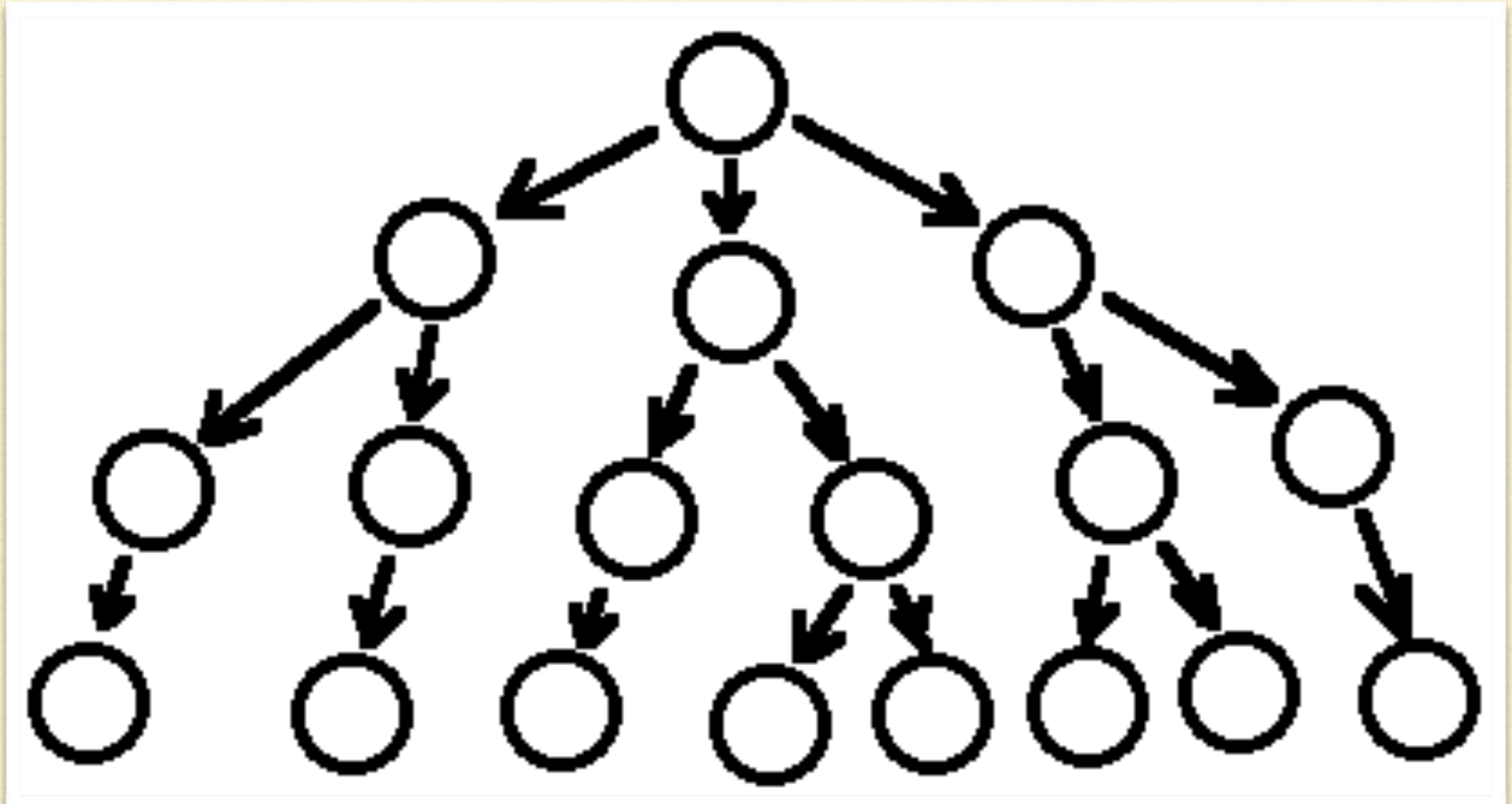
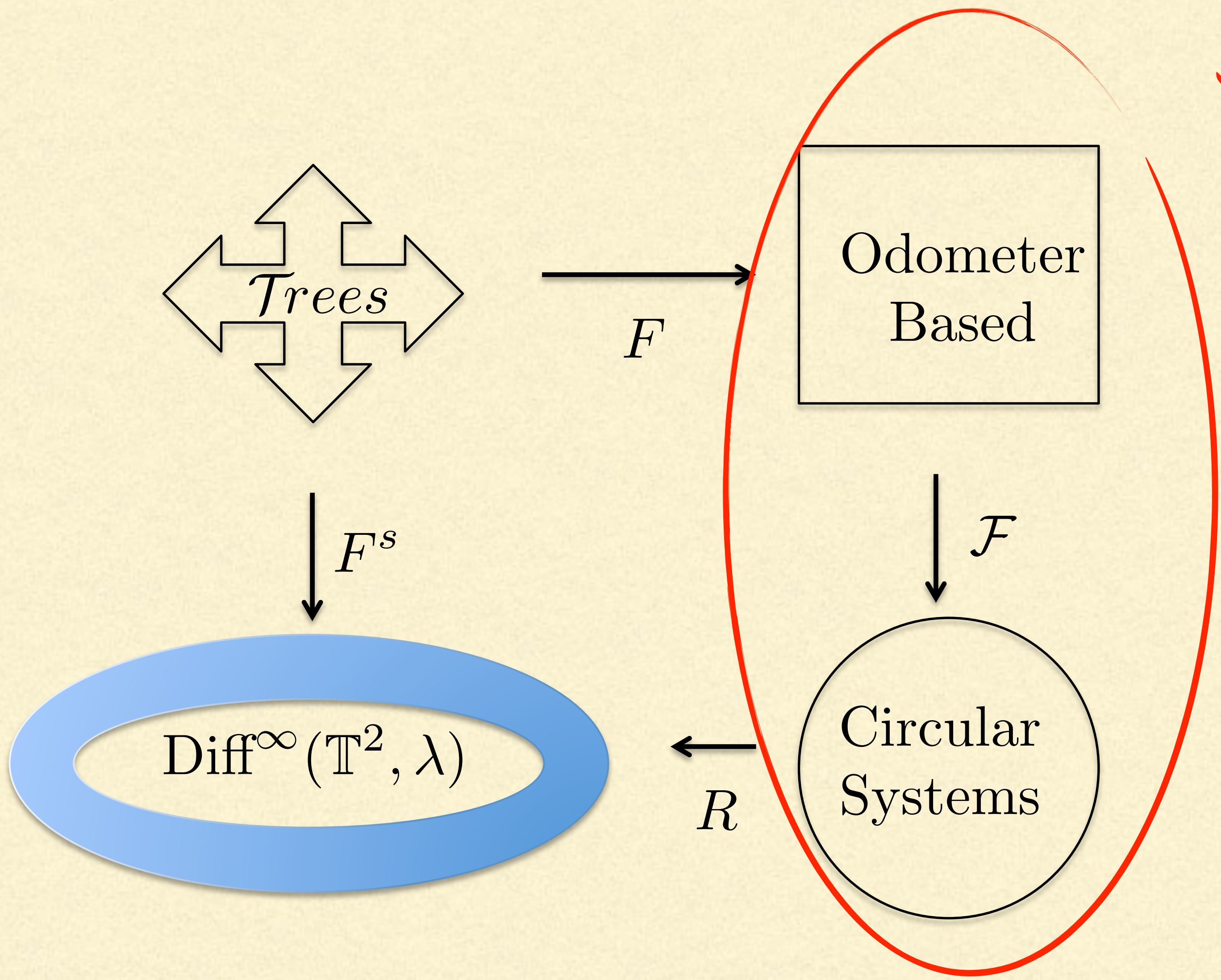


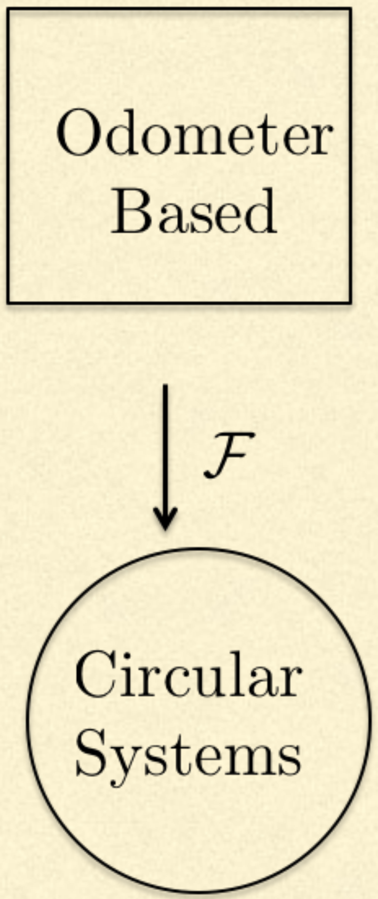
A GLOBAL STRUCTURE THEOREM

RECALL



A global structure theorem

TODAY'S FOCUS



THESE RESULTS ARE IN A PREPRINT

arXiv.org > math > arXiv:1703.07093

Mathematics > Dynamical Systems

From Odometers to Circular Systems: A global structure theorem

Matthew Foreman, Benjamin Weiss

THE IDEA

Instead of looking at MPT's one at a time in a test tube, and proving structure theorems for each one separately, we look at the whole ecosystem of MPT's.

Global Structure Theorems

Consider the space EMPT of ergodic measure preserving transformations of $[0,1]$. This is a dense invariant $\mathcal{G}_\delta$ subset of the group MPT of all measure preserving transformations $[0,1]$.

Structure on EMPT

- Extensions/Factors $\pi : ([0,1], S) \to ([0,1], T)$

Structure on EMPT

- Extensions/Factors $\pi : ([0,1], S) \to ([0,1], T)$

- Weak Mixing Extensions

Structure on EMPT

- Extensions/Factors $\pi : ([0,1], S) \to ([0,1], T)$
- Weak Mixing Extensions
- Group Extensions

Structure on EMPT

- Extensions/Factors $\pi : ([0,1], S) \to ([0,1], T)$
- Weak Mixing Extensions
- Group Extensions
- Compact Extensions

Structure on EMPT

- Extensions/Factors $\pi : ([0,1], S) \to ([0,1], T)$

- Weak Mixing Extensions

- Group Extensions

- Compact Extensions

- General Joinings

A STRANGE QUESTION

Is there an automorphism of EMPT that preserves this structure?

A STRANGE QUESTION

Is there an automorphism of EMPT that preserves this structure?

How rigid is EMPT?

A THEOREM PROVED BY FATHI?

Every automorphism of the group of measure preserving transformations is *inner*.

A THEOREM PROVED BY FATHI?

Every automorphism of the group of measure preserving transformations is *inner*.

Is there a map between large classes of measure preserving transformations that preserves all of the structure?

Form a Category

- Take as objects an invariant set $C \subseteq EMPT$

- Take as morphisms some set of joinings (e.g. extensions, weakly mixing extensions ...)

What categories of EMPT's are functorially isomorphic?

What categories of EMPT's are functorially isomorphic?

Functorial isomorphisms are bijections between

1. objects and objects
2. morphisms and morphisms

and preserve compositions of the *morphisms*.

AN ARTIFICIAL EXAMPLE

Fix:

1. A measure isomorphism $J : [0,1] \to [0,1] \times [0,1]$
2. A weakly mixing transformation S.

Then the function $T \mapsto J^{-1}(T \times S)J$ preserves isomorphism and factors

Not quite a functor without more conditions.

A USEFUL EXAMPLE

The category of odometer based MPTS and the category of circular systems are functorially isomorphic.

A USEFUL EXAMPLE

The category of odometer based MPTS and the category of circular systems are functorially isomorphic.

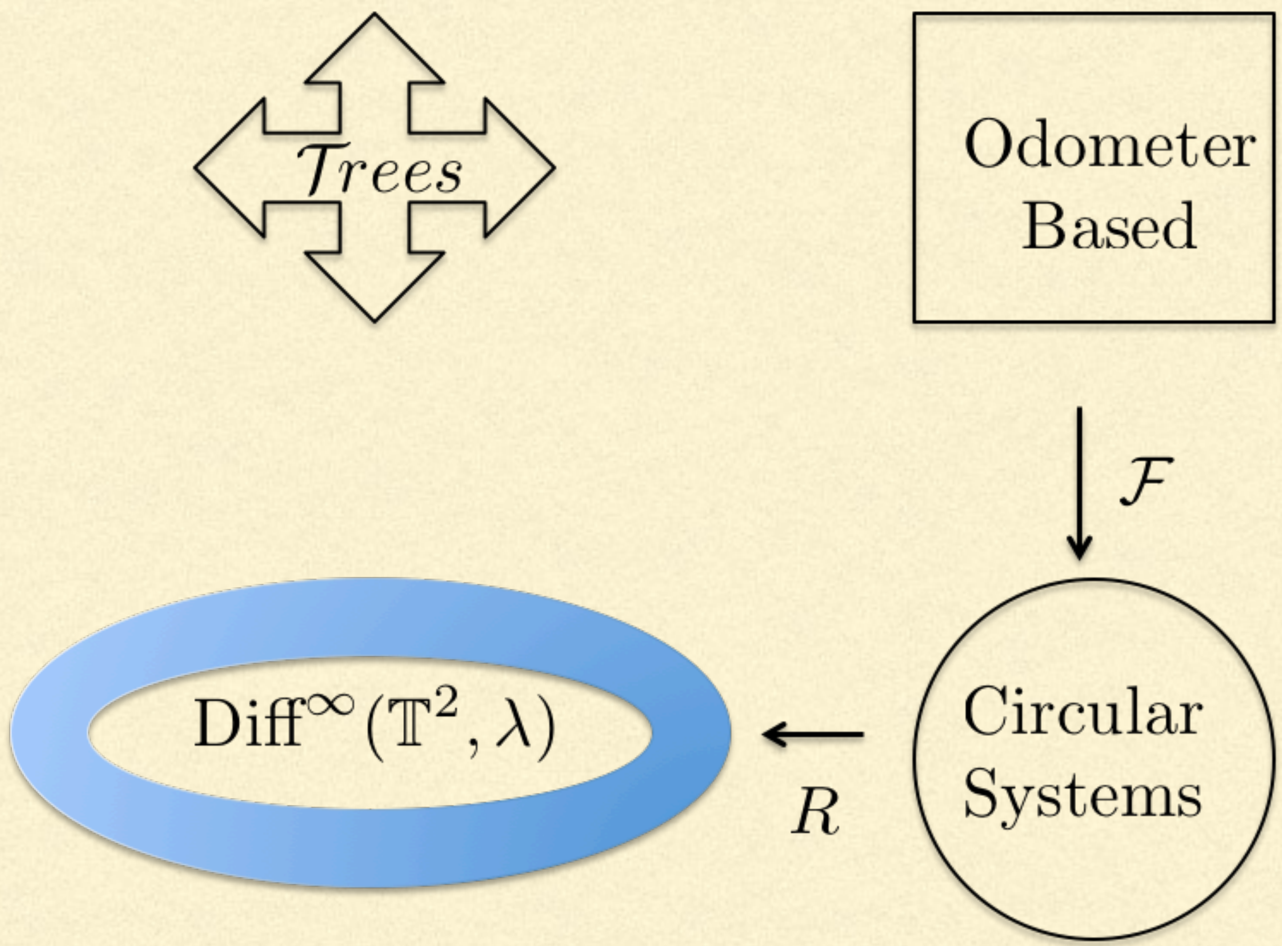
What?

What are the categories?

What are the morphisms?

WTF?

PUTTING THIS TOGETHER WITH THE REALIZATION THEOREM



SOMETHING THAT ODOMETER BASED AND CIRCULAR SYSTEMS HAVE IN COMMON

Both the odometer based systems and the circular systems have an underlying *timing factor*. Either

- An odometer or
- a Liouvillean rotation.

SOMETHING THAT ODOMETER BASED AND CIRCULAR SYSTEMS HAVE IN COMMON

Both the odometer based systems and the circular systems have an underlying *timing factor*. Either

- An odometer or
- a Liouvillean rotation.

Both timing factors have a canonical involution

- For odometers $x \mapsto -x$ drive your car in reverse
- For rotations $z \mapsto \bar{z}$ put a mirror on the real axis

SPECIAL FACTOR MAPS

Let (X, S) and (Y, T) be ergodic MPT's that are either both odometer based or both circular. Suppose that $\pi : X \to Y$ is a factor map. Then π is:

- *synchronous* if it preserves the underlying timing factor

- *anti-synchronous* if it maps the underlying timing factor the underlying timing factor by the canonical involution.

TWO CATEGORIES:

Odometer Based Systems

- **Objects**: Ergodic Odometer Based Transformations based on $\langle k_n : n \in \mathbb{N} \rangle$
- **Morphisms**: Synchronous and anti-synchronous factor maps

Circular Systems

- **Objects**: Ergodic Circular Systems based on $\langle k_n, l_n : n \in \mathbb{N} \rangle$
- **Morphisms**: Synchronous and anti-synchronous factor maps

A GLOBAL STRUCTURE THEOREM

There is a functor $\mathcal{F}$ that is an isomorphism between the odometer based systems and the circular systems.

THEOREM (MORE EXPLICITLY)

There is a functor $\mathcal{F}$ that is an isomorphism between the odometer based systems and the circular systems.

$\mathcal{F}$ takes:

1. isomorphisms to isomorphisms
2. factor maps to factor maps
3. (Glasner) compact factors to compact factors
4. weakly mixing factors to weakly mixing factors

PROPERTIES OF $\mathcal{F}$

$\mathcal{F}$ takes distal towers to distal towers (and preserves height).

PROPERTIES OF $\mathcal{F}$

$\mathcal{F}$ takes distal towers to distal towers (and preserves height).

$\mathcal{F}$ takes rank one transformations to rank one transformations

PROPERTIES OF $\mathcal{F}$

$\mathcal{F}$ takes distal towers to distal towers (and preserves height).

$\mathcal{F}$ takes rank one transformations to rank one transformations

$\mathcal{F}$ preserves simplexes of invariant measures.

DOWNAROWICZ THEOREM

If K is a Choquet simplex then there is an odometer based system $\mathbb{L}$ (a Toeplitz sequence) such that the collection of invariant measures on $\mathbb{L}$ is affinely homeomorphic to K.

AN ACCIDENTAL APPLICATION OF OUR METHOD

Corollary to the Global Structure theorem: Every compact Choquet simplex is the simplex of invariant measures of a circular system.

Better lucky than smart!

TO PROVE THIS THEOREM, WE NEED TO GIVE FINITISTIC CRITERION FOR ALL OF THESE BEHAVIORS THAT WE CAN TRANSFER VIA THE FUNCTOR.

MORE ON SYMBOLIC SHIFTS

WE'LL NEED TO WORK ON A SPECIAL SET OF MEASURE ONE

Suppose that $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ is a construction sequence for a symbolic system $\mathbb{K}$ with each $\mathcal{W}_n$ uniquely readable. Let S be the collection $x \in \mathbb{K}$ such that there are sequences of natural numbers $\langle a_m : m \in \mathbb{N} \rangle$, $\langle b_m : m \in \mathbb{N} \rangle$ going to infinity such that for all m there is an n, $x \upharpoonright [-a_m, b_m) \in \mathcal{W}_n$.

- Suppose that there is a unique invariant measure ν on $S \subseteq \mathbb{K}$, then ν is ergodic

- Suppose that $\mathbb{K}$ is built from $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ and ν is a shift invariant measure on $\mathbb{K}$ concentrating on S. Then for ν-almost every s there is an N for all $n > N$, there are $a_n \leq 0 < b_n$ such that $s \upharpoonright [a_n, b_n) \in \mathcal{W}_n$.

LOCATIONS

Let $s \in S$ and suppose that for some $0 \leq k < q_n$, $s \upharpoonright [-k, q_n - k) \in \mathcal{W}_n$. We define $r_n(s)$ to be the unique k with with this property. We will call the interval $[-k, q_n - k)$ the *principal n-block* of s, and $s \upharpoonright [-k, q_n - k)$ its *principal n-subword*. The sequence of r_n's will be called the *location sequence of s.*

LOCATIONS

Let $s \in S$ and suppose that for some $0 \le k < q_n$, $s \restriction [-k, q_n - k) \in \mathcal{W}_n$. We define $r_n(s)$ to be the unique k with with this property. We will call the interval $[-k, q_n - k)$ the *principal n-block* of s, and $s \restriction [-k, q_n - k)$ its *principal n-subword*. The sequence of r_n's will be called the *location sequence of s*.

If ν concentrates on S, then for ν-almost all s the principal n-subwords exist for all large n.

WHY ARE THESE HELPFUL?

- Suppose that $s, s' \in S$ and $\langle r_n(s) : n \geq N \rangle = \langle r_n(s') : n \geq N \rangle$ and for all $n \geq N$, s and s' have the same principal n-subwords. Then $s = s'$.

- Suppose that we are given a sequence $\langle u_n : M \leq n \rangle$ with $u_n \in \mathcal{W}_n$. If we specify which occurrence of u_n in u_{n+1} is the principal occurrence, and the distances of the principle occurrence to the beginning of u_{n+1} go to infinity, then $\langle u_n : M \leq n \rangle$ determines an $s \in S \subseteq \mathbb{K}$ completely up to a shift k with $|k| \leq q_M$.

WHY ARE THESE HELPFUL?

- A sequence $\langle r_n : N \leq n \rangle$ and sequence of words $w_n \in \mathcal{W}_n$ comes from an infinite word $s \in S$ if both r_n and $q_n - r_n$ go to infinity and that the r_{n+1} position in w_{n+1} is in the r_n position in a subword of w_{n+1} identical to w_n.

  Caveat: just because $\langle r_n : N \leq n \rangle$ is the location sequence of some $s \in S$ and $\langle w_n : N \leq n \rangle$ is the sequence of principal subwords of some $s' \in S$, it does not follow that there is an $x \in S$ with location sequence $\langle r_n : N \leq n \rangle$ and sequence of subwords $\langle w_n : N \leq n \rangle$.

- If $x, y \in S$ have the same principal n-subwords and $r_n(y) = r_n(x) + 1$ for all large enough n, then $y = sh(x)$.

ODOMETER BASED SYSTEMS CONCENTRATE ON S

Let $\mathbb{K}$ be an odometer based system and suppose that ν is a shift invariant measure. Then ν concentrates on S.

ODOMETER BASED SYSTEMS CONCENTRATE ON S

$\vdash$ Let $B = \mathbb{K} \setminus S$. Then B is shift invariant. Suppose that ν gives B positive measure. For $s \in B$ let $a_n(s) \leq 0 \leq b_n(s)$ be the left and right endpoints of the principal n-block of s. Then for all $s \in B$ there is an $N \in \mathbb{N}$ such that:

1. for all $n, -N \leq a_n$ or
2. for all $n, b_n \leq N$.

We assume that ν gives the collection B^* of s such that there is an $N \in \mathbb{N}$ for all $n, -N \leq a_n$ positive measure, the other case is similar.

Define $f : B^* \to \mathbb{N}$ by setting $f(s) =$ least N satisfying item 1. Then f is a Borel function. Let $B_n = f^{-1}(n)$. Then the B_n's are disjoint, $B^* = \bigcup_{n \geq 0} B_n$ and $sh^{-1}(B_n) = B_{n+1}$. Hence for all $n, m, \nu(B_n) = \nu(B_m)$, a contradiction. $\dashv$

NOT TRUE FOR CIRCULAR SYSTEMS

If $\mathbb{K}$ is a circular system then there is an $x \in \mathbb{K}$ such that a left tail of x is constantly b. Shifting this gives the constant $\mathbb{Z}$ sequence consisting of b's in $\mathbb{K}$. Similarly, the constant e sequence is in $\mathbb{K}$. Since these are shift invariant they each support invariant atomic measures.

NOT TRUE FOR CIRCULAR SYSTEMS

If $\mathbb{K}$ is a circular system then there is an $x \in \mathbb{K}$ such that a left tail of x is constantly b. Shifting this gives the constant $\mathbb{Z}$ sequence consisting of b's in $\mathbb{K}$. Similarly, the constant e sequence is in $\mathbb{K}$. Since these are shift invariant they each support invariant atomic measures.

However: If $\mathbb{K}$ is a circular system and ν is a non-atomic shift invariant measure then ν concentrates on S.

SINGLE ORBIT DYNAMICS

SINGLE ORBIT DYNAMICS

Understanding the ergodic theorem in the context of symbolic shifts

Let T be a measure preserving transformation from (X,μ) to (X,μ), where X is a compact metric space. Let $C(X)$ be the space of all complex valued functions. Then a point $x \in X$ is *generic* for T if and only if for all $f \in C(X)$,

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

The Ergodic Theorem tells us that for a given f and ergodic T equation above holds for a set of μ-measure one. Intersecting over a countable dense set of f gives a set of μ-measure one of generic points.

The Ergodic Theorem tells us that for a given f and ergodic T equation above holds for a set of μ-measure one. Intersecting over a countable dense set of f gives a set of μ-measure one of generic points.

For symbolic systems $\mathbb{K} \subseteq \Sigma^{\mathbb{Z}}$ we can describe generic points x as being those x such that the μ-measure of all basic open intervals $\langle u \rangle_0$ is equal to the density of k such that u occurs in x at k.

Let μ be a shift invariant measure on a symbolic system $\mathbb{K}$ defined by a uniquely readable construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ in a finite language Σ. Assume that q_n is the length of the words in $\mathcal{W}_n$.

By μ_m we will denote the discrete measure on the finite set Σ^m given by $\mu_m(u) = \mu(\langle u \rangle)$. By $\hat{\mu}_n(w)$ we will denote the discrete probability measure on $\mathcal{W}_n$ defined by

$$\hat{\mu}_n(w) = \frac{\mu_{q_n}(\langle w \rangle)}{\sum_{w' \in \mathcal{W}_n} \mu_{q_n}(\langle w' \rangle)}.$$

By μ_m we will denote the discrete measure on the finite set Σ^m given by $\mu_m(u) = \mu(\langle u \rangle)$. By $\hat{\mu}_n(w)$ we will denote the discrete probability measure on $\mathcal{W}_n$ defined by

$$\hat{\mu}_n(w) = \frac{\mu_{q_n}(\langle w \rangle)}{\sum_{w' \in \mathcal{W}_n} \mu_{q_n}(\langle w' \rangle)}.$$

Thus $\hat{\mu}_n(w)$ is the relative measure of $\langle w \rangle$ among all $\langle w' \rangle, w' \in \mathcal{W}_n$. The denominator is a normalizing constant to account for spacers at stages $m > n$ and for shifts of size less than q_n.

Explicitly, if $A_n = \{s \in \mathbb{K} : s(0)$ is the start of a word in $\mathcal{W}_n\}$, then the sets $\{sh^j(A_n)\}_{j=0}^{q_n-1}$ are disjoint and their union has a measure that tends to one as n grows to infinity. The set A_n is partitioned into $|\mathcal{W}_n|$ many sets by the words $w \in \mathcal{W}_n$ and $\hat{\mu}_n$ gives their relative size in A_n. Since the measure of an arbitrary finite cylinder set can be calculated along the individual columns represented by a fixed w, it is clear that the $\hat{\mu}_n(w)$ determine uniquely the measure μ.

A word w in $\Sigma^{q_{k+l}}$ determines a unique sequence of words w_j in $\mathcal{W}_k$ such that ,

$$w = u_0 w_0 u_1 w_1 \ldots w_J u_{J+1}.$$

When $w \in \mathcal{W}_{k+l}$, each u_j is in the region of spacers added in $\mathcal{W}_{k+l'}$, $l' \leq l$. We will denote the *empirical distribution* of $\mathcal{W}_k$-words in w by $\mathrm{EmpDist}_k(w)$. Formally:

$$\mathrm{EmpDist}_k(w)(w') = \frac{|\{0 \leq j \leq J : w_j = w'\}|}{J+1}, \; w' \in \mathcal{W}_k.$$

Then *EmpDist* extends to a measure on $\mathcal{P}(\mathcal{W}_k)$ in the obvious way.

A word w in $\Sigma^{q_{k+l}}$ determines a unique sequence of words w_j in $\mathcal{W}_k$ such that ,

$$w = u_0 w_0 u_1 w_1 \ldots w_J u_{J+1}.$$

When $w \in \mathcal{W}_{k+l}$, each u_j is in the region of spacers added in $\mathcal{W}_{k+l'}$, $l' \leq l$. We will denote the *empirical distribution* of $\mathcal{W}_k$-words in w by $\mathrm{EmpDist}_k(w)$. Formally:

$$\mathrm{EmpDist}_k(w)(w') = \frac{|\{0 \leq j \leq J : w_j = w'\}|}{J+1},\ w' \in \mathcal{W}_k.$$

Then *EmpDist* extends to a measure on $\mathcal{P}(\mathcal{W}_k)$ in the obvious way.

Emperical distributions and frequencies can be viewed as the same thing.

FINITIZING THE NOTION OF "GENERIC SEQUENCE"

A sequence $\langle v_n \in \mathcal{W}_n : n \in \mathbb{N}\rangle$ is a *generic sequence of words* if and only if for all k and $\epsilon > 0$ there is an N for all $m, n > N$,

$$\|EmpDist_k(v_m) - EmpDist_k(v_n)\|_{var} < \epsilon.$$

The sequence is generic for a measure μ if for all k:

$$\lim_{n\to\infty} \|\mathrm{EmpDist}_k(v_n) - \hat{\mu}_k\|_{var} = 0$$

where $\| \ \|_{var}$ is the total variation norm on probability distributions.

We can summarize the exact relationship between the empirical distributions and the μ_{q_k} by saying that the empirical distribution is the proportion of occurrences of $w' \in \mathcal{W}_k$ among the k-words that appear in v_n, whereas μ_{q_k} is approximately the density of the locations of the start of k-words in v_n. Letting $u \in \mathcal{W}_k$, d be the density of the positions where an occurrence of u begins in v_n, and d_s be the density of locations of letters in some spacer u_i we see that these are related by:

$$d \quad = \quad \left(\frac{\mathrm{EmpDist}(v_n)(u)}{q_k}\right)(1 - d_s)$$

TAKEAWAYS

- Generic sequence of words determine measures (uniquely)
- If $s \in \mathbb{K}$ is generic for an invariant measure μ, then the sequence of principal subwords of s form a generic sequence.

ERGODICITY

We want a criterion for when a generic sequence determines an *ergodic* measure.

A sequence $\langle v_n : n \in \mathbb{N} \rangle$ with $v_n \in \mathcal{W}_n$ is an *ergodic sequence* if for any k and $\epsilon > 0$ there are $n_0 > k$, and m_0 such that for all $m \geq m_0$, if

$$v_m = u_0 w_0 u_1 w_1 u_2 \ldots u_J w_J u_{J+1}$$

is the parsing of v_m into $\mathcal{W}_{n_0}$ words and spacers u_i then there is a subset $I \subseteq \{0, 1, 2 \ldots J\}$ with $|I|/J > 1 - \epsilon$ and for all $j, j' \in I$

$$\| EmpDist_k(w_j) - EmpDist_k(w_{j'}) \|_{var} < \epsilon. \quad (1)$$

- Ergodic sequences determine ergodic measures
- generic sequences for ergodic measures are ergodic sequences

IF YOU'VE GOTTEN LOST IN THE EPSILONS…

Both genericity and ergodicity can be expressed as counting properties of sequences of finite words.

LIFE IS COMPLICATED

Not all measure preserving symbolic systems come from construction sequences.

LIFE IS COMPLICATED

Not all measure preserving symbolic systems come from construction sequences.

In particular, even if $(\mathbb{K}, \mu)$ and $(\mathbb{L}, \nu)$ come from construction sequences, a joining between them may shift their relative locations (e.g. if it is not synchronous).

LIFE IS COMPLICATED

Not all measure preserving symbolic systems come from construction sequences.

In particular, even if $(\mathbb{K}, \mu)$ and $(\mathbb{L}, \nu)$ come from construction sequences, a joining between them may shift their relative locations (e.g. if it is not synchronous).

In this case we count occurrences of small length words in shifts of relatively large length words.

LIFE IS COMPLICATED

Not all measure preserving symbolic systems come from construction sequences.

In particular, even if $(\mathbb{K}, \mu)$ and $(\mathbb{L}, \nu)$ come from construction sequences, a joining between them may shift their relative locations (e.g. if it is not synchronous).

In this case we count occurrences of small length words in shifts of relatively large length words.

This works, but gets even more technical

CHEATING (A LOT)

- We have a characterization of *relatively independent joinings* in terms of generic sequences

- Compositions of joinings are defined in terms of *relatively independent joinings*, and hence we can characterize the composition of two joinings $\mathcal{J}_1$ and $\mathcal{J}_2$ by finitary criterion.

STATIONARY CODES

Suppose that Σ is a countable language. A *code* of length $2N+1$ is a function $\Lambda : \Sigma^{[-N,N]} \to \Sigma$, where $[-N,N]$ is the interval of integers starting at $-N$ and ending at N.

STATIONARY CODES

Suppose that Σ is a countable language. A *code* of length $2N + 1$ is a function $\Lambda : \Sigma^{[-N,N]} \to \Sigma$, where $[-N, N]$ is the interval of integers starting at $-N$ and ending at N.

Given a code Λ and an $s \in \Sigma^{\mathbb{Z}}$ we define the *stationary code* determined by Λ to be $\bar{\Lambda}(s)$ where:

$$\bar{\Lambda}(s)(k) = \Lambda(s \upharpoonright [k - N, k + N]).$$

STATIONARY CODES

Fix a shift invariant ergodic measure ν on a symbolic shift $\mathbb{K}$.

Suppose we have two codes Λ_0 and Λ_1 that are not necessarily of the same length. Let

- $D = \{s \in \Sigma^{\mathbb{Z}} : \overline{\Lambda}_0(s)(0) \neq \bar{\Lambda}_1(s)(0)\}$
- $d(\Lambda_0, \Lambda_1) = \nu(D)$

Then d is a semi-metric on codes.

STATIONARY CODES

Suppose that $\langle \Lambda_i : i \in \mathbb{N} \rangle$ is a sequence of codes such that $\sum_i d(\Lambda_i, \Lambda_{i+1}) < \infty$. Then there is a shift invariant Borel map $S : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ such that for ν-almost all s, $\lim_{i \to \infty} \overline{\Lambda}_i(s) = S(s)$

STATIONARY CODES

Suppose that $\langle \Lambda_i : i \in \mathbb{N} \rangle$ is a sequence of codes such that $\sum_i d(\Lambda_i, \Lambda_{i+1}) < \infty$. Then there is a shift invariant Borel map $S : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ such that for ν-almost all s, $\lim_{i\to\infty} \overline{\Lambda}_i(s) = S(s)$

Hence a convergent sequence of stationary codes determines a factor of $(\Sigma^{\mathbb{Z}}, \mathcal{B}, \nu, sh)$.

STATIONARY CODES

Conversely Suppose that $\pi : \mathbb{K} \to \mathbb{L}$ is a factor map. Then there is a sequence of finite codes $\langle \Lambda_i : i \in \mathbb{N} \rangle$ such that

$$\lim_{i \to \infty} \bar{\Lambda}_i = \pi$$

almost everywhere.

COMPUTING WITH STATIONARY CODES

Let Λ_0 and Λ_1 be codes. Define $\bar{d}(\bar{\Lambda}_0(s), \bar{\Lambda}_1(s))$ to be

$$\overline{\lim}_{N\to\infty} \frac{|\{k \in [-N, N] : \bar{\Lambda}_0(s)(k) \neq \bar{\Lambda}_1(s)(k)\}|}{2N+1}$$

COMPUTING WITH STATIONARY CODES

Let Λ_0 and Λ_1 be codes. Define $\bar{d}(\bar{\Lambda}_0(s), \bar{\Lambda}_1(s))$ to be

$$\overline{\lim}_{N\to\infty} \frac{|\{k \in [-N, N] : \bar{\Lambda}_0(s)(k) \neq \bar{\Lambda}_1(s)(k)\}|}{2N+1}$$

Lemma: Suppose that $(\Sigma^{\mathbb{Z}}, sh, \nu)$ is ergodic and that Λ_0 and Λ_1 be codes. Then for almost all $s \in S$:

$$d(\Lambda_0, \Lambda_1) = \bar{d}(\bar{\Lambda}_0(s), \bar{\Lambda}_1(s))$$

NB

Stationary codes only go from left to right. So to compare T and its inverse we need to look at the reversing map discussed on Monday.

WE ARE ALMOST READY FOR PRECISE DEFINITIONS THAT ALLOW US TO STATE PRECISELY THE THEOREM AND GIVE SOME IDEAS OF ITS PROOF

BACK TO OUR UGLY BABY

We want to understand circular words.

$$w = \prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (b^{q-j_i} w_j^{l-1} e^{j_i})$$

BACK TO OUR UGLY BABY

We want to understand circular words.

Whole Word:	$\prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (b^{q-j_i} w_j^{l-1} e^{j_i})$
2-subsection:	$\prod_{j=0}^{k-1} (b^{q-j_i} w_j^{l-1} e^{j_i})$
1-subsection:	$(b^{q-j_i} w_j^{l-1} e^{j_i})$
0-subsection:	w_j^{l-1}

BOUNDARY SECTIONS

Suppose that $w = \mathcal{C}(w_0, w_1, \ldots w_{k-1})$. Then w consists of blocks of w_i repeated $l-1$ times, together with some b's and e's that are not in the w_i's. The *interior* of w is the portion of w in the w_i's. The remainder of w consists of blocks of the form b^{q-j_i} and e^{j_i}. We call this portion the *boundary* of w.

BOUNDARY SECTIONS

Suppose that $w = \mathcal{C}(w_0, w_1, \ldots w_{k-1})$. Then w consists of blocks of w_i repeated $l-1$ times, together with some b's and e's that are not in the w_i's. The *interior* of w is the portion of w in the w_i's. The remainder of w consists of blocks of the form b^{q-j_i} and e^{j_i}. We call this portion the *boundary* of w.

If w is an n-word then we denote the boundary region as ∂_n.

BOUNDARY PORTIONS OF WORDS ARE SMALL

The proportion of a circular word w that belongs to its boundary is $1/l$. Moreover the proportion of the word that is within q letters of boundary of w is $3/l$.

BOUNDARY PORTIONS OF WORDS ARE SMALL

The proportion of a circular word w that belongs to its boundary is $1/l$. Moreover the proportion of the word that is within q letters of boundary of w is $3/l$.

Boundary regions are summable proportions of the words.

BOUNDARY PORTIONS OF WORDS ARE SMALL

The proportion of a circular word w that belongs to its boundary is $1/l$. Moreover the proportion of the word that is within q letters of boundary of w is $3/l$.

This is the reason that invariant measures on circular systems concentrate on S

COMPLEX CONJUGACY IS AN ISOMORPHISM BETWEEN A ROTATION AND ITS INVERSE.

WE NEED TO REPRESENT IT AS A SYMBOLIC MAP BETWEEN THE SYMBOLIC REPRESENTATION OF THE ROTATION AND ITS INVERSE.

WE NEED A SYMBOLIC REPRESENTATION OF COMPLEX CONJUGACY

$$w = \prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (b^{q-j_i} w_j^{l-1} e^{j_i})$$

so

$$\mathrm{rev}(w) = \prod_{i=1}^{q} \prod_{j=1}^{k} (e^{q-j_i} \, \mathrm{rev}(w_{k-j})^{l-1} b^{j_i}).$$

Since $q - j_i = j_{q-i} \pmod q$:

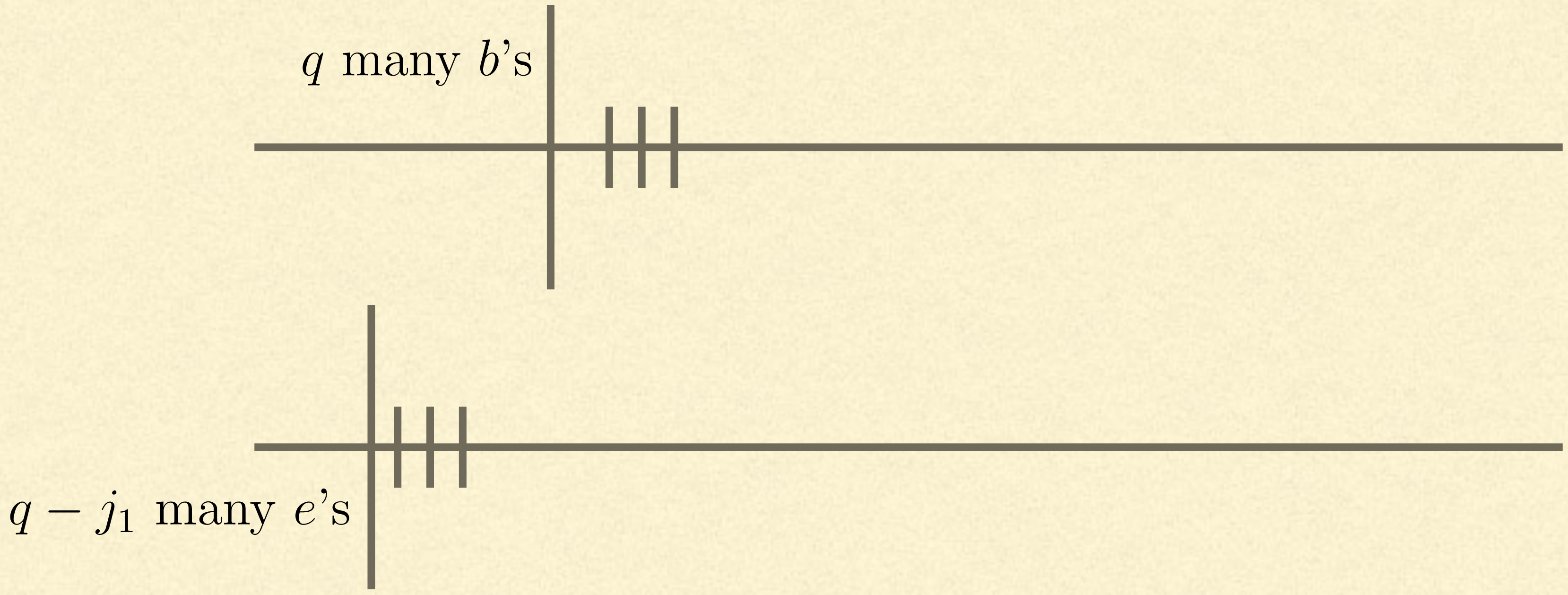
$$\mathrm{rev}(w) = \prod_{i=1}^{q} \prod_{j=1}^{k} (e^{q-j_i} \, \mathrm{rev}(w_{k-j})^{l-1} b^{j_i}).$$

We can reindex again and get another form:

$$\mathrm{rev}(w) = \prod_{i=0}^{q-1} \prod_{j=0}^{k-1} (e^{q-j_{i+1}} \, \mathrm{rev}(w_{k-j-1})^{l-1} b^{j_{i+1}}).$$

Let $w \in \mathcal{W}^c_{n+1}$ and view w as sitting at location $[0, q_{n+1}) \subseteq \mathbb{Z}$. Let $q = q_n$ and $k = k_n$. Consider $sh^{-j_1}(\mathrm{rev}(w))$ as being the word $\mathrm{rev}(w)$ in location $[j_1, q_{n+1} + j_1)) \subseteq \mathbb{Z}$. For all but at most $2kq$ of the occurrences of an n-subword w_j of w starting in a location $r \in [0, q_{n+1})$, the reversed word $\mathrm{rev}(w_{k-j-1})$ occurs in $sh^{-j_1}(\mathrm{rev}(w))$ starting at r.

Let $w \in \mathcal{W}^c_{n+1}$ and view w as sitting at location $[0, q_{n+1}) \subseteq \mathbb{Z}$. Let $q = q_n$ and $k = k_n$. Consider $sh^{-j_1}(\text{rev}(w))$ as being the word $\text{rev}(w)$ in location $[j_1, q_{n+1} + j_1)) \subseteq \mathbb{Z}$. For all but at most $2kq$ of the occurrences of an n-subword w_j of w starting in a location $r \in [0, q_{n+1})$, the reversed word $\text{rev}(w_{k-j-1})$ occurs in $sh^{-j_1}(\text{rev}(w))$ starting at r.



This is where it is convenient that both K and it inverse move in the same direction.

THE NATURAL MAP

We let $A_0 = 0$ and

$$A_{n+1} = A_n - (p_n)^{-1}.$$

THE NATURAL MAP

We let $A_0 = 0$ and

$$A_{n+1} = A_n - (p_n)^{-1}.$$

Define a stationary code $\bar{\Lambda}_N : \mathcal{K} \to \Sigma \cup \{b, e\}$ that approximates elements of $\mathrm{rev}(\mathbb{K})$ by:

$$\Lambda_n(s) = \begin{cases} sh^{A_n + 2r_n(s) - (q_n - 1)}(\mathrm{rev}(s))(0) & \text{if } r_n(s) \text{ is defined} \\ b & \text{otherwise} \end{cases}$$

THEOREM

Suppose that $\langle k_n, l_n : n \in \mathbb{N}\rangle$ is a circular coefficient sequence. Then the sequence of stationary codes $\langle \overline{\Lambda}_n : n \in \mathbb{N}\rangle$ converges to a shift invariant function $\overline{\natural} : \mathcal{K} \to (\{*\} \cup \{b, e\})^{\mathbb{Z}}$ that induces an isomorphism $\natural$ from $\mathcal{K}$ to $\mathrm{rev}(\mathcal{K})$.

OBSERVATIONS

Let $\phi : \mathcal{K} \to [0,1]$ be the isomorphism between $(\mathbb{K}, sh)$ and the rotation R_α. Then:

- $\phi\natural\phi^{-1}$ is the map $z \mapsto -z + \beta$ for some β.
- $\natural$ is an involution.

What is β?

FINALLY PRECISE STATEMENTS

If $\mathbb{K}$ and $\mathbb{L}$ are either odometer based or circular systems let $\mathbb{K}^\pi$ and $\mathbb{L}^\pi$ be the corresponding odometer or rotation systems on which they are based.

SYNCHRONOUS JOININGS: ODOMETERS

Let $\mathbb{K}$ and $\mathbb{L}$ be odometer based systems with the same coefficient sequence, and ρ a joining between $\mathbb{K}$ and $\mathbb{L}^{\pm 1}$. Then ρ is *synchronous* if ρ joins $\mathbb{K}$ and $\mathbb{L}$ and the projection of ρ to a joining on $\mathbb{K}^\pi \times \mathbb{L}^\pi$ is the graph joining determined by the identity map (the diagonal joining of the odometer factors); ρ is *anti-synchronous* if ρ is a joining of $\mathbb{K}$ with $\mathbb{L}^{-1}$ and its projection to $\mathbb{K}^\pi \times (\mathbb{L}^{-1})^\pi$ is the graph joining determined by the map $x \mapsto -x$.

SYNCHRONOUS JOININGS: CIRCULAR SYSTEMS

Let $\mathbb{K}^c$ and $\mathbb{L}^c$ be circular systems with the same coefficient sequence and ρ a joining between $\mathbb{K}^c$ and $(\mathbb{L}^c)^{\pm 1}$. Then ρ is *synchronous* if ρ joins $\mathbb{K}^c$ and $\mathbb{L}^c$ and the projection to a joining of $(\mathbb{K}^c)^\pi$ with $(\mathbb{L}^c)^\pi$ is the graph joining determined by the identity map of $\mathcal{K}$ with $\mathcal{L}$, the underlying rotations; ρ is *anti-synchronous* if it is a joining of $\mathbb{K}^c$ with $(\mathbb{L}^c)^{-1}$ and projects to the graph joining determined by $\mathrm{rev}() \circ \natural$ on $\mathcal{K} \times \mathcal{L}^{-1}$.

THE CATEGORIES

Let $\mathcal{O}B$ be the category whose objects are ergodic odometer based systems with coefficients $\langle k_n : n \in \mathbb{N}\rangle$. The morphisms between objects $(\mathbb{K}, \mu)$ and $(\mathbb{L}, \nu)$ will be synchronous graph joinings of $(\mathbb{K}, \mu)$ and $(\mathbb{L}, \nu)$ or anti-synchronous graph joinings of $(\mathbb{K}, \mu)$ and $(\mathbb{L}^{-1}, \nu)$. We call this the *category of odometer based systems.*

THE CATEGORIES

Let $\mathcal{C}B$ be the category whose objects consists of all ergodic circular systems with coefficients $\langle k_n, l_n : n \in \mathbb{N}\rangle$. The morphisms between objects $(\mathbb{K}^c, \mu^c)$ and $(\mathbb{L}^c, \nu^c)$ will be synchronous graph joinings of $(\mathbb{K}^c, \mu^c)$ and $(\mathbb{L}^c, \nu^c)$ or anti-synchronous graph joinings of $(\mathbb{K}^c, \mu^c)$ and $((\mathbb{L}^c)^{-1}, \nu^c)$. We call this the *category of circular systems.*

THE MAIN THEOREM

For a fixed circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$ the categories $\mathcal{O}B$ and $\mathcal{C}B$ are isomorphic by a function $\mathcal{F}$ that takes synchronous joinings to synchronous joinings, anti-synchronous joinings to anti-synchronous joinings, isomorphisms to isomorphisms and weakly mixing extensions to weakly mixing extensions. Moreover for an odometer system $\mathbb{K}$ the simplex of invariant measures is affinely homeomorphic to the simplex of invariant measures for $\mathcal{F}(\mathbb{K})$.

TO PROVE THIS

We need to:

- Define $\mathcal{F}$ on objects
- Define $\mathcal{F}$ on morphisms
- Prove that $\mathcal{F}$ preserves composition
- Prove that $\mathcal{F}$ is a bijection

NOTATION

- q_n is defined by the AK-formulas. It is the length of the circular n-words.
- K_n is defined as $\prod_{m<n} k_m$. This is the length of odometer n-words.

NOTATION

Notational convention:
$\mathbb{K}$ vs $\mathbb{K}^c$ and ν vs ν^c

DEFINING $\mathcal{F}$ ON OBJECTS

Fix a circular coefficient sequence $\langle k_n, l_n : n \in \mathbb{N} \rangle$.

Suppose we are given an odometer based construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ with coefficients $\langle k_n : n \in \mathbb{N} \rangle$.

We inductively build a circular construction sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$ by induction.

Define a construction sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N}\rangle$ and bijections $c_n : \mathcal{W}_n \to \mathcal{W}_n^c$ by induction as follows:

1. Let $\mathcal{W}_0^c = \Sigma$ and c_0 be the identity map.
2. Suppose that $\mathcal{W}_n, \mathcal{W}_n^c$ and c_n have already been defined.

   $$\mathcal{W}_{n+1}^c = \{\mathcal{C}_n(c_n(w_0), c_n(w_1), \dots c_n(w_{k_n-1})) : w_0 w_1 \dots w_{k_n-1} \in \mathcal{W}_{n+1}\}.$$

   Define the map c_{n+1} by setting

   $$c_{n+1}(w_0 w_1 \dots w_{k_n-1}) = \mathcal{C}_n(c_n(w_0), c_n(w_1), \dots c_n(w_{k_n-1})).$$

DEFINING $\mathcal{F}$ ON OBJECTS

Define a map $\mathcal{F}$ from the set of odometer based systems (viewed as subshifts) to circular systems (viewed as subshifts) as follows. Suppose that $\mathbb{K}$ is built from a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$. Define

$$\mathcal{F}(\mathbb{K}) = \mathbb{K}^c$$

where $\mathbb{K}^c$ has construction sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$.

DEFINING $\mathcal{F}$ ON OBJECTS

Define a map $\mathcal{F}$ from the set of odometer based systems (viewed as subshifts) to circular systems (viewed as subshifts) as follows. Suppose that $\mathbb{K}$ is built from a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$. Define

$$\mathcal{F}(\mathbb{K}) = \mathbb{K}^c$$

where $\mathbb{K}^c$ has construction sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$.

Since the c_n's are invertible $\mathcal{F}$ is a bijection.

DEFINING $\mathcal{F}$ ON OBJECTS

Define a map $\mathcal{F}$ from the set of odometer based systems (viewed as subshifts) to circular systems (viewed as subshifts) as follows. Suppose that $\mathbb{K}$ is built from a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$. Define

$$\mathcal{F}(\mathbb{K}) = \mathbb{K}^c$$

where $\mathbb{K}^c$ has construction sequence $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$.

But how do we define $\mathcal{F}$ on measures?

GENETIC MARKERS

Suppose that u, v are words in $\mathcal{W}_n$ and $\mathcal{W}_{n+1}$ respectively and u occurs as an n-subword of v in a particular location. Viewing v as a concatenation $w_0 w_1 \ldots w_{n_k - 1}$ of n-subwords, there is a j such that $u = w_j$. Let $j_n^* = j$ and call j_n^* the *genetic marker* of u in v.

GENETIC MARKERS

Suppose that $u \in \mathcal{W}_n$, $v \in \mathcal{W}_{n+k}$ and u is an n-subword of v occurring at a particular location. Then there is a sequence of words $u_n = u$, u_{n+1}, $\dots u_{n+k-1}, u_{n+k} = v$ such that u_i is an $n+i$-subword of v at a definite location and the location of u in v is inside u_i. Let j^*_{n+i} be the genetic marker of u_{n+i} inside u_{n+i+1}. We call the sequence $\vec{j}^* = \langle j^*_n, j^*_{n+1}, \dots j^*_{n+k-1}\rangle$ the *genetic marker* of u in v. If $\vec{j}^*$ is the genetic marker of some n-word inside an m-word, we will call it an (n, m)-genetic marker.

GENETIC MARKERS

The point is that if u occurs as a subword of v then the genetic marker $\langle j^*_n, j^*_{n+1} \ldots j^*_{n+k-1}\rangle$ of that occurrence codes its location inside v.

GENETIC MARKERS INSIDE ODOMETER WORDS

Suppose that $s \in \mathbb{K}$ has principal n-blocks $\langle w_n : n \in \mathbb{N} \rangle$. Each w_{n+1} is a concatenation of words $v_0 v_1 \ldots v_{k_n - 1}$. Let

$$j'_n =_{def} \frac{r_{n+1}(s) - r_n(s)}{K_n} \tag{1}$$

or equivalently

$$r_{n+1}(s) = r_n(s) + j'_n K_n. \tag{2}$$

Each w_{n+1} is a concatenation of words $v_0 v_1 \ldots v_{k_n - 1}$, and we see that $s(0)$ belongs to $v_{j'_n}$. In particular, the genetic marker of w_n inside w_{n+k} is the sequence $\langle j'_n, j'_{n+1}, \ldots j'_{n+k-1} \rangle$.

INFINITE GENETIC MARKERS

Given a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ for an odometer based or circular system $\mathbb{K}$, $s \in S$ and an occurrence of an n-word u in s. Then we can inductively define an infinite sequence of words $\langle u_m : n \leq m \in \mathbb{N} \rangle$, letting $u_n = u$, and u_{m+1} to be the $m+1$-subword of s that contains u_m. For each $n < m$ we get a genetic marker $\langle j_n^*, j_{n+1}^*, \dots j_{m-1}^* \rangle$, and these cohere as m goes to infinity. We define the *infinite genetic marker* to be $\vec{j}^* = \langle j_m^* : n \leq m \in \mathbb{N} \rangle$.

INFINITE GENETIC MARKERS

Given a construction sequence $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ for an odometer based or circular system $\mathbb{K}$, $s \in S$ and an occurrence of an n-word u in s. Then we can inductively define an infinite sequence of words $\langle u_m : n \leq m \in \mathbb{N} \rangle$, letting $u_n = u$, and u_{m+1} to be the $m+1$-subword of s that contains u_m. For each $n < m$ we get a genetic marker $\langle j^*_n, j^*_{n+1}, \dots j^*_{m-1} \rangle$, and these cohere as m goes to infinity. We define the *infinite genetic marker* to be $\vec{j}^* = \langle j^*_m : n \leq m \in \mathbb{N} \rangle$.

Given two words occurring in s, their genetic markers agree on a tail.

RECONSTRUCTING WITH GENETIC MARKERS

If we are given a sequence of words $\langle u_m : n \leq m \rangle$, with $u_m \in \mathcal{W}_m$, and an infinite sequence $\langle j_m : n \leq m \rangle$ such that the genetic marker j_m denotes an instance of u_m in u_{m+1} then we can find an $s \in \mathbb{K}$ with $\langle u_m : m \geq n \rangle$ as a tail of its principal subwords. If $\mathbb{K}$ is odometer then s is unique up to a shift of size less than or equal to K_m

GENETIC MARKERS FOR CIRCULAR WORDS ARE MORE COMPLICATED

GENETIC MARKERS FOR CIRCULAR WORDS ARE MORE COMPLICATED

They code *regions* of words.

GENETIC MARKERS FOR CIRCULAR WORDS

Given u and v as above, we can consider the construction of $c_{n+k}(v)$ starting with the collection $\{c_n(u) : u \text{ is an } n\text{-subword of } v\}$. Each of the genetic markers $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$ of a subword u of v determines a *region* of n-subwords of $c_{n+k}(v)$.

GENETIC MARKERS FOR CIRCULAR WORDS

Explicitly, in the first step of the construction we put u into the $(j_n^*)^{th}$ argument of $\mathcal{C}_n$. At the next step we put the result into the j_{n+1}^* argument of $\mathcal{C}_{n+1}$ and so on. Thus we see that there are bijections between

1. sequences $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$ with $0 \leq j_m^* < k_m$,
2. n-subwords u of v,
3. the regions of v^c occupied by the occurrences of powers $(u^c)^{l_n - 1}$ where u^c is the element of $\mathcal{W}_n^c$ determined by $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^* \rangle$.

GENETIC MARKERS IN CIRCULAR SYSTEMS

Genetic markers can be used to reconstruct elements s in circular systems, but the ambiguity of referring to *regions* rather than specific words has to be accounted for.

GENETIC MARKERS LET US COMPUTE THE RELATIVE FREQUENCY OF CORRESPONDING REGIONS OF ODOMETER AND CIRCULAR WORDS

Suppose that u^c occurs in v^c with genetic marker $\langle j^*_n, j^*_{n+1}, \ldots j^*_{n+k-1}\rangle$. Then the number of occurrences of u^c in v^c with the same genetic marker $\langle j^*_n, j^*_{n+1}, \ldots j^*_{n+k-1}\rangle$ is

$$\prod_{n}^{n+k-1} q_i(l_i - 1).$$

COMPUTING EMPIRICAL DISTRIBUTIONS

Since particular (n, m)-genetic markers $\langle j^*_n, j^*_{n+1}, \dots j^*_{n+k-1}\rangle$ correspond to powers of u^c's that occur with the same multiplicity in v^c, independently of the marker, we see that for a given u and v:

$$\frac{|\{\text{occurrences of } u \text{ in } v\}|}{|\{n\text{-subwords of } v\}|} = \frac{|\{\text{occurrences of } c_n(u) \text{ in } c_{n+k}(v)\}|}{|\{\text{circular } n\text{-subwords of } c_{n+k}(v)\}|} \tag{1}$$

COMPUTING EMPIRICAL DISTRIBUTIONS

Since particular (n, m)-genetic markers $\langle j_n^*, j_{n+1}^*, \dots j_{n+k-1}^* \rangle$ correspond to powers of u^c's that occur with the same multiplicity in v^c, independently of the marker, we see that for a given u and v:

$$\frac{|\{\text{occurrences of } u \text{ in } v\}|}{|\{n\text{-subwords of } v\}|} = \frac{|\{\text{occurrences of } c_n(u) \text{ in } c_{n+k}(v)\}|}{|\{\text{circular } n\text{-subwords of } c_{n+k}(v)\}|} \tag{1}$$

In other words:

$$\mathrm{EmpDist}(v)(u) = \mathrm{EmpDist}(c_{n+k}(v))(c_n(u)).$$

COMPUTING DENSITIES

For sets $A \subseteq [0, K_m)$ and $A^c \subseteq [0, q_m)$ we denote their densities by:

$$\begin{aligned} d_m(A) &= |A|/K_m \\ d_m^c(A^c) &= |A^c|/q_m \end{aligned}$$

Then d_m and d_m^c can be viewed as discrete probability measures on the sets $[0, K_m)$ and $[0, q_m)$ respectively.

COMPUTING DENSITIES

Let S^* be a collection of (n,m)-genetic markers, g the total number of (n,m)-genetic markers and $d = |S^*|/g$. Let:

- $A = \{k \in [0, K_m) :$ some $u \in \mathcal{W}_n$ with genetic marker in S^* begins at k in $w\}$
- $A^c = \{k \in [0, q_m) :$ some $u^c \in \mathcal{W}_n^c$ with genetic marker in S^* begins at k in $w^c\}$,

COMPUTING DENSITIES

Then:

$$\begin{aligned}
d_m(A) &= \frac{d}{K_n} \\
d_m^c(A^c) &= \frac{d}{q_n}\prod_{p=n}^{m-1}(1-1/l_p) \\
d_m(A) &= \left(\frac{d_m^c(A^c)}{\prod_{p=n}^{m-1}(1-1/l_p)}\right)\left(\frac{q_n}{K_n}\right) \\
d_m^c(A^c) &= d_m(A)\left(\prod_{p=n}^{m-1}(1-1/l_p)\right)\left(\frac{K_n}{q_n}\right).
\end{aligned}$$

The analogous equations between densities and invariant measures are:

$$d^c_m(A^c) = d_m(A)\left(\prod_{p=n}^{m-1}(1-1/l_p)\right)\left(\frac{K_n}{q_n}\right)$$

and

$$\nu^c(\langle c_n(u)\rangle) = \left(\frac{K_n}{q_n}\right)\nu(\langle u\rangle)\left(1-\sum_n^{\infty} d^{\partial_n}\right)$$

OUR TASK

We have our map $\mathbb{K} \mapsto \mathcal{F}(\mathbb{K})$.
But we need to define $\mathcal{F}(\mathbb{K}, \nu)$?

OUR TASK

We have our map $\mathbb{K} \mapsto \mathcal{F}(\mathbb{K})$.
But we need to define $\mathcal{F}(\mathbb{K}, \nu)$?

Notational convention:
$\mathbb{K}$ vs $\mathbb{K}^c$ and ν vs ν^c

TWO MAPS TU AND UT

- We define maps $TU : S \to S^c$ and $UT : S^c \to S$.
- The map TU will be one-to-one but UT will not
- $UT \circ TU$ will be the identity map

DEFINING TU

Let $s \in S$. Let u_n be the principal n-subword of s. The sequence $\langle u_n : n \in \mathbb{N}\rangle$ determines a sequence of circular words $\langle u_n^c : n \in \mathbb{N}\rangle$ which we assemble to define $TU(s)$. Let $\vec{j} = \langle j_n : n \in \mathbb{N}\rangle$ be the infinite genetic marker of $s(0)$. To describe $TU(s)$ completely we need to define $\langle r_n^c : n \in \mathbb{N}\rangle$. Set $r_0^c = 0$, and inductively define r_{n+1}^c to be the $(r_n^c)^{th}$ position in the first occurrence of an n-word with genetic marker j_n in u_{n+1}^c. Set $TU(s)$ to be the element of $\mathbb{K}^c$ with principal subwords $\langle u_n^c : n \in \mathbb{N}\rangle$ and location sequence $\langle r_n^c : n \in \mathbb{N}\rangle$.

DEFINING UT

UT associates an element of $\mathbb{K}$ to each element of S^c. Given such an $s^c \in S^c$, let $\langle u_n^c : n \geq N \rangle$ be its sequence of principal n-subwords. For each $n \geq N$, u_n^c occurs as $u_{j_n^*}$ in the preword corresponding to u_{n+1}^c. Let $u_n = c_n^{-1}(u_n^c)$. Then the sequence of words $\langle u_n : n \in \mathbb{N} \rangle$ and genetic markers $\langle j_n^* : n \geq N \rangle$ determine an element of $s \in \mathbb{K}$ except for the location of 0 in the double ended sequence. (The sequence is double ended because $s \in S^c$.)

DEFINING UT

This location is determined arbitrarily subject to the requirement that the sequence of u_n's are the principal n-blocks of s ($n \geq N$) and the j_n^* the sequence of genetic markers of these n-blocks. Let $\bar{0}$ be a sequence of zeros of length N. Then $\bar{0}^\frown\langle j_n^* : n \geq N\rangle$ is a well-defined member of the odometer $\mathcal{O}$ associated with $\mathbb{K}$. The sequence $\bar{0}^\frown\langle j_n^* : n \geq N\rangle$ determines a sequence $\langle r_n : n \in \mathbb{N}\rangle$. Hence the pair $\langle u_n : n \geq N\rangle$ and $\bar{0}^\frown\langle j_n^* : n \geq N\rangle$ determines a unique element s of $\mathbb{K}$ which we will denote $UT(s^c)$. It is easy to check that $UT \circ TU = id$ and that for each $s \in S$, there is a perfect set of s^c with $UT(s^c) = s$.

FACTS ABOUT UT

Let $s \in S$. Then $\{TU(sh^k(s)) : k \in \mathbb{Z}\} \subseteq \{sh^k(TU(s)) : k \in \mathbb{Z}\}$. If $s \in S$, $s^c = TU(s)$ and $u \in \mathcal{W}_n$, then there is a canonical correspondence between occurrences of u in s and finite regions of s^c where u^c occurs. The occurrences of u^c in these finite regions have the same infinite genetic marker $\langle j_m : m > n \rangle$ in s^c as u does in s.

WHY ALL THIS WORK?

Let $\langle v_n : n \in \mathbb{N} \rangle$ be a sequence with $v_n \in \mathcal{W}_n$. Let $v_n^c = c_n(v_n)$. Then:

1. $\langle v_n : n \in \mathbb{N} \rangle$ is an ergodic sequence iff $\langle v_n^c : n \in \mathbb{N} \rangle$ is an ergodic sequence.

2. $\langle v_n : n \in \mathbb{N} \rangle$ is a generic sequence for a measure ν iff $\langle v_n^c : n \in \mathbb{N} \rangle$ is a generic sequence for a measure ν^c. In case either sequence is generic, the measures ν and ν^c satisfy

$$\nu^c(\langle c_n(u) \rangle) = \left(\frac{K_n}{q_n} \right) \nu(\langle u \rangle) \left(1 - \sum_n^{\infty} \nu^c(\partial_m) \right).$$

WHY?

Since particular (n,m)-genetic markers $\langle j_n^*, j_{n+1}^*, \ldots j_{n+k-1}^*\rangle$ correspond to powers of u^c's that occur with the same multiplicity in v^c, independently of the marker, we see that for a given u and v:

$$\frac{|\{\text{occurrences of } u \text{ in } v\}|}{|\{n\text{-subwords of } v\}|} = \frac{|\{\text{occurrences of } c_n(u) \text{ in } c_{n+k}(v)\}|}{|\{\text{circular } n\text{-subwords of } c_{n+k}(v)\}|}$$

The analogous equations between densities and invariant measures are:

$$d_m^c(A^c) = d_m(A)\left(\prod_{p=n}^{m-1}(1-1/l_p)\right)\left(\frac{K_n}{q_n}\right)$$

and

$$\nu^c(\langle c_n(u)\rangle) = \left(\frac{K_n}{q_n}\right)\nu(\langle u\rangle)\left(1-\sum_n^\infty d^{\partial_n}\right)$$

TRANSFERRING MEASURES

Let $\langle \mathcal{W}_n : n \in \mathbb{N} \rangle$ be a uniquely readable construction sequence for an odometer based system $\mathbb{K}$ and $\langle \mathcal{W}_n^c : n \in \mathbb{N} \rangle$ be the associated circular construction sequence for $\mathbb{K}^c$. Then there is a canonical affine homeomorphism $\nu \mapsto \nu^c$ between shift invariant measures ν concentrating on $\mathbb{K}$ and non-atomic, shift invariant measures ν^c such that

$$\nu^c(\langle c_n(u) \rangle) = \left(\frac{K_n}{q_n} \right) \nu(\langle u \rangle) \left(1 - \sum_n^{\infty} \nu^c(\partial_m) \right)$$

holds between ν and ν^c.

LOTS OF CALCULATIONS GO HERE…

BIG PICTURE

If ν and ν^c are corresponding measures on $\mathbb{K}$ and $\mathbb{K}^c$ and $s \in \mathbb{K}$ is arbitrary then s is generic for ν iff $TU(s)$ is generic for ν^c. The point s is generic just in case its sequence of principal subwords is generic for ν. This holds just in case the sequence of principal subwords of $TU(s)$ is generic; i.e. $TU(s)$ is generic.

DOWNAROWICZ THEOREM

Every non-empty compact metrizable Choquet simplex is affinely homeomorphic to the simplex of invariant probability measures for a dyadic Toeplitz flow

DOWNAROWICZ THEOREM

Every non-empty compact metrizable Choquet simplex is affinely homeomorphic to the simplex of invariant probability measures for a dyadic Toeplitz flow

1. Toeplitz systems are special cases of odometer based systems.
2. We can apply our functor to the symbolic shift associated with the Toeplitz system.
3. Our functor preserves the affine structure of the space of measures.

COROLLARY

Every non-empty compact metrizable Choquet simplex is affinely homeomorphic to the simplex of invariant probability measures for a circular system.

WHAT'S LEFT?

This is only about half of the paper.

The rest of the paper is devoted to dealing with

- extending $\mathcal{F}$ to morphisms, synchronous and anti-synchronous
- showing $\mathcal{F}$ preserves composition
- showing $\mathcal{F}$ preserves weakly mixing/compact extension
- showing $\mathcal{F}$ is continuous

JOININGS

A synchronous joining $\mathcal{J}$ between $\mathbb{K}$ and $\mathbb{L}$ (based on the same coefficient sequence $\langle k_n : n \in \mathbb{N}\rangle$) can be itself viewed as a measure on

$$\{(s,t) : s \in \mathbb{K}, t \in \mathbb{L} \text{ and } \pi(s) = \pi(t)\}.$$

This latter can itself be viewed as an odometer based system.

We know how to lift such measures via $\mathcal{F}$.

This gives the map $\mathcal{F}$ on synchronous morphisms.

JOININGS

More explicitly, we need to lift joinings $\mathcal{J}$ of $\mathbb{K}$ with $\mathbb{L}$.

- we choose a pair $(s,t) \in \mathbb{K} \times \mathbb{L}$ that is generic for $\mathcal{J}$.
- Let $\langle (u_n, v_n) : n \in \mathbb{N} \rangle$ be the sequence of principal subwords of (s,t)
- Argue the map sending $\langle (u_n, v_n) : n \in \mathbb{N} \rangle$ to $\langle (u_n^c, v_n^c) : n \in \mathbb{N} \rangle$ sends ergodic sequences generic for $\mathcal{J}$ to ergodic sequences generic for some joining $\mathcal{J}^c$ of $\mathbb{K}^c$ with $\mathbb{L}^c$

ANTI-SYNCHRONOUS JOININGS

We have to understand the relationship

1. $\mathbf{P}^{-}$, the collection of anti-synchronous joinings ρ of $\mathbb{K}$ and $\mathbb{L}^{-1}$.

2. $\mathbf{P}^{\natural}$, the collection of anti-synchronous joinings ρ^c of $\mathbb{K}^c$ and $(\mathbb{L}^c)^{-1}$.

THIS INVOLVES A COMPLICATED ANALYSIS OF HOW WORDS OCCUR IN THE NATURAL MAP.

RECALL

We let $A_0 = 0$ and

$$A_{n+1} = A_n - (p_n)^{-1}.$$

We need to lift joinings $\mathcal{J}$ of $\mathbb{K}$ with $\mathbb{L}^{-1}$. (Remember we are identifying $\mathbb{L}^{-1}$ with $\mathrm{rev}(\mathbb{L})$.)

- we choose a pair $(s, t) \in \mathbb{K} \times \mathrm{rev}(\mathbb{L})$ that is generic for $\mathcal{J}$.
- Let $\langle (u_n, \mathrm{rev}(v_n)) : n \in \mathbb{N}\rangle$ be the sequence of principal subwords of (s, t)
- Define $(u_n, v_n)^c$ to be $(u_n^c, sh^{A_n}(\mathrm{rev}(v_n^c))$.
- Argue the map sending $\langle (u_n, \mathrm{rev}(v_n)) : n \in \mathbb{N}\rangle$ to $\langle (u_n, \mathrm{rev}(v_n))^c : n \in \mathbb{N}\rangle$ sends ergodic sequences generic for $\mathcal{J}$ to ergodic sequences generic for some joining $\mathcal{J}^c$ of $\mathbb{K}^c$ with $\mathrm{rev}(\mathbb{L}^{-1})$

WHAT'S LEFT

- Graph joinings get sent to graph joinings–so factor maps get sent to factor maps
- Compositions are preserved
- weakly mixing and compact extensions

WHAT'S LEFT

- Graph joinings get sent to graph joinings–so factor maps get sent to factor maps
- Compositions are preserved
- weakly mixing and compact extensions

Part 2 is complicated by the fact that genetic markers refer to regions in circular systems

For the first item (very roughly)

1. If $\mathcal{J}$ is a graph joining of $\mathbb{K}$ with $\mathbb{L}$, then we can well approximate (in the sense of $\mathcal{J}$-measure) of any open set $\mathbb{L} \times \langle v \rangle$ by a finite union of open sets $\bigcup(\langle u_i \rangle \times \mathbb{K})$.

2. We lift this approximation to an approximation of $\mathbb{L}^c \times \langle v^c \rangle$ by a finite union of open sets of form $\langle u_i^c \rangle_l \times \mathbb{K}^c$.

WHAT'S LEFT

- Compositions are preserved
- weakly mixing and compact extensions

- Compositions are preserved
- weakly mixing and compact extensions

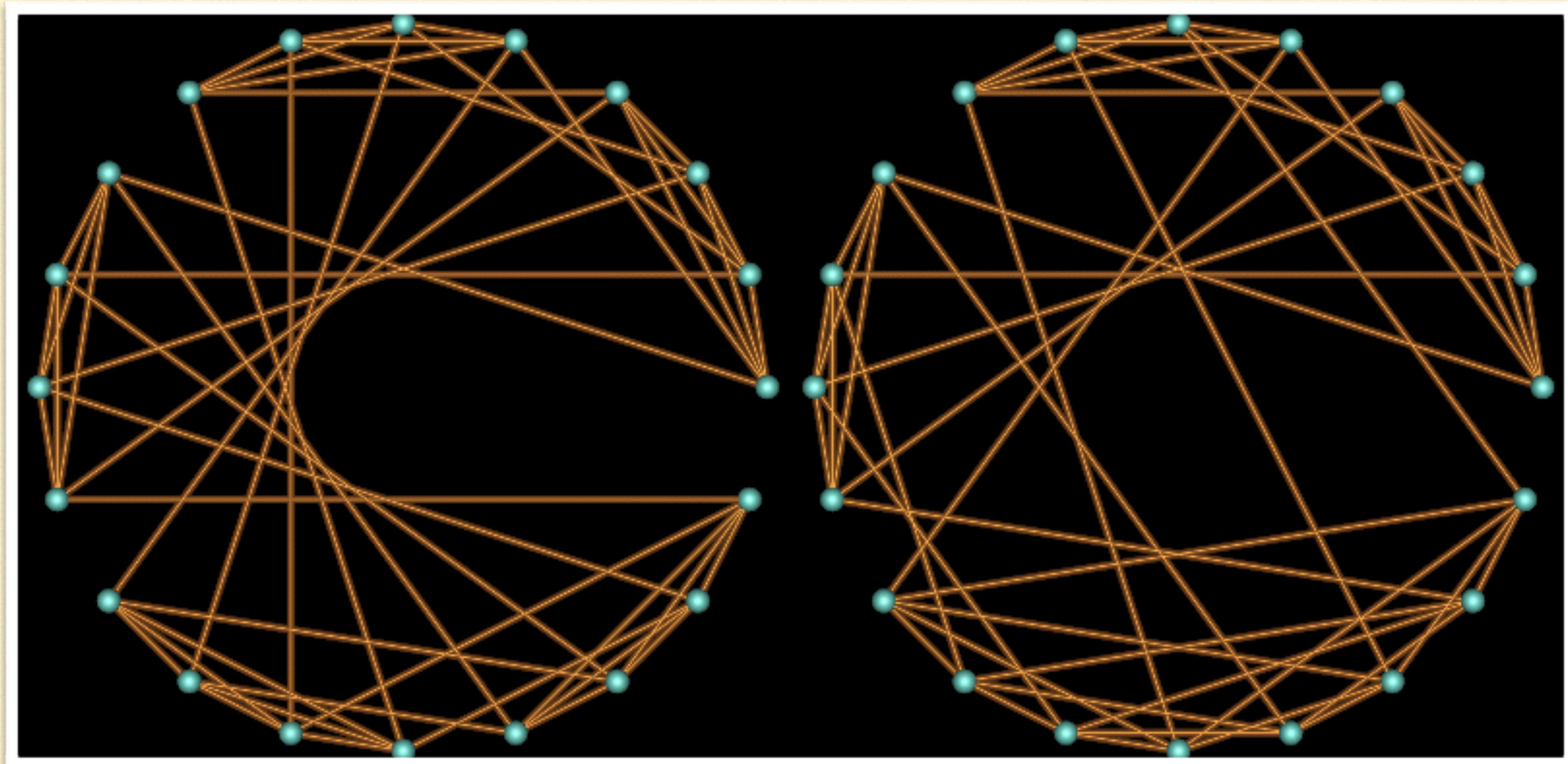
1. Both of these involve analyzing relatively independent joinings in terms of generic sequences of words.
2. This involves computing conditional frequencies of words
3. It gets complicated.

A COUPLE OF OPEN PROBLEMS

1. Characterize circular systems explicitly in ergodic-theoretic terms (in analogy to the spectral characterization of odometer based systems).

2. How much do the categories depend on the sequence $\langle k_n, l_n : n \in \mathbb{N}\rangle$ rather than the limiting irrational $\alpha = \lim \alpha_n$?

3. Do we NEED Liouvillean irrationals for this part of the argument to work? (Hermann's theorem isn't relevant.)



END OF GLOBAL STRUCTURE LECTURE


It's time for the intermission!
SODA POP
HOT
Let's grab ourselves a snack!