alpha, \alpha)$. The goal is to show that $v'' = \tau_0(v')$, and that $\tau(v') = v'$.

For the first part, it is useful to define the second return time of an x in I as $n_2(x) = n(x) + n(R^{n(x)}(x))$, and in general, the i^{th} return time as $n_i(x) = n_{i-1}(x) + n(R^{n_{i-1}}(x)(x))$. Please note that this return time is with respect to R, not R', as every application of R' to a point in I is guaranteed to result in a point in I.

Suppose $v'_i = 0$; then $R'^i(0)$ is in $[0, 1-\alpha)$; hence $R^{n_i(0)}(0)$, which is the same point, is in $[0, 1-\alpha)$, $R^{n_i(0)+1}(0)$ is in $[\alpha, 1) \subset [1-\alpha, 1)$, and $R^{n_i(0)+2}$ is in $[2\alpha-1, \alpha)$, which is again in $[0, \alpha)$, so it must be $R'^{i+1}(0)$, or in other words, $R^{n_{i+1}(0)}(0)$. Therefore, $n_{i+1}(0) - n_i(0) = 2$. By definition, since $R^{n_i(0)}(0) \in [0, 1-\alpha)$, $v''_{n_i(0)} = 1$, and since $R^{n_i(0)+1}(0) \in [1-\alpha, 1)$, $v''_{n_i(0)+1} = 0$.

Suppose $v'_i = 1$; then $R'^i(0)$ is in $[1-\alpha, \alpha)$, so $R^{n_i(0)}(0)$ is in $[1-\alpha, \alpha) \subset [1-\alpha, 1)$ and $R^{n_i(0)+1}(0)$ is in $[0, \alpha)$, so it must be $R^{n_{i+1}(0)}$. So we have $n_{i+1}(0) - n_i(0) = 1$ and $v''_{n_i(0)} = 0$.

The above analysis tells us that if $v'_i = 0$, then we need to apply R twice to get back to I, but we also know the itenerary of the point under these two R. If $v'_i = 1$, then we need to apply R once to get back to I. Therefore, $v'' = \tau_0(v')$ where $\tau_0(1) = 0$ and $\tau_0(0) = 10$. So $v = S\tau_0(v')$.

Now we need to show that v' is the fix point of τ. We do this by looking at R' in a different context.

We do not change v' if we make a homothety of the interval $[0, \alpha)$ of ratio $1/\alpha$. Remembering the action of R' on the interval $[0, \alpha)$, we know that R' becomes a rotation of $(2\alpha-1)/\alpha = 2 - 1/\alpha = 1-\alpha$ on the interval $[0, 1)$. And $v'_n = 1$ if $R'^n(0) \in [\alpha, 1)$, $v'_n = 0$ if $R'^n(0) \in [0, \alpha)$.

Let Q be the induced transformation of R' on $[0, \alpha)$, $m(x)$ being the first return time; by a similar argument, we have $Q(x) = x + 1 - \alpha$ and $m(x) = 1$ if $x \in [0, 2\alpha-1)$, and $Q(x) = x + 1 - 2\alpha$ and $m(x) = 2$ if $x \in [2\alpha-1, \alpha)$.

Let w be the sequence defined by $w_n = 0$ if $Q^n(0) \in [0, 2\alpha-1)$, $w_n = 1$ if $Q^n(0) \in [2\alpha-1, \alpha)$. By the same method as above, we can check that $v' = \tau_1(w)$, where $\tau_1(0) = 0$, $\tau_1(1) = 01$.

We use the same transformation on Q, multiplying by $\frac{1}{\alpha}$ again. Then Q becomes $(1-\alpha)/\alpha = \frac{1}{\alpha} - 1 = \alpha$, which is the rotation R itself, and that $w_n = 1$ if $Q^n(0) = R^n(0) \in [1-\alpha, 1)$, and $w_n = 0$ if $R^n(0) \in [0, 1-\alpha)$. Note that w is the "complement" of v'', meaning that $w_i + v_i = 1$ for all i.

Now we apply the Rauzy induction again to come back to R' and v', and see $w = \tau_2(v')$, where $\tau_2(0) = 01$, $\tau_2(1) = 1$.

Hence $v' = \tau_1\tau_2(v') = \tau(v')$, where $\tau(0) = 001$, $\tau(1) = 01$.

$\square$

2.11 ZQ: Geometric Representation of the Fibonacci Substitution

Lecture date: 2014-04-14.

In this section, σ is the Fibonacci substitution $0 \mapsto 01$ and $1 \mapsto 0$.

Its unique fixed point u is the Fibonacci sequence $0100101\ldots$ Note that the length of $\sigma^n(0)$ is the n^{th} Fibonacci nubmer f_n, given by the recursion formulas $f_0 = 1$, $f_1 = 2$, $f_{n+1} = f_n + f_{n-1}$. This elementary result follows directly from the fact that $\sigma^{n+1}(0) = \sigma^n(0)\sigma^{n-1}(0)$.

We give some remarks on the ergodic and spectral properties without proofs, since they have been presented in earlier talks.

Remark 2.86. The substitution σ is primitive, so by Proposition 5.4.4(presented by Alex) the system (X_u, S) is uniquely ergodic. The dominant eigenvalue of the incidence matrix is the golden ratio $\alpha_0 = \frac{1+\sqrt{5}}{2} = 1+\alpha$, which when we apply the Perron-Frobenius theorem tells us the limit of the frequencies of the letters 0 and 1 in u.

Let R and $\overline{R}$ be the irrational rotation of angle α defined respectively on the interval $[0, 1)$ and on the torus $\mathbb{T}^1$. The irrational rotation $\overline{R}$ is a translation on a compact group(the unit circle on the complex plane with the subspace topology); hence, by the same proof as in Proposition 5.2.17(presented by Jay), it is uniquely ergodic. The invariant measure, which is the pushforward measure of the Lebesgue measure on $[0, 1)$, gives a strictly positive measure to every open set. This, together with unique ergodicity, which tells us that every point is equidistributed with respect to the Lebesgue measure, imply that $\overline{R}$, as a dynamical system, is minimal, meaning that for every point $x \in \mathbb{T}^1$, the orbit of x is dense in $\mathbb{T}^1$.

It is easy to check that these properties of $\overline{R}$ are shared by R.

Now we give a result on the relation of the Fibonacci sequence and v, the P-name of α under the rotation R. We show that they are in fact the same sequence.

Proposition 2.87. *We have $u_n = 0$ whenever $R^n(\alpha) \in [1-\alpha, 1)$, $u_n = 1$ whenever $R^n(\alpha) \in [0, 1-\alpha)$. In other words, u is the P-name of α under R.*

Proof. We just have to identify the sequence u with the sequence v in the previous proposition. Recall that v' is the fix point of the substitution τ, where $\tau(0) = 001$ and $\tau(1) = 01$, so that v' starts with $00100101\ldots.$ We have $\tau_0\tau(0) = 10100$, $\tau_0\tau(1) = 100$.

We will prove by induction that $\tau_0\tau^n(0)$ is made with a 1 followed by $\sigma^{2n+1}(0)$ minus its last letter, and that $\tau_0\tau^n(1)$ is made with a 1 followed by $\sigma^{2n-1}(1)\sigma^{2n-1}(0)$ minus its last letter. For $n = 1$, $\sigma^3(0) = 01001$, and $\sigma(1)\sigma(0) = 001$, so that the base case is established. The induction step follows once we write u in its reverse decomposition of $u : \sigma^{2n+1}(0) = \sigma^{2n}(0)\sigma^{2n-1}(0) = \sigma^{2n-1}(0)(\sigma^{2n-2}(0)\sigma^{2n-1}(0))$. Hence v begins with $\sigma^{2n+1}(0)$ for all n, and so $v = u$.

□

Corollary 2.88. *The complexity function of the Fibonacci sequence is*

$$p_u(n) = n + 1 \tag{96}$$

for every n. We say that the Fibonacci sequence is a ***Sturmian sequence****; it has the lowest possible complexity for a nonperiodic sequence.*

Proof. Recall that $P_1 = [0, 1-\alpha)$ and $P_0 = [1-\alpha, 1)$, and we have shown that u, the fix point of the Fibonacci substitution, is the P-name of the point α under R. We first show that a word $w_0 \dots w_{n-1}$ occurs in u if and only if $\cap_{i=0}^{n-1} R^{-i} P_{w_i}$ is nonempty.

If a word $w_0 \dots w_{n-1}$ occurs in u, then by definition there exists m such that $R^m(\alpha) \in P_{w_0}$, $R^{m+1}(\alpha) \in P_{w_1}$, $\dots$, $R^{m+n-1}(\alpha) \in P_{w_{n-1}}$. Remembering that R is invertible, we have $\cap_{i=0}^{n-1} R^{-i} P_{w_i} \neq \emptyset$.

Conversely, suppose $\cap_{i=0}^{n-1} R^{-i} P_{w_i}$ is not empty, so that there is some $b \in \cap_{i=0}^{n-1} R^{-i} P_{w_i}$. Since α is irrational, for suitable sequence k, $R^k(\alpha)$ is as close to b as possible. Then $R^{k-i}(\alpha) \in R^{-i} P_{w_i}$, for all i, so that $u_{k+i} = w_i$, and $w_0 w_1 \dots w_{n-1}$ occurs in u.

The sets $\cap_{i=0}^{n-1} R^{-i} P_{w_i}$, when w ranges over $\mathcal{L}_n(u)$, are disjoint intervals, and they partition $[0, 1)$. The partition of the interval $[0, 1)$ by them is the partition of the interval by the points $R^{-i}(0)$, $1 \leq i \leq n$, or more explicitly, $-n\alpha$ mod 1. This can be seen by applying the map $-\alpha$ to the end points of the intervals $[0, 1-\alpha)$ and $[1-\alpha, 1)$. There are $n+1$ nonempty intervals, so there are $n+1$ words with length n in u. □

Proposition 2.89. *The system (X_u, S, μ) associated with the Fibonacci sequence is a coding of the rotation R on the interval, or $\overline{R}$ on the torus, preserving Lebesgue measure λ.*

Proof. To prove that a dynamical system is a coding of the other, we just need to find a homeomorphism which is defined outside a set that is at most countable, and commutes with the transformations on the two dynamical systems.

Let P be the partition $P_1 = [0, 1-\alpha)$, $P_0 = [1-\alpha, 1)$ and $P(x)$ be the P-name of x. We will prove that $P(x)$ is a homeomorphism between $[0, 1)$ and X_u.

First please recall the definition of the topology on X_u. It is the subspace topology of the product topology on $\{0, 1\}^{\mathbb{N}}$, or equivalently defined by the metric $d(u, v) = 2^{-m}$ where $m = \min\{|n| : u_n \neq v_n\}$. A basis for open sets in X_u are the cylinders in X_u. The topology on $[0, 1-\alpha)$ is just the usual topology given by the Euclidean metric.

We first show that $P(x)$ is continuous on the set $[0, 1[/D$ for a countable set D. We will use the fact that $P(\alpha) = u$ and $P(R^n(\alpha)) = S^n u$.

Suppose $R^{n_k}(\alpha) \to x$, using the metric definition, it is clear that

$$P(R^{n_k}(\alpha)) \to P(x). \tag{97}$$

The reason is that the linear function $x \to x + \alpha$ preserves the metric on $[0, 1)$, so that if the sequence can get as close to x, then applying the transformation

R repeatedly ensures that x and the sequence end up in the same interval P_i. There is a caveat: If x is in the orbit of α, then this continuity may fail, because P_0 and P_1 are semi-closed. For example, $1-\alpha$ is in the orbit of α, if we allow R^n for negative n. If we have a sequence that oscillates around $1-\alpha$ and converges to $1-\alpha$, then the limit of the P-names does not converge.

As a consequence, $P(x) = \lim P(R^{n_k}(\alpha)) = \lim S^{n_k}u$ is then a point of X_u. The set $\{n\alpha, n \in \mathbb{N}\}$ is dense in $[0,1)$, so we see that $P([0,1[/D) \subset X$, where D is the countable set $\{n\alpha, n \in \mathbb{Z}\}$ mod 1 in $[0,1)$. Note that this also shows $PR(x) = SP(x)$ for all $x \in [0,1)$, since $PR(x) = P(\lim R^{n_k+1}(\alpha)) = \lim P(R^{n_k+1}(\alpha)) = \lim S(R^{n_k}(\alpha)) = S\lim(R^{n_k}(\alpha)) = SP(x)$.

Next we show that P is invertible. Note that the observation from the previous corollary tells us that the interval $[0,1)$ is partitioned into disjoint intervals by the points $R^{-i}(0)$, $1 \leq i \leq n$, and each word of length n corresponds to one of these n intervals. If two points have the same P-name, then they must appear in the same interval in the partition for all n. The fact that α is irrational guarantees that the partitions will have decreasing lengths as n increases. As a result, two points in $[0,1)$ with the same P-name are identical. This guarantees that $P(x)$ is invertible on $P([0,1[/D)$.

Now we show P^{-1} is also continuous. This will actually be a result of the continuity of P. After deleting a countable number of points, which are the points in the orbit of u under the shift S, every point $v \in X_u$ can be written under the form $v = \lim S^{n_k}u$, for a diverging sequence n_k. By the sequential compactness of real numbers, the bounded sequence $R^{n_k}(\alpha)$ has a convergent subsequence, so we can choose a subsequence n'_k such that $R^{n'_k}(\alpha)$ converges. Then $v = \lim S^{n_k}(u) = \lim S^{n'_k}(u) = \lim P(R^{n'_k}(\alpha)) = P(\lim R^{n'_k}(\alpha))$ where in the last equality we have used the fact that P is continuous. Applying P^{-1} on both sides yields the result that P^{-1} is continuous.

Hence $P(x)$ is a bicontinuous bijection, except on a countable set, and $PR = SP$, while P sends the only invariant measure λ for R to the only invariant measure for S.

As for $\overline{R}$, it is semi-topologically conjugate to R, and semi-topological conjugacy is transitive. $\square$

Remark 2.90. The systems (X_u, S), $([0,1), R)$ and $(\mathbb{T}^1, \overline{R})$ are not mutually topologically conjugate: $\overline{R}$ and S are continuous while R is not, which is seen by looking at points on both sides of the point α, but topological conjugacy must preserve continuity of transformation.

Between S and $\overline{R}$, the topology of X_u is generated by clopen sets while the topology of $\mathbb{T}^1$ is not, since its topology is the subspace topology of the Euclidean topology on complex plane.

The system $(\mathbb{T}^1, \overline{R})$ has a discrete spectrum and the eigenvalues are all the $e^{2i\pi n\alpha}$, $n \in \mathbb{Z}$. For a proof of this fact, see results in Lemma 1.6.2 in the book. Also, $\overline{R}$ has rank one, though no explicit sequence of stacks generating the system has been found. Note that the Fibonacci substitution can be used to produce a sequence of stacks, in the same fashion as the Rokhlin stacks were constructed for the Morse sequence, but this would only give R a rank at most 2.

2.12 KD: Chacon Sequence - Mixing Properties

Lecture date: 2014-04-21.

The upshot of this discussion will be to state the definitions of two different properties of ergodic systems - namely weak and strong mixing - and then demonstrate a symbolic dynamical system which is weakly, but not strongly mixing.

Definition 2.91. A system (X, T, μ) is **weakly mixing** if for any two measurable sets A, B we have:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(A \cap T^k B) - \mu(A)\mu(B)| = 0.$$

Definition 2.92. A system (X, T, μ) is **strongly mixing** if for any two measurable sets A, B we have:

$$\lim_{n\to\infty} \mu(A \cap T^k B) = \mu(A)\mu(B).$$

It is a simple exercise in elementary analysis to see that strong mixing implies weak mixing. To see how both notions relate to ergodicity, we will state the following two results without giving their proof (the interested reader will find them, respectively, to be Theorem 1.17 and 1.26 in [4]).

Lemma 2.93. *A system (X, T, μ) is ergodic iff for any two measurable sets A, B we have:*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(A \cap T^k B) = \mu(A)\mu(B).$$

Lemma 2.94. *A system (X, T, μ) is weakly mixing iff all the eigenfunctions in $L^2(X, \mu)$ are the constants and 1 is a simple eigenvalue.*

From more simple considerations, weak mixing and strong mixing both imply ergodicity, but the converse is not the case.

We will now proceed to construct the apparatus necessary to demonstrate that weak mixing does not imply strong mixing.

Definition 2.95. Over the alphabet $\mathcal{A} = \{0, 1, 2\}$, the **Chacon substitution** δ is defined by:

$$0 \to 0012, 1 \to 12, 2 \to 012.$$

Clearly, the Chacon substitution is a primitive one; if we label the fixed point of δ starting with 0 by v, then (X_v, S_v) is a uniquely ergodic system (by Alex's second talk). We will briefly digress and study what the starting blocks of v look like through a study of δ.

Lemma 2.96. *For $n \geq 0$,*

$$\delta^n(1)\delta^n(2) = 12S_v(\delta^n(0)).$$

Proof. We proceed to prove the lemma by induction. For $n = 0$ we get

$$\delta^0(1)\delta^0(2) = 12 = 12S_v(0) = 12S_v(\delta^0(0)),$$

as desired.

Now suppose that for some $n \geq 0$ we have

$$\delta^n(1)\delta^n(2) = 12S_v(\delta^n(0)).$$

Applying δ to both sides preserves the equality. The LHS then easily equals $\delta^{n+1}(1)\delta^{n+1}(2)$. The RHS transforms as follows:

$$\delta(12S_v(\delta^n(0))) = \delta(1)\delta(2)\delta(S_v(\delta^n(0))).$$

We can now observe that $\delta(S_v(\delta^n(0))) = S_v^4(\delta^{n+1}(0))$, as the single shift in $S_v(\delta^n(0))$ eliminates the leading 0 in $\delta^n(0)$, and under δ that 0 maps to a 4-letter word. Hence,

$$\delta(12S_v(\delta^n(0))) = \delta(1)\delta(2)S_v^4(\delta^{n+1}(0)) = 12012S_v^4(\delta^{n+1}(0)).$$

Finally, the S_v^4 in the last expression removes the string 0012 from the beginning of $\delta^{n+1}(0)$, which implies that $012S_v^4(\delta^{n+1}(0)) = S_v(\delta^{n+1}(0))$. Combining everything, we get

$$\delta^{n+1}(1)\delta^{n+1}(2) = 12S_v(\delta^{n+1}(0)),$$

as desired. □

Corollary 2.97. *If we put $w_n = \delta^n(0)$, then we have $w_{n+1} = w_n w_n 12S_v(w_n)$ for all $n \geq 0$.*

Proof. For $n \geq 0$,

$$w_{n+1} = \delta^{n+1}(0) = \delta^n(\delta(0)) = \delta^n(0012) =$$

$$= \delta^n(0)\delta^n(0)\delta^n(1)\delta^n(2) = w_n w_n \delta^n(1)\delta^n(2).$$

By the preceding Lemma, we have $\delta^n(1)\delta^n(2) = 12S_v(\delta^n(0)) = 12S_v(w_n)$, which gives us

$$w_{n+1} = w_n w_n 12S_v(w_n),$$

as desired. □

It would serve us well to define the numbers $h_n = \frac{3^{n+1}-1}{2}$; clearly by the preceding Corollary, $|w_n| = h_n$.

Having constructed the system (X_v, S_v) we will now proceed to demonstrate a system which is topologically semi-conjugate to it, which will be weakly, but not strongly mixing.

Definition 2.98. Let the substitution σ on the alphabet $\mathcal{A} = \{0, 1\}$ be defined as follows:

$$0 \to 0010, 1 \to 1.$$

It is easy to check that if we denote $b_n = \sigma^n(0)$, then

$$b_{n+1} = \sigma^{n+1}(0) = \sigma^n(\sigma(0)) = \sigma^n(0010) =$$

$$= \sigma^n(0)\sigma^n(0)\sigma^n(1)\sigma^n(1) = b_n b_n 1 b_n.$$

Hence, $|b_n| = h_n$. If we then take a sequence of sequences u_n with first h_n letters given by b_n, these will converge to a limit, which we will denote by u. It is worth noting that u is the fixed point of σ; unfortunately, σ is not a primitive substitution, so we will have to resort to other means of establishing the unique ergodicity of the system (X_u, S_u). The next Lemma and its Corollary will suffice to that end; we defer the rather technical proof of the lemma to the end of the exposition.

Lemma 2.99. *(X_v, S_v) and (X_u, S_u) are semi-topologically conjugate.*

Corollary 2.100. *(X_u, S_u) is a uniquely ergodic system with a measure μ which is precisely the factor-frequency measure on cylinders.*

Proof. The unique ergodicity of the system follows from the preceding Lemma and the fact that topological semi-conjugacy preserves unique ergodicity. The second part of the claim will follow once we construct the conjugating map $f : X_v \to X_u$ in the proof of the Lemma. $\square$

We will now demonstrate that (X_u, S_u) is a rank one system. Consider the following geometric construction, illustrated in Figure 7. All the intervals that we are working with are left-closed-right-open.

Step 1) Put $P_0 = [0, \frac{2}{3})$, $P_1 = [\frac{2}{3}, 1)$ and $F_0 = P_0$. Here, F_0 is the base of out 0-th stack and the whole 0-th stack itself.

Step 2) At each successive step, vertically cut stack n into three equal columns and take the leftmost part of P_1 which has not been used up yet and has length equal to $\frac{2}{3^{n+2}}$; call that piece a_{n+1}. Note that the length of a_{n+1} will coincide with the lengths of all the intervals obtained from the cutting of the n-stack. With all the pieces thus defined, obtain the $(n+1)$-stack by putting the second column of the n-stack on top of the first, then putting a_{n+1} on top of the second column, and finally putting the third column of the n-stack on top of a_{n+1}.

Going back to the definition of a rank one system, it is worth noting how our geometric construction fits in it before we proceed to identify it with (X_u, S_u). Clearly, the heights of the stacks satisfy the relationship

$$h_{n+1} = 3h_n + 1,$$

(these are the same h_n's as before). Furthermore, if we write $\alpha_{n,1} = 0$, $\alpha_{n,2} = 1$ and $C_{n,2,1} = a_{n+1}$, then by directly substituting into the definition of a rank one system, we see that the transitions and equations match exactly. The inequality

$$\sum_{n=0}^{\infty} \frac{h_{n+1} - 3h_n}{h_{n+1}} < +\infty$$

is satisfied as the LHS equals to

$$\sum_{n=0}^{\infty} \frac{1}{h_{n+1}} = \sum_{n=0}^{\infty} \frac{2}{3^{n+2}-1},$$

which trivially converges by a comparison to a geometric series. It is also worth noting that if we label the transition map between levels of the stacks by T, we have ascribed an image under T to each point in $[0,1)$ as the sum of the lengths of the a_n equals to $\frac{1}{3}$.

If we now map each point in $[0,1)$ to its P-name under T we obtain (up to sets of measure 0) a bijection between $[0,1)$ and X_u. Furthermore, by a simple induction, we can check that the measure of each level in the geometric stacks will coincide with the measure of the points in X_u which give the P-names of the level. All of those facts pieced together give us that (X_u, S_u) is a rank one system; for simplicity we will keep thinking of the method of construction of the stacks in terms of "cutting" and "columns." We will also duplicate the notation of F_n as the base of the n-stack of (X_u, S_u).

Lemma 2.101. *For $p \geq n \geq 1$,*

$$\mu(S_u^{h_p} F_n \cap F_n) - (\mu(F_n))^2 \geq \frac{1}{12}\mu(F_n).$$

Proof. First of all, observe that

$$\mu(F_n) = \frac{2}{3^{n+1}} \leq \frac{1}{4}$$

for $n \geq 1$. Hence, to prove the lemma, it suffices to demonstrate that

$$\mu(S_u^{h_p} F_n \cap F_n) \geq \frac{1}{3}\mu(F_n).$$

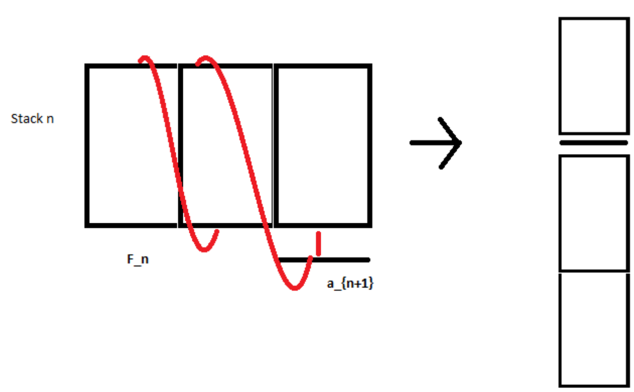

Figure 7: Geometric construction of the stacks.

For the next point reference Figure 8. For $p \geq n$, consider the p-stack cut into 3 "equal" columns. In the p-stack, all of the pieces of F_n were stacked vertically (with many other levels between them); when we cut the p stack into 3 equal parts, we cut each level in it which used to belong to F_n into 3 equal parts as well. Then, applying $S_u^{h_p}$ to the small sub-levels (thirds of the levels in the p-stack) in the first column just sends them over to the same level in the p stack, but in the second column. Hence, $S_u^{h_p}$ sends a subset of F_n of measure at least $\frac{1}{3}\mu(F_n)$ back into F_n which finalizes our proof. ☐

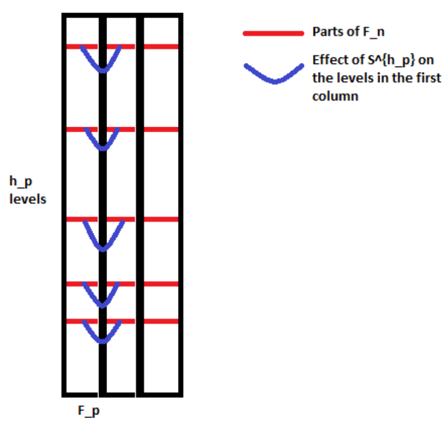

Figure 8: Visualization of the action of $S_u^{h_p}$ on the subsets of F_n in the p-stack.

Corollary 2.102. (X_u, S_u, μ) *is not strongly mixing.*

Proof. The preceding Lemma gives us that

$$limsup_{m\to\infty}\mu(S_u^m F_n \cap F_n) - (\mu(F_n))^2 \geq \frac{1}{12}\mu(F_n),$$

for $n \geq 1$, which means that the limit $lim_{m\to\infty}\mu(S_u^m F_n \cap F_n)$ cannot be $(\mu(F_n))^2$. Applying the definition of strong mixing, we reach the conclusion of the Corollary. ☐

To conclude the main point of this section, we will demonstrate that (X_u, S_u) is weakly mixing.

Lemma 2.103. (X_u, S_u) *is weakly mixing.*

Proof. To prove the result, we will resort to the spectral deinition of weak mixing. Let λ be an eigenvalue of (X_u, S_u) with an eigenfunction $f : \|f\| = 1$. By an argument similar to those encoutered in every integration theory textbook, we can approximate f in the L^2 sense by functions f_n: $\|f_n\| = 1$, f_n is constant on the levels of the n-stack and supported only on the levels of the n stack. Let $C_{1,n}, C_{2,n}, C_{3,n}$ be the three columns in which the n-stack is cut. Then, by all the properties of f_n and the fact that $\mu(C_{1,n}) = \mu(C_{2,n}) = \mu(C_{3,n})$, it follows that

$$\int_{C_{1,n}} |f_n|^2 = \int_{C_{2,n}} |f_n|^2 = \int_{C_{3,n}} |f_n|^2 = \frac{1}{3}.$$

A few more preliminaries are in order. First, applying $S_u^{h_n}$ to the sub-levels in $C_{1,n}$ just sends them to the same level of the n-stack in $C_{2,n}$. Similarly, applying $S_u^{h_n+1}$ to the sub-levels in $C_{2,n}$ just sends them to the same level of the n-stacl in $C_{3,n}$ (the difference comes from the addition of one extra level corresponding to the interval a_{n+1}).

Now, consider the expression

$$\int_{C_{1,n}} |f_n(S_u^{h_n}) - \lambda^{h_n} f_n|^2.$$

On one hand, by our last comment and the fact that f_n is constant on the levels of the n-stack, $f_n(S_u^{h_n}) = f_n$ on $C_{1,n}$. Hence,

$$\int_{C_{1,n}} |f_n(S_u^{h_n}) - \lambda^{h_n} f_n|^2 = \int_{C_{1,n}} |f_n|^2 |1 - \lambda^{h_n}|^2 = \frac{1}{3}|1 - \lambda^{h_n}|^2.$$

On the other hand, we have the following estimates (we use $|a+b|^2 \leq 2[|a|^2 + |b|^2]$):

$$\begin{aligned}
&\int_{C_{1,n}} |f_n(S_u^{h_n}) - \lambda^{h_n} f_n|^2 \leq \\
&\leq 2\left[\int_{C_{1,n}} |f_n(S_u^{h_n}) - f(S_u^{h_n})|^2 + \int_{C_{1,n}} |f(S_u^{h_n}) - \lambda^{h_n} f_n|^2\right] = \\
&= 2\left[\int_{C_{1,n}} |f_n(S_u^{h_n}) - f(S_u^{h_n})|^2 + \int_{C_{1,n}} |\lambda^{h_n} f - \lambda^{h_n} f_n|^2\right] \leq \\
&\leq 2\left[\int_{X_u} |f_n(S_u^{h_n}) - f(S_u^{h_n})|^2 + \int_{X_u} |\lambda^{h_n} f - \lambda^{h_n} f_n|^2\right] = \\
&= 4\|f - f_n\|^2,
\end{aligned}$$

where we use that f is a λ eigenfunction, $|\lambda| = 1$ and S_u is measure-preserving.

As the f_n approximate f in the L^2 sense, $4\|f - f_n\|^2$ goes to 0 as n goes to ∞. Combining this with our prior result, we get that $|1 - \lambda^{h_n}|$ goes to 0 as n goes to ∞, which means that λ^{h_n} goes to 1 as n goes to ∞. Repeating exactly the same argument but for $C_{2,n}$ will yield that λ^{h_n+1} goes to 1 as n goes to ∞. This implies that $\lambda = 1$ and since our system is ergodic (by unique ergodicity), 1 is a simple eigenvalue (by Alex's first talk). This fits exactly in the spectral definition of weak mixing and concludes our proof. $\square$

To bring closure to this presentation, we provide a proof of the fact that (X_v, S_v) and (X_u, S_u) are topologically semi-conjugate.

Proof. First, observe that all 2's in v occur after 1's, and all 1's occur exactly before 2's. Hence, the following argument reconstructs sequences in X_v from just the positions of the 1's in them; the forward and backward transitions will satisfy the conditions of topological semi-conjugacy.

Define the map $f : X_v \to X_u : f(\alpha)_i = \alpha_i \pmod 2$. In other words, we zero all the 2's. It is clear that f is continuous and $f \circ S_v = S_u \circ f$. Furthermore, by the structural properties of v and u we established earlier, both v and u are minimal and $f(v) = u$. This in turn implies that if we naturally extend the deinifition of f to words, $f(\mathcal{L}(v)) = f(\mathcal{L}(u))$. Hence, the range of f is trully a subset of X_u, and furthermore, the map is surjective.

The next fact we will use is Proposition 5.5.5 in [1]; we have that $p_u(n) = 2n - 1$ for $n \geq 2$. Hence, by my previous lecture, there will be no more than 3 sequences in X_u with more than one pre-image under S_u. Denote the set of those sequences by $P, |P| \leq 3$. Also denote all the set of all sequences in X_v that map to P by F; it is clear that F is an at most countable set. We will show that f defines a bijection between $X_v \backslash F$ and $X_u \backslash P$, and its inverse is also continuous (which will suffice for topological semi-conjugacy).

Let g be the map from X_u to sequences over $\mathcal{A} = \{0, 1, 2\}$ which sends the first digit of the sequence to 0, then simply copies all the 1's, and writes a 2 in each position where a 0 was preceded by a 1 in the argument sequence. All the remaining zeros are left as they are.

Now $f^{-1} = S_v \circ g \circ S_u^{-1}$ is a continuous function with domain $X_u \backslash P$ and range the sequences over $\mathcal{A} = \{0, 1, 2\}$ (the continuity follows from the continuity of S_u^{-1}, which is trivially established by looking at long-enough prefixes). Our problem is reduced to showing that $f \circ f^{-1} = Id_{X_u}$ on $X_u \backslash P$, $f^{-1} \circ f = Id_{X_v}$ on $X_v \backslash F$ and that the range of f^{-1} is a subset of X_v.

1) $f(f^{-1}(t)) = f(S_v(g(S_u^{-1}(t)))) = S_u(f(g(\alpha_t t))$ for some $\alpha_t \in \{0, 1\}$. Now, $S_u \circ f \circ g$ just erases the first digit of the sequence (as $f \circ g$ only changes the leading digit); hence, we get $f(f^{-1}(t)) = t$ for all $t \in X_u \backslash P$.

2) $f^{-1}(f(q)) = S_v(g(S_u^{-1}(f(q))))$. In that expression, first, f zeros all the 2's in q producing an element of $X_u \backslash P$; then S_u^{-1} returns the unique preimage of $f(q)$ in X_u by appending a digit to the start of the sequence. Finally, $S_v \circ g$ erases the first appended digit, but only after correctly and uniquely reconstructing q from the structure of the 1's in $S_u^{-1}(f(q))$ (here, we use that $f(q) \in X_u \backslash P$ to claim the uniqueness of the reconstruction). Hence, $f^{-1}(f(q)) = q$ for all $q \in X_v \backslash F$.

3) Finally, by the surjectivity of f and 2) we have that the range of f^{-1} is exactly X_v, which concludes our proof.

□

3 Chapter 6

Lecture date: 2014-05-05.

3.1 RJ: Sturmian Sequences: Frequency and minimality

In this section, we will discuss the notion of the frequency of **1** in a Sturmian sequence. We will see that is is well-defined, irrational, and that if two sequences have the same frequency, then they have the same language. Furthermore, we will also see that the frequency of a sequence is only dependent upon the language, which we will use to deduce that Sturmian sequences are minimal.

Theorem 3.1. *The frequency of* **1** *in a Sturmian sequence u, defined as the limit of*

$$\frac{|u_0 u_1 \cdots u_{n-1}|_1}{n}$$

as n tends to infinity, is well defined and irrational.

Proof. Let u be Sturmian. Write a_n as the minimum number of **1** that occurs in a factor of length n of u. Because u is Sturmian it is balanced. Thus, we know that

$$|u_0 u_1 \cdots u_{n-1}|_1 \in \{a_n, a_n + 1\}.$$

Thus, it is enough to show that the limit of a_n/n exists and is irrational.
A word of length $kq + r$ can be split into k words of length q and one word of length r. Thus, we get the inequality

$$ka_q \leq a_{kq+r} \leq k(a_q + 1) + r.$$

Let $n > q^2$. Then we can write $n = kq + r$ where $k \geq q$ and $0 \leq r < q$. Because $r < k$, we have

$$\frac{r}{n} < \frac{k}{n} \leq \frac{1}{q}.$$

The second part of this inequality gives

$$\frac{a_n}{n} = \frac{a_{kq+r}}{kq + r} \leq \frac{a_q}{q} + \frac{2}{q}.$$

It is also immediate that $ra_q - n$ is negative, because we have

$$n \geq a_n \geq ka_q > ra_q.$$

Thus, we have

$$\frac{n}{q}(a_q - 1) = ka_q + \frac{1}{q}(ra_q - n) \leq a_n.$$

We divide by n and obtain

$$\frac{a_q}{q} - \frac{1}{q} \leq \frac{a_n}{n} \leq \frac{a_q}{q} + \frac{2}{q}.$$

Thus, we see that $\{a_n/n\}_{n\in\mathbb{N}}$ is a Cauchy sequence, and thus converges to some limit α.
Now, suppose that $\alpha = p/q$. Because we have

$$ka_n \leq a_{kn} < a_{kn} + 1 \leq k(a_n + 1),$$

we have that if n divides n', then

$$\frac{a_n}{n} \leq \frac{a_{n'}}{n'} < \frac{a_{n'} + 1}{n'} \leq \frac{a_n + 1}{n}.$$

In particular, this gives that the sequence

$$\left\{\frac{a_{2^n q}}{2^n q}\right\}$$

is increasing, and

$$\left\{\frac{a_{2^n q} + 1}{2^n q}\right\}$$

is decreasing. Thus, we obtain

$$\frac{a_q}{q} \leq \frac{a_{2^n q}}{2^n q} < \frac{a_{2^n q} + 1}{2^n q} \leq \frac{a_q + 1}{q}.$$

But this sequence must converge to p/q, which is only possible if $a_q = p$ and $a_{2^n q} = 2^n p$ for all n, or $a_q + 1 = p$ and $a_{2^n q} + 1 = 2^n p$ for all n.
First, we show it is impossible to have $a_{2^n q} = 2^n a_q$ for all n. Because the sequence is not periodic, at least one word U of length q in the language of u is such that $|U|_1 = a_q + 1$. We see this by starting at the beginning, and noting that the first n letters completely determine our sequence if there is no word of length n with $|U|_1 = a_q + 1$. Because u is recurrent, this word occurs an infinite number of times, so that it must occur in two positions congruent $\bmod q$ by the infinite pigeon hole principle. Thus, we can find a word of length $2^n q$ which can be split into words of length q, which contains at least two occurrences of U. The number of **1** in this word is at least $2^n a_q + 2$, so that $a_{2^n q} > 2^n a_q$, which contradicts the first case.
To show that $a_{2^n q} + 1$ cannot equal $2^n(a_q + 1)$ for all n, we do essentially the same thing. We can find some word W with $|W|_1 = a_q$, else a_q would not be the true minimum. Again, because u is recurrent, we can find two occurrences of W in positions congruent $\bmod q$, call this word V. Thus, we have

$$|V|_1 \leq 2^n(a_q + 1) - 2,$$

because W occurs twice in this word. Because V is a word of length $2^n q$, we have

$$a_{2^n q} + 1 \leq 2^n(a_q + 1) - 2,$$

and $a_{2^n q} + 1 \neq 2^n(a_q + 1)$.

□

Now we define a property of balancedness for sets of finite words. These results will be used more later, but for now allow us to conclude that Sturmian sequences are minimal.

Definition 3.2. We say that a set E of finite words is **balanced** if for any pair of words U, V in E, and for any words U', V' of the same length that occur in U, V, we have that

$$\left| |U'|_1 - |V'|_1 \right| \leq 1.$$

Proposition 3.3. *A balanced set of words of length n contains at most $n+1$ distinct words.*

Proof. We proceed by induction. The base case is obvious, for if $n = 1$, the only possible words are 0 and 1. Assume the result is true for n. Assume for contradiction that there are at least $n+3$ distinct words of length $n+1$ in a balanced set of words E. Each of these words contains a prefix of length n. There are at most $n+1$ distinct prefixes by the inductive hypothesis. Enumerate these are $u_1, \ldots, u_{n+1}$. Then, from an easy counting argument we see that there must be at least two of these prefixes of length n, say u_1 and u_2, such that $u_1 0, u_1 1, u_2 0, u_2 1$ all occur in our set of distinct words. Call such words – words which are followed by both a 0 and a 1 – special words. Because our sequence is Sturmian, we see that for words of length $k < n-1$, there is only one special word. In particular, there is only one special word of length $n-1$. Now, note we can write

$$u_1 = au_1', u_2 = bu_2',$$

where $a, b \in \mathcal{A}$. Then u_1' and u_2' are special words of length $n-1$. As there is only one such word, we have that $u_1' = u_2' = U$. Thus, in order for all of our words to be distinct, we have $a \neq b$. Then, we have

$$0U0, 1U1$$

both in our sequence, which is impossible. $\square$

Proposition 3.4. *If u and v are Sturmian sequences with the same frequency, then they have the same language.*

Proof. Left as an exercise to the reader. □

Proposition 3.5. *The frequency of* **1** *in a Sturmian sequence depends only on the language, not on the sequence itself. From this, we can see that Sturmian sequences are minimal.*

Proof. Consider a sequence u with language $\mathcal{L}(u)$. Take any word of length n from the language. Because the language is balanced, the frequency of **1** for this word is a_n or $a_n + 1$, where a_n is as before. Thus, the limit of a_n/n is the same for this word as any other word of length n in the language.
Let u be a Sturmian sequence. Suppose there is some factor W of u such that W does not appear with bounded gaps. Let $\{n_k\}$ be the sequence of positions of the end of the word, which goes to infinity. Then we create a sequence by $\{S^{n_k}u\}$, where we shift u by n_k at each step. Because our space of sequences is compact, this subsequence converges to a limit v whose language $\mathcal{L}(v)$ by definition does not contain u. But because v is in the orbit closure of u, they have the same language, a contradiction. □

3.2 AZ: Basic Properties of Sturmian Sequences

Lecture date: 2014-05-05.

We begin by recalling a few preliminary facts about sequences from Chapter 1. Let u be a sequence with values in a finite alphabet $\mathcal{A}$.

The language of u is the set $\mathcal{L}(u)$ of finite words that occur in u, and we denote by $\mathcal{L}_n(u)$ the set of words of length n that occur in u.

Recall from Definition 1.6 that the complexity function of u is the function p_u which associates the number Card $\mathcal{L}_n(u)$ of distinct words of length n that occur in u to each integer n. p_u is an increasing function, and if u is eventually periodic (also called ultimately periodic; see Definition 1.4), p_u is bounded. If there is an n such that $p_u(n+1) = p_u(n)$, then u is eventually periodic.

We can see that if u is not eventually periodic, we must have $p_u(n) \geq n+1$ because p_u has strictly increasing integral values and $p_u(1) \geq 2$.

Definition 3.6. A sequence u is called **Sturmian** if it has complexity $p_u(n) = n+1$. We will denote by Σ' the set of Sturmian sequences.

Because $p_u(1) = 2$, Sturmian sequences contain only two letters, so we may fix the alphabet $\mathcal{A} = \{0, 1\}$.

Proposition 3.7. *A Sturmian sequence is* ***recurrent****.*

Recall from Definition 1.2 that a sequence is recurrent if each word in $\mathcal{L}(u)$ occurs an unbounded number of times.

Proof. Suppose that a word U, of length n, occurs in a Sturmian sequence u a finite number of times, and does not occur after rank N. Let v be the sequence defined by $v_k = u_{k+N}$. It is clear that the language of v is contained in that of u, and does not contain U. Hence we must have $p_v(n) \leq n$. This implies that v is eventually periodic, and hence so is u, a contradiction. □

Lemma 3.8. *If u is Sturmian, then exactly one of the words* ***00****,* ***11*** *does not occur in u.*

Proof. We have $p_u(2) = 3$, so there are exactly three words of length 2 occurring in u. By the previous proposition, **0** and **1** each occur an infinite number of times in u, which implies that **01** and **10** both occur in u. But **00** and **11** are the two other words of length 2, and exactly one of them must occur. □

Definition 3.9. If U is a finite word over the alphabet $\mathcal{A}$, we denote by $|U|$ the length of U, and $|U|_{\mathbf{a}}$ the number of occurrences of the letter **a** in U.

Definition 3.10. A sequence u over the alphabet $\{\mathbf{0}, \mathbf{1}\}$ is **balanced** if, for any pair of words U, V of the same length occurring in u, we have $||U|_{\mathbf{1}} - |V|_{\mathbf{1}}| \leq 1$.

Remark 3.11. For any word in $\mathcal{L}_n(u)$, the number of occurrences of **1** can have one of two consecutive integral values - this is the smallest possible number of values for non-periodic sequences. We will show that this property is equivalent to being Sturmian.

Lemma 3.12. *If the sequence u is not balanced, there is a (possibly empty) word W such that $\mathbf{0}W\mathbf{0}$ and $\mathbf{1}W\mathbf{1}$ occur in u.*

Proof. If u is not balanced, we can find two words A and B of length n such that $|A|_{\mathbf{1}} - |B|_{\mathbf{1}} > 1$. First, we prove that we can suppose that $|A|_{\mathbf{1}} - |B|_{\mathbf{1}} = 2$.

Let A_k (respectively B_k) denote the suffix of length k of A (respectively B), and put $d_k = |A_k|_{\mathbf{1}} - |B_k|_{\mathbf{1}}$. We have $d_n > 1$ and $d_0 = 0$, and $|d_{k+1} - d_k| = 0$ or 1. We can see from an intermediate value argument in integer-valued functions that there is k such that $d_k = 2$.

Suppose now that A and B are words of minimal length with $d_n = 2$. We then write $A = \mathbf{a}_0\mathbf{a}_1...\mathbf{a}_{n-1}$ and $B = \mathbf{b}_0\mathbf{b}_1...\mathbf{b}_{n-1}$. We must have $\mathbf{a}_0 = \mathbf{a}_{n-1} = \mathbf{1}$ and $\mathbf{b}_0 = \mathbf{b}_{n-1} = \mathbf{0}$; otherwise, we could find a shorter pair by removing some prefix. We then have $\mathbf{a}_k = \mathbf{b}_k$ for all remaining letters (just repeatedly remove the last letter of both words and note that d_n at each step is equal to 1). □

Theorem 3.13. *A sequence u is Sturmian if and only if it is a non-eventually periodic balanced sequence over two letters.*

Proof. We begin by proving the "if" direction by proving its contrapositive. We show that, if u is not Sturmian, then u is not balanced. Let n_0 be the smallest integer such that $p_u(n_0 + 1) \geq n_0 + 3$. We have $n_0 \geq 1$ because $p_u(1) = 2$.

Because $p_u(n_0) = n_0 + 1$, there are at least two words U and V of length n_0 that can be extended on the right in two ways. Because n_0 is the smallest integer with this property, U and V differ only in the first letter. Without loss of generality, there is a word W such that $U = \mathbf{0}W$ and $V = \mathbf{1}W$, so both $\mathbf{0}W\mathbf{0}$ and $\mathbf{1}W\mathbf{1}$ occur in u, so u is not balanced.

We now prove the converse - if u belongs to Σ', it is a non-eventually periodic balanced sequence over two letters. We again use proof by contradiction - we suppose that u is not balanced. By Lemma 3.12, there is a word W such that $\mathbf{1}W\mathbf{1}$ and $\mathbf{0}W\mathbf{0}$ occur in u. Consider the minimal such word with length $n + 1$, and write $W = w_0w_1...w_n$.

Note that, due to our choice of W and n, if we have a pair of words U, V of the same length such that $||U|_{\mathbf{1}} - |V|_{\mathbf{1}}| \geq 2$, then their length is at least $n + 3$. This is true because we apply the same argument as in the proof of Lemma 3.12 to see that (without loss of generality) there is a word W' such that $U = \mathbf{0}W'\mathbf{0}$ and $V = \mathbf{1}W'\mathbf{1}$. This means that $\mathbf{0}W\mathbf{0}$ and $\mathbf{1}W\mathbf{1}$ are a pair of non-balanced words of minimal length.

W cannot be empty by Lemma 3.8. Additionally, we have $w_0 = w_n$ because the words $\mathbf{0}W\mathbf{0}$ and $\mathbf{1}W\mathbf{1}$ would contain both $\mathbf{00}$ and $\mathbf{11}$, which contradicts the statement of Lemma 3.8 because u is Sturmian. Furthermore, we must have $w_k = w_{n-k}$ for all k. If not, take the smallest k such that $w_k \neq w_{n-k}$ and suppose without loss of generality that $w_k = \mathbf{0}$ and $w_{n-k} = \mathbf{1}$. We then see that $\mathbf{0}w_0...w_{k-1}\mathbf{0}$ and $\mathbf{1}w_{n-k+1}...w_n\mathbf{1}$ is a non-balanced pair of smaller length, which contradicts our choice of W. Note that both of these words are indeed in $\mathcal{L}_u$ because $\mathbf{0}W\mathbf{0}$ and $\mathbf{1}W\mathbf{1}$ are in $\mathcal{L}_u$. Thus W is a palindrome.

Because u is Sturmian, we know that there are $n + 2$ words of length $n + 1$. W is the only word of this length that can be extended on the right in two ways and on the left in two ways. Once against using the fact that u is Sturmian, exactly one of $\mathbf{0}W$ and $\mathbf{1}W$ can be extended in two ways on the right. Suppose it is $\mathbf{0}W$. Then $\mathbf{0}W\mathbf{0}$, $\mathbf{0}W\mathbf{1}$, and $\mathbf{1}W\mathbf{1}$ all occur in u, but $\mathbf{1}W\mathbf{0}$ does not. Let i be the rank of an occurrence of $\mathbf{1}W\mathbf{1}$. We now prove a lemma.

Lemma 3.14. *The word* $\mathbf{0}W$ *cannot occur in* $u_i u_{i+1} ... u_{i+2n+3}$.

The length of $u_i u_{i+1} ... u_{i+2n+3}$ is $2n+4$, and the length of $\mathbf{1}W\mathbf{1}$ is $n+3$, and the length of $\mathbf{0}W$ is $n+2$. Thus, the statement of this lemma is that the first letter of an occurrence of $\mathbf{0}W$ cannot occur in an occurrence of $\mathbf{1}W\mathbf{1}$ in u.

Proof. Suppose that the beginning of $\mathbf{0}W$ overlaps with $\mathbf{1}W\mathbf{1}$. Let w_k be the letter on which $\mathbf{0}W$ begins, as $\mathbf{0}W$ cannot begin on the first or last letter of $\mathbf{1}W\mathbf{1}$. We then have $\mathbf{0}w_0...w_{n-k} = w_k w_{k+1}...w_n\mathbf{1}$. But this implies that $w_k = \mathbf{0}$ and $w_{n-k} = \mathbf{1}$, which is a contradiction because W is a palindrome. $\square$

Continuation of the proof of Theorem 3.13: We see that there are $n+3$ words of length $n+2$ occurring in $u_i u_{i+1} ... u_{i+2n+3}$. But there are $n+3$ words of length $n+2$ in u and $\mathbf{0}W$ does not occur due to the above lemma. Thus, at least one word occurs twice. But all of these worsd can be extended to the right in only one way, as $\mathbf{0}W$ is the only word of length $n+2$ that can be extended to the right in two ways. Thus u is eventually periodic, which is a contradiction. $\square$

References

[1] N. Fogg and V. Berthé. *Substitutions in Dynamics, Arithmetics and Combinatorics*. Lecture Notes in Mathematics. Springer, 2002.

[2] P. Halmos. *Lectures on Ergodic Theory*. AMS Chelsea Publishing Series. Chelsea Publishing Company, 1956.

[3] P. Kůrka. *Topological and Symbolic Dynamics*. Collection SMF. Société mathématique de France, 2003.

[4] P. Walters. *An Introduction to Ergodic Theory*. Graduate texts in mathematics. U.S. Government Printing Office, 2000.