

Flat stacks, Joining-Closure and Genericity

Jonathan L.F. King
 University of Florida, Gainesville FL 32611-2082, USA
 squash@math.ufl.edu
 Webpage <http://www.math.ufl.edu/~squash/>

29 July, 2001

ABSTRACT: [Preliminary version.] In the usual Halmos topology on the group of transformations (\mathbb{Z} -actions), we show that a joinings version of the Weak-Closure Theorem holds for the “generic T ” (that is, for a residual set of T). Specifically, the closure of the set of off-diagonal self-joinings of T is the *full* simplex of self-joinings. This also holds for infinite-dimensional self-joinings.

An ingredient in the proof is showing that if T has “flat stacks”, then the off-diagonal self-joinings are dense in the set of *ergodic* self-joinings. A second ingredient is a combinatorial lemma, fancifully called the “Chameleon lemma”, involving words and blockings.

§A Introduction

In the past two decades it has become evident that properties of the self-joinings (henceforth just called “joinings”) of a transformation provide important tools-of-classification of such maps. My purpose here, *Strong Joining-Closure*, theorem 4, is to show that a certain “joinings extension” of the *Weak-Closure Theorem* holds generically.

Section A and §B transmogrify the problem to a discrete combinatorial statement, the *Chameleon Lemma* of §C, which is dispatched in §D. Here is a bird’s eye view.

§Outline

A Introduction	A1
Weak-Closure Theorem	A2
EJCI Theorem	A2
SJCI Theorem	A2
Words & Blockings	A2
Rank-1 and Flat Stacks	A3
Criteria for rank-1 and flat stacks	A3
Rotation FS Theorem	A4

AMS Subject Classification: Primary: 28D05. Secondary: 54E52. *Key words and phrases:* Joinings, Baire category, Flat stacks, Weak closure theorem.

Joinings	A5
Graph and Diagonal Joinings	A5
Self-joinings of a transformation	A5
Ergodic joinings	A5
Topologies on $\Omega(\mu)$ and $\mathcal{J}(\mu)$	B6
Standard topology on $\mathcal{J}(\mu)$	B6
Finite-order suffices	B6
B Tools	B6
Typicality Lemma	B7
Establishing the <i>EJCI Theorem</i>	B7
Shift-words	B7
FS Uniformity Lemma	B8
Reduction of SJCI to 2-mixtures	B8
C Strong Joining-Closure is Ω-residual	C9
Prolegomenon	C9
Frequencies estimate a diagonal joining	C10
Conditional proof of <i>SJCI Theorem</i>	C10
Word Combinatorics	C10
Chameleon Lemma	D11
Frequency Mixture Corollary	D11
D Reformulating the Chameleon Lemma	D11
Blockings revisited	D11
Joint-word Prop’n	D12
Swimming Pools	D12
Gym-shift Prop’n	D12
Same-Shift Prop’n	D12
Nested Gym Construction	D13
Blocked-Chameleon Lemma	D13
Composition Prop’n	D13
Proof of <i>Blocked-Chameleon</i> for $\mathfrak{D}=4$	D13
E End-Notes	E14

A *measure* will mean “Lebesgue probability measure”. We will work on a non-atomic measure space (X, \mathcal{X}, μ) with a (measure-preserving) map $T: X \rightarrow X$ which is invertible*. Let $\Omega = \Omega(\mu)$ denote the group, under composition, of such transformations. We have subgroups $\Omega \supset \mathbf{C}(T) \supset \mathbb{Z}(T)$. Here $\mathbf{C}(T)$, the *commutant* of T , comprises those transformations S such that $ST = TS$, and $\mathbb{Z}(T)$ means the set $\{T^e\}_{e \in \mathbb{Z}}$ of powers of T .

Idiosyncrasies. Expressions “ $d =: s$ ” and “ $s := d$ ” each mean that d is the definition of symbol s .

For real numbers, let “ b exceeds c ” mean $b > c$, whereas the weaker $b \geq c$ condition is “ b dominates c ”. Say that b is δ -*close* to c , written $b \overset{\delta}{\approx} c$,

*It turns out that the maps which commute with a rank-1 must necessarily be invertible, [6]. For this reason, there will be no loss of generality in considering only invertible transformations.

if $|b - c| < \delta$. For vectors \mathbf{b} and \mathbf{c} over a common index set, say \mathbf{b} is δ -close to \mathbf{c} if each $b_i \overset{\delta}{\approx} c_i$. For equal length words \mathbf{B} and \mathbf{C} , let $\mathbf{B} \overset{\delta}{\approx} \mathbf{C}$ mean that $\bar{d}(\mathbf{B}, \mathbf{C}) < \delta$. (“Word” and “ \bar{d} -distance” are defined below.)

An interval of integers $[a .. b)$ means $[a, b) \cap \mathbb{Z}$, with analogous notation for open/closed intervals. Use \mathbb{N} for $[0 .. \infty)$ and use $[0 .. \infty]$ for $\mathbb{N} \cup \{\infty\}$. For G a set, let $\#G$ mean its cardinality; $\#\text{Stooges} = 3$.

Employ notation $G^{\times K}$ to mean the K -fold cartesian power of a set, partition, field (σ -algebra), or transformation. Let Id be the identity transformation.

I will first state three theorems, defining later the terms *rank-1*, *flat stacks*, and the set $\mathcal{J}(\mu)$ of *joinings*. Also postponed are the definitions of the standard topologies on Ω and $\mathcal{J}(\mu)$. Agree to $\text{Cl}(\cdot)$ for the closure operator on each of these two spaces.

1: Weak-Closure Theorem ([6, P.365]). *Suppose that T is rank-1. Then*

$$\heartsuit 1: \quad \mathbf{C}(T) \subset \text{Cl}(\mathbb{Z}(T)),$$

where $\text{Cl}()$ is the closure operator on Ω . ◇

An analog of the sets $\mathbf{C}(T)$ and $\mathbb{Z}(T)$ are the sets $\mathcal{J}_2^{\text{Gra}}(T)$ and $\mathcal{J}_2^{\text{Dia}}(T)$ of 2-fold “graph” and “diagonal” joinings of T . In the language of joinings, then, ($\heartsuit 1$) can be written as

$$\heartsuit 1': \quad \mathcal{J}_2^{\text{Gra}}(T) \subset \text{Cl}(\mathcal{J}_2^{\text{Dia}}(T)).$$

The graph joinings are a special case of ergodic joinings, suggesting a question raised in [6] a decade and a half ago.

2: Question. *If T is rank-1, must*

$$\mathcal{J}_2^{\text{Erg}}(T) \subset \text{Cl}(\mathcal{J}_2^{\text{Dia}}(T)) ? \quad \diamond$$

While still unresolved, the conclusion *is* known for the “flat stacks” maps, a subset of RANK-1. Indeed for flat stacks it holds for ∞ -order.

3: EJCl Theorem (Ergodic Joining-Closure). *Suppose that T has flat stacks. Then*

$$\heartsuit 2: \quad \mathcal{J}_\infty^{\text{Erg}}(T) \subset \text{Cl}(\mathcal{J}_\infty^{\text{Dia}}(T)).$$

I.e., each ergodic ∞ -order self-joining is a limit of diagonal joinings. ◇

Ergodic joinings comprise the extreme points of the simplex of *all* joinings. By beefing up $\mathcal{J}_\infty^{\text{Erg}}(T)$ to the full simplex, the theorem below strengthens ($\heartsuit 2$), but at a cost. The price paid is to conclude this, alas, no longer for all flat-stack maps, but only for a residual set of such T .

4: SJCl Theorem (Strong Joining-Closure). *The generic T has the inclusion*

$$\heartsuit 3: \quad \mathcal{J}_\infty(T) \subset \text{Cl}(\mathcal{J}_\infty^{\text{Dia}}(T)).$$

Each ∞ -order self-joining η (not necessarily ergodic) is a limit of diagonal joinings. ◇

For a dimension $\mathfrak{D} \in [2 .. \infty]$, if T satisfies

$$\heartsuit 2': \quad \mathcal{J}_\mathfrak{D}^{\text{Erg}}(T) \subset \text{Cl}(\mathcal{J}_\mathfrak{D}^{\text{Dia}}(T)), \text{ respectively,}$$

$$\heartsuit 3': \quad \mathcal{J}_\mathfrak{D}(T) \subset \text{Cl}(\mathcal{J}_\mathfrak{D}^{\text{Dia}}(T)),$$

then say that T is \mathfrak{D} -*ergodic-dense*, respectively, \mathfrak{D} -*simplex-dense*.

Words & Blockings

Over a (finite) alphabet P , an L -word \mathbf{H} is a sequence

$$\mathbf{H} = h_0 h_1 \cdots h_i \cdots h_{L-1}$$

of letters $h_i \in P$. Use $\mathbf{H}(i) = h_i$ for the i^{th} letter in a word. Notation $\mathbf{H}[3 .. 6) = h_3 h_4 h_5$ indicates a substring. The length $\text{Len}(\mathbf{H})$, above, is L . An expression such as $\mathbf{H} \overset{\text{Len}}{\leq} 5$ indicates that $\text{Len}(\mathbf{H}) \leq 5$.

For \mathbf{H} a letter or word, let $\mathbf{H}^{\otimes \mathfrak{R}}$ be the word formed by concatenating \mathfrak{R} copies of \mathbf{H} , e.g., $[\mathbf{zb}]^{\otimes 3}$ is \mathbf{zbzbzb} . Also, in characterizing flat-stacks and in the words arguments of \mathfrak{C} , \mathfrak{D} , the symbol “ \mathfrak{R} ” will always denote a positive integer. It is called a *repetition number*.

Let “ $\mathbf{H} \oplus e$ ” denote \mathbf{H} shifted (i.e., rotated) by e positions. Define this by example for an L -word \mathbf{H} :

$$\mathbf{H} \oplus 3 = h_3 h_4 h_5 \dots h_{L-2} h_{L-1} h_0 h_1 h_2.$$

Evidently $\mathbf{H} \oplus [e+L]$ equals $\mathbf{H} \oplus e$.

Joint-words. Given L -words B and C , let $\left[\begin{array}{c} B \\ C \end{array} \right]$ denote the **joint-word** over alphabet $P^{\times 2}$ consisting of letter-pairs. As a 3-dimensional example, if H is abcde12345 then

$$\left[\begin{array}{c} H \\ H \oplus 18 \\ H \oplus 1 \end{array} \right] = \begin{array}{c} abcde12345 \\ 45abcde123 \\ bcde12345a \end{array}$$

over the alphabet $P^{\times 3}$. A \mathfrak{D} -tuple \vec{e} of integers $(e_0, e_1, \dots, e_{\mathfrak{D}-1})$ will be called a **shift**. A word H and shift \vec{e} determine a **shift-word**,

$$5: \quad H^{\vec{e}} := \left[\begin{array}{c} H \oplus e_0 \\ H \oplus e_1 \\ \vdots \\ H \oplus e_{\mathfrak{D}-1} \end{array} \right].$$

Length Convention. For words we use symbols H, A, B, C, D . A script version, $\mathcal{H}, \mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$, automatically denotes the length of the corresponding word. Use L as a general length.

To avoid trivialities, all word-lengths are positive. (Some definitions fail for the empty word.)

The d -bar metric. The **\bar{d} -distance** between two L -words is $\bar{d}(B, C) := \frac{\#G}{L}$, where G is the set of indices i with $b_i \neq c_i$.

Blockings. Consider a set $G \subset \mathbb{N}$. Given a length $L \in \mathbb{N}$, let “the density of G in $[0..L]$ ” mean

$$\text{Den}_L(G) := \#(G \cap [0..L])/L.$$

Write $\underline{\text{Den}}(G)$ for the **lower density** of G in \mathbb{N} ; it is $\liminf_{L \rightarrow \infty} \text{Den}_L(G)$. Define **upper density** $\overline{\text{Den}}(G)$ analogously. If $\overline{\text{Den}}(G)$ equals $\underline{\text{Den}}(G)$ then use $\text{Den}(G)$ to denote the common density.

A **blocking** $\Upsilon := ([\ell_i .. r_i])_{i=1}^I$ of an interval $[0..L]$ is a list of intervals with $\ell_i, r_i \in \mathbb{N}$ and $\ell_i < r_i \leq \ell_{i+1}$ and $\bigcup_1^I [\ell_i .. r_i] \subset [0..L]$. Usually both L and I are ∞ ; define the general case now for use in §C. This Υ is an **ε -blocking** of $[0..L]$ if the lower density of $\bigcup_1^I [\ell_i .. r_i]$ dominates $1-\varepsilon$. An ε -blocking is also called a **$[1-\varepsilon]$ -cover**.

Consider an \mathcal{H} -word H and a name (an infinite string) $x = x[0.. \infty)$. A blocking $\Upsilon := ([\ell_i .. \ell_i + \mathcal{H}])_{i=1}^\infty$ is a “**H-blocking**” of x if

$$*: \quad x[\ell_i .. \ell_i + \mathcal{H}] = H$$

for each i . Naturally, if $\underline{\text{Den}}(\bigcup \Upsilon) \geq 1 - \varepsilon$ then Υ is called an “ **ε -H-blocking**”. Lastly, retaining the density condition but weakening $(*)$ to

$$x[\ell_i .. \ell_i + \mathcal{H}] \stackrel{\approx}{\approx} H$$

yields a “ **\bar{d} - ε -H-blocking**” of x .

Rank-1 and Flat Stacks

Given a Rohlin stack Ξ with \mathcal{H} many levels, let $P = P_\Xi$ be the partition of X into the levels of Ξ and the stack complement $X \setminus \Xi$; so P has $\mathcal{H}+1$ atoms. A map T is **rank-1** if: *For each partition Q and $\varepsilon > 0$ there exists a T -stack Ξ whose partition ε -refines Q , i.e*

$$6: \quad P \succ_\varepsilon Q.$$

A function f on a measurable set G is “**0.9-constant**” if f is constant on some subset G' with $\mu(G')/\mu(G) \geq 0.9$.

Each point $x \in \text{Base}(\Xi)$ has a first-return time, $f(x)$, to $\text{Base}(\Xi)$. Say that T has (**asymptotically flat stacks**) if: *For all Q , all small ε and large repetition integer \mathfrak{R} , there exists Ξ such that (6) and*

$$7: \quad f() \text{ is } \frac{\mathfrak{R}-1}{\mathfrak{R}}\text{-constant on } \text{Base}(\Xi).$$

Let $\mathbf{FS} \subset \mathbf{\Omega}$ denote the set of flat-stack maps. It is well known that \mathbf{FS} is $\mathbf{\Omega}$ -residual and it will also follow from the argument in §C and §D.

Criteria for rank-1 and flat stacks. The stack-shaving argument of [7, Lemma 1.4] produces, for rank-1 and *mutatis mutandis* for flat stacks, the following cut & stack and \bar{d} versions of the two properties.

For a process T, P , a Rohlin stack Ξ is **P -pure** if each of its \mathcal{H} levels lies entirely within some P -atom; consequently, it determines an \mathcal{H} -word H over the alphabet P . Each $x \in X$ inherits an H -blocking where, letting ℓ_i be the i^{th} visit-time of $T\text{-orbit}(x)$ to $\text{Base}(\Xi)$, the i^{th} block is $[\ell_i .. \ell_i + \mathcal{H}]$.

Cut&Stack criteria. T has a generating partition P and a refining sequence of P -pure stacks

$$\Xi_1 \subset \Xi_2 \subset \dots \subset \Xi_n \subset \dots \subset X$$

with heights $\mathcal{H}_n \nearrow \infty$ and $\mu(\Xi_n) \nearrow 1$. Let H_n denote the T, P -word of Ξ_n . If T has **FS** then the stacks can be chosen so that, for f_n the first-return map,

$$8: \quad f_n(\cdot) \text{ is } \frac{\mathfrak{R}_n - 1}{\mathfrak{R}_n} \text{-constant on } \text{Base}(\Xi_n),$$

for integers $\mathfrak{R}_n \nearrow \infty$. Evidently $T^{\mathcal{H}_n}$ converges in Ω to Id . A map T with such a sequence is said to be **rigid**.

Two \bar{d} -criteria. Transformation T is rank-1 IFF for every partition P : $\forall \varepsilon \exists H$ such that for a.e P -name $x \in X$:

$$6': \quad \text{There is a } \bar{d}\text{-}\varepsilon\text{-}H\text{-blocking of } x[0 .. \infty).$$

Further, T has flat stacks if: $\forall \varepsilon$ and \mathfrak{R} there exists H so that a.e x has a $\bar{d}\text{-}\varepsilon\text{-}H\text{-blocking}$ with

$$7': \quad \underline{\text{Den}}\left\{i \in \mathbb{N} \mid r_i = \ell_{i+1}\right\} \geq \frac{\mathfrak{R} - 1}{\mathfrak{R}}.$$

Along the name x , in other words, most $\bar{d}\text{-}\varepsilon$ -copies of H are immediately followed by another $\bar{d}\text{-}\varepsilon$ -copy.

The $\bar{d}\text{-FS}$ condition is perhaps best understood by example, which rotations provide.

9: Rotation FS Theorem (Katok [4]). *On the unit circle Z , each irrational rotation R has flat-stacks.* \diamond

Proof. Take two distinct points $q_1, q_2 \in Z$ and let P be the (generating) partition having the two intervals $[q_1, q_2)$ and $[q_2, q_1)$, along the circle, as atoms. For a point z in the circle and a length L , let $L\text{-word}(z)$ mean the P -word of the $R\text{-}L$ -orbit of z .

Fix an $\varepsilon > 0$ and large repetition \mathfrak{R} . Pick an open interval $I := (z_0 - 2\delta, z_0 + 2\delta)$ in Z , short enough that for each positive integer L and each point $q \in \{q_1, q_2\}$: *The $R^{-1}\text{-}L$ -orbit of q hits I fewer than $\frac{\varepsilon}{2} \cdot L$ times.* Consequently

$\forall L$ and $\forall x \in I$: *The $L\text{-word}(x)$ is $\bar{d}\text{-}\varepsilon$ -close to H ,*

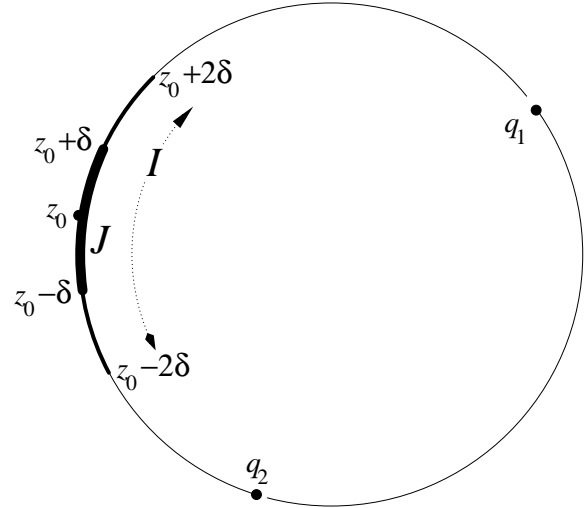


Figure 10: Each irrational rotation has flat stacks.

With clockwise direction being positive, we have a tiny interval I (shaded) and centered subinterval J half its length (thickly shaded). Exponent L is taken so that R^L rotates by less than δ/\mathfrak{R} . Consequently the $R^L\text{-}\mathfrak{R}$ -orbit of each point x in J lies entirely in I .

where $H := L\text{-word}(z_0)$.

Inside I , consider interval $J := (z_0 - \delta, z_0 + \delta)$. By compactness of Z there is a number, say ϑ , so that intervals $J, R^{-1}(J), \dots, R^{-\vartheta}(J)$ cover Z . We can have taken L so large that $\frac{\vartheta}{L} < \varepsilon$. Since \mathfrak{R} is frozen, we can have chosen L so that R^L is close to Id ; so close, that each of the points $R^L(x), R^{2L}(x), \dots, R^{\mathfrak{R}L}(x)$ is in I , whenever x started in J .

The upshot is that each point $y \in Z$ hands us a blocking

$$[a_1 .. b_1), [a_2 .. b_2), [a_3 .. b_3), \dots, \quad \text{with } b_j := a_j + \mathfrak{R}L,$$

where a_{j+1} is the smallest time dominating b_j for which $R^{a_{j+1}}(y)$ is in J . Automatically (7') holds, since each word $y[a_j .. b_j)$ comprises \mathfrak{R} consecutive $\bar{d}\text{-}\varepsilon$ -copies of H . \blacklozenge

Joinings

In order to define finite-fold and infinite-fold joinings[†] at once, fix a dimension $\mathfrak{D} \in [2..\infty]$. Below, the index j ranges over $j \in [0..\mathfrak{D}]$. (The dimension \mathfrak{D} is Fraktur D, the joining index j is Fraktur j.)

With an ergodic T in mind, we define these six sets

$$\begin{array}{ccc} \mathcal{J}^{\text{Gra}}(\mu) & \subset & \mathcal{J}(\mu) \\ 11: & \cup & \cup \\ \mathcal{J}^{\text{Dia}}(T) & \subset & \mathcal{J}^{\text{Gra}}(T) \subset \mathcal{J}^{\text{Erg}}(T) \subset \mathcal{J}(T) \end{array}$$

of \mathfrak{D} -fold joinings. Rematerialize the dependence on \mathfrak{D} when necessary, e.g write $\mathcal{J}_{\mathfrak{D}}^{\text{Erg}}(T)$. When the transformation is understood I may omit it, e.g writing \mathcal{J}^{Gra} in place of $\mathcal{J}^{\text{Gra}}(T)$.

Let $\mathcal{X}_0 \subset \mathcal{X}$ be a countable algebra of sets which generates field \mathcal{X} . A set $\vec{C} \subset X^{\times \mathfrak{D}}$ is a **cylinder set** if it can be written as a product

$$\vec{C} = \prod_{j \in [0..\mathfrak{D}]} B_j, \quad \text{with each } B_j \in \mathcal{X}_0,$$

and $B_j = X$ for all but finitely many j . There are but countably many cylinder sets.

A “ \mathfrak{D} -fold **joining of μ** ” is a measure η on $(X^{\times \mathfrak{D}}, \mathcal{X}^{\times \mathfrak{D}})$ which projects to μ in each coordinate of $[0..\mathfrak{D}]$. A joining is determined by its values on cylinder sets. Use $\mathcal{J}(\mu)$ for the collection of all \mathfrak{D} -fold joinings. Since the average of two joinings, e.g $\frac{1}{3}\eta + \frac{2}{3}\nu$, is itself a joining, this set $\mathcal{J}(\mu)$ is seen to be a simplex.

Graph and Diagonal Joinings

One \mathfrak{D} -joining of special importance is *the* diagonal joining $\Delta_{\mathfrak{D}}$, defined by

$$\Delta_{\mathfrak{D}}(\prod_j B_j) := \mu(\cap_j B_j).$$

Warping $\Delta_{\mathfrak{D}}$ by transformations will give rise to a graph joining.

Suppose that \vec{S} is a tuple $(S_j)_{j \in [0..\mathfrak{D}]}$ of maps $S_j \in \Omega$. There is a corresponding **graph joining** $\gamma^{\vec{S}}$

[†]Introductions to joinings can be found in in papers by Rudolph and del Junco, [11] [1], by Goodson [2], in papers by Lemanczyk, Parreau, and Thouvenot [10] [13] and two papers of mine, [7] [8].

specified on cylinder sets by

$$\gamma^{\vec{S}}(\prod_j B_j) := \mu(\cap_j S_j^{-1}(B_j)).$$

Use $\mathcal{J}^{\text{Gra}}(\mu)$ for the set of all graph joinings. As we only consider invertible maps, institute the convention that graph joinings $\gamma^{\vec{S}}$ are always written with S_0 equal to *Id*. (Replace each S_j by $S_j S_0^{-1}$.) With this convention, there is a bijection between such \mathfrak{D} -tuples \vec{S} and $\mathcal{J}^{\text{Gra}}(\mu)$.

A special case merits attention —when \vec{S} comprises various powers of a map T , i.e,

$$\vec{S} = (T^{e_j})_{j \in [0..\mathfrak{D}]}, \quad \text{with each } e_j \in \mathbb{Z}$$

and $e_0 = 0$. Let $\Delta^{\vec{e}}$ denote the resulting $\gamma^{\vec{S}}$ joining, i.e

$$\Delta^{\vec{e}}(\prod_j B_j) := \mu(\cap_j T^{-e_j}(B_j)).$$

Each such $\Delta^{\vec{e}}$ is called a \mathfrak{D} -fold **diagonal joining** of T . (The diagonal joining is when \vec{e} is the all-zero tuple.) Unimaginatively, use $\mathcal{J}^{\text{Dia}}(T)$ for the ensemble of diagonal joinings. When needed, the dependence on T will be signaled by writing $\Delta_T^{\vec{e}}$.

Self-joinings of a transformation

A joining $\eta \in \mathcal{J}(\mu)$ is a (*self-joining of T*) if it is T -invariant, i.e

$$\eta(\prod_j T^{-1}(B_j)) = \eta(\prod_j B_j)$$

for each cylinder set. Let $\mathcal{J}(T)$ be this set of joinings; it is a subsimplex of $\mathcal{J}(\mu)$. Analogously, $\mathcal{J}^{\text{Gra}}(T)$ is the set of T -invariant graph joinings. Evidently a graph joining $\gamma^{\vec{S}}$ is T -invariant iff each S_j commutes with T .

Ergodic joinings. Henceforth assume that T is ergodic. A joining ξ is an **ergodic joining** of T if the system

$$T^{\times \mathfrak{D}}: (X^{\times \mathfrak{D}}, \mathcal{X}^{\times \mathfrak{D}}, \xi) \circlearrowleft$$

is ergodic. Abbreviate this system as $(T^{\times \mathfrak{D}}: \xi)$.

Evidently $\Delta_{\mathfrak{D}}$ is ergodic, because the ergodic system $(T: \mu)$ is isomorphic to $(T^{\times \mathfrak{D}}: \Delta_{\mathfrak{D}})$ via the mapping $x \mapsto (x, x, \dots)$. More generally, the mapping

$$x \mapsto (S_0(x), S_1(x), S_2(x), \dots)$$

is an isomorphism from $(T: \mu)$ to $(T^{\times \mathfrak{D}}: \gamma^{\mathfrak{S}})$, revealing that $\gamma^{\mathfrak{S}}$ is an ergodic joining.

The upshot is that diagram (11) is verified.

Topologies on $\Omega(\mu)$ and $\mathcal{J}(\mu)$

Say that a net $(R_\lambda)_{\lambda \in \Lambda}$ of transformations $R_\lambda \in \Omega$ converges to S if, measuring symmetric difference,

$$\mu\left(R_\lambda^{-1}(B) \triangle S^{-1}(B)\right) \xrightarrow{\lambda} 0$$

for each set $B \in \mathcal{X}_0$ (equivalently, for each $B \in \mathcal{X}$). This defines the **coarse topology** on Ω ; it is simply the strong operator topology, with each map S interpreted as a unitary operator on the Hilbert space $\mathbb{L}^2(\mu)$.

Our measure μ is Lebesgue, so Ω is a **Polish space** in that it is homeomorphic to a complete separable metric space (e.g [3, pp.62] or [9, lemma 6]), and is non-void. In particular, the Baire Category Theorem obtains: *Each residual subset of Ω is dense*. A subset $\Omega' \subset \Omega$ is **residual** if $\Omega \setminus \Omega'$ is meager. In a topological space, a subset M is **meager** if $M \subset \bigcup_1^\infty C_n$, for some countable family of closed, interiorless sets C_n .

Standard topology on $\mathcal{J}(\mu)$. Say that a net of joinings ν_λ converges to η if

$$\nu_\lambda(\vec{C}) \xrightarrow{\lambda} \eta(\vec{C}), \quad \text{for each cylinder } \vec{C}.$$

A subsbasis for this topology is the countable collection of sets

$$12: \quad U(\vec{C}, q, \varepsilon) := \left\{ \eta \in \mathcal{J}(\mu) \mid |\eta(\vec{C}) - q| < \varepsilon \right\},$$

where \vec{C} is a cylinder, and q and ε are rational numbers in the unit interval.

Endowed with this topology, $\mathcal{J}(\mu)$ is a Polish space. It is also compact, and hence it is a Choquet simplex.

Remark. It is routine to check that the joinings topology, when restricted to $\mathcal{J}_2^{\text{Gra}}(\mu)$, equals the coarse topology carried across from Ω to $\mathcal{J}_2^{\text{Gra}}(\mu)$ via the bijection $S \mapsto \gamma^{(Id, S)}$. Thus $(\heartsuit 1')$ indeed is a restatement of $(\heartsuit 1)$ in the guise of joinings.

For an irrational rotation R on the circle Z , one can check that $\text{Cl}(\mathcal{J}_2^{\text{Dia}}(R))$ is simply the set of $\gamma^{(Id, S)}$,

as S ranges over the rotations on Z . Thus R is not 2-simplex-dense (e.g, this closure does not contain an average of two diagonal joinings) and so the conclusion of SJCl can not be strengthened to all of **FS**. \square

Finite-order suffices. An ∞ -order joining is determined by its finite-order marginal measures. Hence EJCl is equivalent to showing, for each finite \mathfrak{D} , that every rank-1 T is \mathfrak{D} -ergodic-dense. Similarly, SJCl is equivalent to taking a finite \mathfrak{D} and showing that residually many T are \mathfrak{D} -simplex-dense; then intersecting over the countably many Ω -residual sets.

Henceforth the order \mathfrak{D} is finite, i.e in $[2 .. \infty)$, and is implicit in the notation.

For example, we will just write *ergodic-dense* for \mathfrak{D} -ergodic-dense, and write \mathcal{J}^{Erg} for $\mathcal{J}_{\mathfrak{D}}^{\text{Erg}}(T)$.

§B Tools

Our inroad to studying ergodic joinings will be to study the “typical points” of an arbitrary ergodic process $(S, Q: Y, \xi)$. When applied, this S, Q process will be $T^{\times \mathfrak{D}}, P^{\times \mathfrak{D}}$, with Y being $X^{\times \mathfrak{D}}$.

Frequencies. Over an alphabet Q , suppose that words \mathbf{B} and \mathbf{C} have lengths $\mathcal{B} \leq \mathcal{C}$. With \mathbf{I} denoting the interval $[0 .. \mathcal{C} - \mathcal{B}]$, define

$$\text{Freq}(\mathbf{B} \text{ in } \mathbf{C}) := \frac{\#\{i \in \mathbf{I} \mid C[i .. i + \mathcal{B}] = \mathbf{B}\}}{\#\mathbf{I}}.$$

A point $y \in Y$ is **generic** if

$$\text{Freq}(\mathbf{B} \text{ in } y[0 .. \infty)) := \lim_{\mathcal{C} \rightarrow \infty} \text{Freq}(\mathbf{B} \text{ in } y[0 .. \mathcal{C}])$$

exists and equals $\xi(\mathbf{B})$, for each word \mathbf{B} . Here $\xi(\mathbf{B})$ is the ξ -mass of the S, Q -cylinder set that \mathbf{B} determines.

Let’s now examine a stronger notion of genericity.

Typical points. A word C is ε -good for ξ if

$$\text{Freq}(\mathbf{B} \text{ in } C) \stackrel{\varepsilon}{\approx} \xi(\mathbf{B}),$$

for each word $\mathbf{B} \stackrel{\text{Len}}{\leq} 1/\varepsilon$. An ε -bad C , consequently, admits a word $\mathbf{B} \stackrel{\text{Len}}{\leq} \frac{1}{\varepsilon}$ whose frequency in C differs by more than ε from $\xi(\mathbf{B})$.

A **min- L -blocking**, given a length L , is a blocking $([\ell_i .. r_i])_{i=1}^\infty$ where each length $r_i - \ell_i$ dominates L . A point $y \in Y$ is **typical** (for ξ) if, for all ε and δ , there exists a large enough L so that the following statement obtains.

Whenever $([\ell_i .. r_i])_{i=1}^\infty$ is a min- L -blocking with each $y[\ell_i .. r_i]$ an ε -bad word, then its upper density is δ -dominated, i.e

$$\overline{\text{Den}}\left(\bigcup_{i=1}^\infty [\ell_i .. r_i]\right) \leq \delta.$$

13: Typicality Lemma. An ergodic process $(S, Q: Y, \xi)$, after removing an invariant nullset from Y , has every point $y \in Y$ typical. *Proof.* In End-Notes. \diamond

Establishing the EJCI Theorem

In the sequel T, P is a flat-stack process as in paragraph (8). Shrink each \mathfrak{R}_n , if necessary, so that repetition numbers $\mathfrak{R}_n \nearrow \infty$ still witness **FSness**, and now also bound stack-mass in that

$$14: \quad \mu(\Xi_n) > 1 - \frac{1}{\mathfrak{R}_n}.$$

Remove a nullset from X so that now for each $x \in X$ and each n , the H_n -blocking $([\ell_i .. \ell_i + \mathcal{H}_n])_{i=1}^\infty$ of x has a density

$$\text{Den}\left(\bigcup_{i=1}^\infty [\ell_i .. \ell_i + \mathcal{H}_n]\right) = \mu(\Xi_n).$$

Below, once a stage n has been chosen, then automatically $\Xi, H, \mathcal{H}, \mathfrak{R}$ will represent $\Xi_n, H_n, \mathcal{H}_n, \mathfrak{R}_n$ as in (8). Also, S is $T^{\times \mathfrak{D}}$ and Q is $P^{\times \mathfrak{D}}$.

Shift-words. Taking $\mathfrak{D}=3$ as an example, consider any three points $x_0, x_1, x_2 \in X$ and, fixing a large n , their n -blockings. Most H -blocks $[\ell .. r)$ on x_0 are **amiable** in being entirely covered by (parts of) two

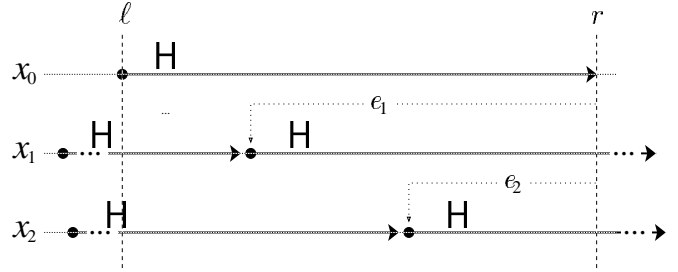


Figure 15: An amiable H -block $x_0[\ell .. r)$ is entirely covered by H -blocks on x_1 and on x_2 . As drawn here, $e_1 > e_2 > e_0 = 0$. Let $H^{\vec{e}}$ denote the resulting $P^{\times 3}$ -word.

consecutive H -blocks on x_1 , and by two consecutive on x_2 , as Fig. 15 shows. Temporarily, let r_j denote the unique position in $[\ell .. r)$ so that $[r_j - \mathcal{H} .. r_j)$ is an H -block on x_j . Let $e_j := r - r_j$ be the corresponding shift-length; thus e_0 is zero.

Let $x := x_0$. Over our alphabet $Q = P^{\times 3}$ we have two Q -names

$$y := (x_j)_{j \in [0..3)} \quad \text{and} \quad x^{\vec{e}} := (T^{e_j}(x))_{j \in [0..3)}.$$

When restricted to our amiable block, they give equal Q -words,

$$y[\ell .. r) = x^{\vec{e}}[\ell .. r) =: H^{\vec{e}}.$$

This, since each x_j has no gap between its H -blocks that touch $[\ell .. r)$. The common Q -word, $H^{\vec{e}}$, can be thought of as three P -words written one atop another. It is the joint-word

$$H^{\vec{e}} = \begin{bmatrix} H \oplus 0 \\ H \oplus e_1 \\ H \oplus e_2 \end{bmatrix},$$

which was called a shift-word in (5).

Shift-words mostly cover y . The above shift \vec{e} has each $e_i \in [0.. \mathcal{H})$. Say that such an \vec{e} is \mathcal{H} -**bounded**.

How much of $y[0.. \infty)$ is covered by shift-words $H^{\vec{e}}$, for varying \mathcal{H} -bounded shifts \vec{e} ? Letting δ denote $\mu(X \setminus \Xi)$, it is straightforward that at least density

$$\mu(\Xi) - [\mathfrak{D} - 1] \cdot [\delta + \frac{1}{\mathfrak{R}}]$$

is so covered. Rematerializing the stage n and using (14) shows that at least density

$$16: \quad 1 - \frac{2\mathfrak{D}}{\mathfrak{R}_n}$$

of $y[0.. \infty)$ is covered by H_n -shift-words.

Mild uniformity for FS. Arguably, the “reason” that an irrational rotation R has flat stacks (9) is due to the set $\mathbb{Z}(R)$ of powers being an equicontinuous family. A general **FS** map T has the following milder uniformity property.

17: **FS Uniformity Lemma.** *For all ε and all large n :*

Shift-word $H_n^{\vec{e}}$ is ε -good for the $\Delta^{\vec{e}}$ diagonal joining,

for **each** \mathcal{H}_n -bounded shift \vec{e} . ◇

Proof. Take an $x \in X$; it is generic for $(T, P: \mu)$. For each shift \vec{e} , then, the point

$$x^{\vec{e}} := (T^{e_0}(x), T^{e_1}(x), \dots, T^{e_{\mathfrak{D}-1}}(x))$$

is generic for the $(S, Q: \Delta^{\vec{e}})$ process.

Take a stage n large enough that

$$\frac{1}{\mathfrak{R}} < \frac{\varepsilon}{4} \quad \text{and} \quad \mathcal{H} > \frac{4}{\varepsilon^2}.$$

(Recall that \mathfrak{R} means \mathfrak{R}_n , etc.) For our \mathcal{H} -bounded \vec{e} , at least density $1 - \frac{2}{\mathfrak{R}}$ of the Q -name $x^{\vec{e}}$ is covered by copies of $H^{\vec{e}}$. For each Q -word $\mathbf{B} \stackrel{\text{Len}}{\leq} \frac{1}{\varepsilon}$, consequently,

$$\text{Freq}(\mathbf{B} \text{ in } x^{\vec{e}}[0.. \infty)) \quad \text{differs from} \quad \text{Freq}(\mathbf{B} \text{ in } H^{\vec{e}})$$

by at most $\frac{2}{\mathfrak{R}} + \frac{1/\varepsilon}{\mathcal{H}}$. This is less than $\frac{2\varepsilon}{4} + \frac{\varepsilon}{4}$, so $H^{\vec{e}}$ is ε -good for $(S, Q: \Delta^{\vec{e}})$. ◇

We now combine the preceding tools.

Proof of EJCI Theorem. Fix an ergodic $\xi \in \mathcal{J}_{\mathfrak{D}}^{\text{Erg}}(T)$.

It suffices to fix a positive ε and produce a shift \vec{e} and Q -word which

Goal: \dots is ε -good for ξ and
is ε -good for $\Delta^{\vec{e}}$ as well.

For specificity, take $\mathfrak{D} = 3$. Choose a point $y = (x_0, x_1, x_2)$ which is $(S, Q: \xi)$ -typical.

An ε -good word $H^{\vec{e}}$. Take n so that $\mathcal{H} = \mathcal{H}_n$ is large enough for the *Typicality Lemma* in that at most density $\frac{1}{3}$ of y can be covered by min- \mathcal{H} -blocks which are ε -bad for ξ . Also take $\mathfrak{R} = \mathfrak{R}_n$ big so that, courtesy (16), at least $\frac{2}{3}$ of y is covered by H -shift-words. Hence there exists an \mathcal{H} -bounded shift \vec{e} such that: $H^{\vec{e}}$ is ε -good for ξ .

Chasing one’s tail? Is this Q -word $H^{\vec{e}}$ also ε -good for the $\Delta^{\vec{e}}$ joining? A natural avenue would be to again appeal to the *Typicality Lemma*; but there is no reason for the current \mathcal{H}_n to be sufficiently large when applied to the $(S, Q: \Delta^{\vec{e}})$ process. And were we to enlarge \mathcal{H}_n (by increasing n) to be large enough, then this likely would alter \vec{e} . Happily, the argument can be rendered non-circular[†] via **FSness**.

Using mild uniformity. Since ε was known in advance, we could have initially taken n big enough for (17) and \mathcal{H}_n big enough for ξ . Since the resulting \vec{e} is \mathcal{H}_n -bounded automatically, our $H^{\vec{e}}$ is ε -good for $(S, Q: \Delta^{\vec{e}})$. ◆

Reduction of SJCI to 2-mixtures

A probability vector is a tuple (b_1, \dots, b_M) of non-negative reals which sum to 1.

A **2-mixture** is a joining of the form

$$18: \quad a\Delta^{\vec{p}} + b\Delta^{\vec{q}}, \quad \text{with } (a, b) \text{ a probability vector,}$$

where b is rational and \vec{p} and \vec{q} are each \mathfrak{D} -shifts. Let $\text{Two}(T)$ denote the set of such mixtures.

[†]It is this potential circularity which obstructs one approach to question (2) of whether every rank-1 is ergodic-dense.

19: **Reduction Lemma.** *Suppose T has flat stacks. If*

$$19': \quad \text{Cl}(\mathcal{J}^{\text{Dia}}(T)) \supset \text{Two}(T)$$

then T is simplex-dense. \diamond

Proof. For shifts $\vec{p}, \vec{q}, \vec{r}$ and a probability vector (a, b, c) , we can write the 3-mixture

$$a\Delta^{\vec{p}} + b\Delta^{\vec{q}} + c\Delta^{\vec{r}} \stackrel{\text{note}}{=} [1 - c]\nu + c\Delta^{\vec{r}}$$

by defining the 2-mixture

$$\nu := \frac{a}{1 - c}\Delta^{\vec{p}} + \frac{b}{1 - c}\Delta^{\vec{q}}.$$

Courtesy (19') we can first pick a diagonal $\Delta^{\vec{s}}$ near to ν , then choose a diagonal near to

$$[1 - c]\Delta^{\vec{s}} + c\Delta^{\vec{r}}.$$

Sending “near” $\searrow 0$, so to speak, says that

$$\text{Cl}(\text{Two}(T)) \supset \text{THREE}(T).$$

Thus $\text{Cl}(\mathcal{J}^{\text{Dia}})$ owns all the 3-fold mixtures. Iterating the reduction shows that $\text{Cl}(\mathcal{J}^{\text{Dia}})$ contains all mixtures of finitely many diagonal joinings.

Because T has flat stacks, the *EJCI Theorem* assures that \mathcal{J}^{Dia} is dense in \mathcal{J}^{Erg} . Consequently, $\text{Cl}(\mathcal{J}^{\text{Dia}})$ contains each finite-fold mixture of ergodic self-joinings of T . As $\text{Cl}(\mathcal{J}^{\text{Dia}})$ is closed, it must therefore equal all of $\mathcal{J}(T)$. \blacklozenge

Remark. The set of diagonals $\Delta^{\vec{e}}$ with \vec{e} non-negative is dense in $\mathcal{J}^{\text{Dia}}(T)$, since T is rigid. We thus redefine $\text{Two}(T)$ to only comprise rational 2-mixtures with non-negative shifts (each $e_j \geq 0$), and the *Reduction Lemma* will persist. Consequently:

All shifts, $\vec{e}, \vec{s}, \vec{p}, \dots$, are non-negative

for the remainder of the paper. \square

§C Strong Joining-Closure is Ω -residual

To exhibit a residual set of \mathfrak{D} -simplex-dense maps T , it suffices to fix shifts \vec{p} and \vec{q} as well as positive rationals $a + b = 1$, then show:

20: *There is a residual set of T for which the mixture $a\Delta_T^{\vec{p}} + b\Delta_T^{\vec{q}}$ is in $\text{Cl}(\mathcal{J}^{\text{Dia}}(T))$.*

After all, there are but countably many triples (\vec{p}, \vec{q}, b) , so residually many T must fill the hypothesis of the foregoing *Reduction Lemma*. Since **FS** is itself residual, the conclusion of simplex-density will hold residually.

There is no loss of generality in assuming, as done in the sequel, that $b = \frac{2}{7}$.

Prolegomenon. Given a partition P on (X, μ) , each \mathfrak{D} -joining η engenders a vector with $[\#P]^{\mathfrak{D}}$ many components, namely, the vector

$$21: \quad Q \rightarrow [0, 1]: \mathbf{q} \mapsto \eta(\mathbf{q})$$

as \mathbf{q} ranges over the atoms $\mathbf{q} \in Q := P^{\times \mathfrak{D}}$. If joinings η and ν have ε -close vectors then we say that η and ν are ε -***P***-close, i.e

$$\forall \mathbf{q} \in P^{\times \mathfrak{D}}: |\eta(\mathbf{q}) - \nu(\mathbf{q})| < \varepsilon.$$

Because of the strict “ $<$ ”, the set

$$\text{Open}_\varepsilon^P := \left\{ T \mid \begin{array}{l} \text{There exists a shift } \vec{s} \text{ such that } \Delta_T^{\vec{s}} \\ \text{is } \varepsilon\text{-}P\text{-close to } \frac{5}{7}\Delta_T^{\vec{p}} + \frac{2}{7}\Delta_T^{\vec{q}} \end{array} \right\},$$

is an Ω -open set of maps. So $\text{Open}_\varepsilon^P$ is residual **if** we can establish that $\text{Open}_\varepsilon^P$ is dense in Ω .

Intersecting all the $\text{Open}_\varepsilon^P$ over a countable dense family of partitions P and countably many $\varepsilon \searrow 0$, will produce the residual set desired by (20). The upshot is that we can freeze the following data.

In the sequel, we have fixed: Shifts \vec{p} and \vec{q} , an $\varepsilon \in (0, 1)$ and partition P on X , as well as a non-empty open set $\mathbf{U} \subset \Omega$. Also, Q denotes $P^{\times \mathfrak{D}}$, the cartesian-power partition.

Our goal now is to construct a T in \mathbf{U} and a shift \vec{s} for which:

20bis: *Diagonal-joining $\Delta_T^{\vec{s}}$ is ε -*P*-close to mixture $\frac{5}{7}\Delta_T^{\vec{p}} + \frac{2}{7}\Delta_T^{\vec{q}}$.*

We will produce T and \vec{s} by computing frequencies.

Frequencies estimate a diagonal joining. Here is a device to approximate the η -vector of (21) when η is some diagonal $\Delta_T^{\vec{e}}$. Imagine, for T , that we have a P -pure Rohlin stack Ξ of some height \mathcal{H} , and let H be its P -word. Suppose that $\mathcal{H} \gg m := \max(\vec{e})$, where

$$\max(\vec{e}) := \max \{e_j \mid j \in [0 .. \mathfrak{D}]\}.$$

Each position $i \in [0 .. \mathcal{H} - m]$ determines a Q -letter whose \mathfrak{D} many P -letters are

$$*: \quad H(i+e_0), H(i+e_1), H(i+e_2), \dots, H(i+e_{\mathfrak{D}-1}).$$

For a letter $q \in Q$ we can compute the frequency of $i \in [0 .. \mathcal{H} - m]$ such that Q -letter (*) equals q . This frequency will be 2ε -close to $\Delta_T^{\vec{e}}(q)$ if

$$\mu(X \setminus \Xi) < \varepsilon \quad \text{and} \quad \frac{m}{\mathcal{H}} < \varepsilon.$$

Indeed, 2ε -closeness holds if we average over all of $[0 .. \mathcal{H}]$, by conceptually wrapping H into a circle and computing the frequency of q in the shift-word $H^{\vec{e}}$, defined in (5).

Let $\mathbb{F}(H^{\vec{e}})$ denote the mapping

$$21': \quad Q \longrightarrow [0, 1]: q \mapsto \text{Freq}(q \text{ in } H^{\vec{e}}),$$

where we regard each letter q as a Q -word of length 1.

How can we use the foregoing discussion to restate (20bis)? Our task now is to find a $T \in \mathbf{U}$ and P -pure Rohlin stack Ξ , with $\mu(\Xi) > 1 - \varepsilon$, whose T -word H fulfills the following for some shift \vec{s} :

$$\begin{aligned} & \text{Vector } \mathbb{F}(H^{\vec{s}}) \text{ is } \varepsilon\text{-close to the average} \\ & \frac{5}{7}\mathbb{F}(H^{\vec{p}}) + \frac{2}{7}\mathbb{F}(H^{\vec{q}}). \\ 20\text{ter: } & \text{Furthermore, each of } \max(\vec{s}), \max(\vec{p}) \text{ and} \\ & \max(\vec{q}) \text{ is less than } \varepsilon \cdot \text{Len}(H). \end{aligned}$$

We now argue the *SJCl Theorem*, making use of the *Chameleon Lemma* (following the argument) which is a combinatorial frequency result. Handed words A and H , say that H is an ε -**A-word**, or is ε -**A-blocked**, if interval $[0 .. \mathcal{H}]$ admits an ε -**A-blocking**

$$\Upsilon := ([\ell_i .. \ell_i + A])_{i=1}^I$$

where each substring $H[\ell_i .. \ell_i + A]$ equals A and $\text{Den}(\cup \Upsilon)$ dominates $1 - \varepsilon$.

Conditional proof of SJCl Theorem

Since the family of ergodic maps is Ω -dense, we can pick an ergodic $S \in \mathbf{U}$. As \mathbf{U} is open, we can *shrink the given* ε so that every perturbation of S to a nearby map T , by altering S on (at most) mass ε , necessarily has $T \in \mathbf{U}$.

Let Ξ'' be an S -Rohlin-stack with

$$22: \quad \mu(X \setminus \Xi'') < \frac{\varepsilon}{2} \quad \text{and} \quad \mu(\text{Base}(\Xi'')) < \frac{\varepsilon}{2}.$$

Columnate Ξ'' into P -pure columns. Remove a little mass from Ξ'' by thinning the columns so that, now, the column-widths are commensurate with each other. This can be done preserving (22).

Further subdivide the columns so that they now have a common width. Stack these columns, one atop the next, to form Ξ' , a P -pure stack. Finally, let A be the P -word of Ξ' .

A cut & stack recipe. Pick a positive number $\delta < \mu(X \setminus \Xi')$. The *Frequency Mixture Corollary* (24), below, provides a shift \vec{s} and δ -**A-word** H satisfying this statement about letter-frequencies:

$$\mathbb{F}(H^{\vec{s}}) \approx \frac{\delta}{7}\mathbb{F}(H^{\vec{p}}) + \frac{2}{7}\mathbb{F}(H^{\vec{q}}).$$

Since H is $[1 - \delta]$ -covered by **A**-copies, we can (using no more than mass δ of $X \setminus \Xi'$ as spacer) use H as a template to cut & stack Ξ' into a stack Ξ of height \mathcal{H} , producing a (new) map T .

In going from S to T we made alterations on the top level of stack Ξ'' and the stack complement. This, courtesy (22), has at most ε mass; our T is indeed in \mathbf{U} . Furthermore, $\mu(\Xi) \geq \mu(\Xi'') > 1 - \varepsilon$.

Since $\delta < \varepsilon$, inequality (24.1) assures that triple T, Ξ, H fulfills (20ter), thus establishing the *Strong Joining-Closure Theorem*. \blacklozenge

Word Combinatorics

The *Chameleon Lemma* below produces a ‘‘chameleon word’’ D which adapts its letter-frequencies to different shifts \vec{e} , and does so changing neither its length \mathcal{D} nor its shift \vec{s} . It is this independence which permits its corollary concerning mixtures of frequencies.

23: Chameleon Lemma. Given a word A and $\delta > 0$, there is a length \mathcal{D} and shift \vec{s} satisfying the following: For each $\frac{1}{5}$ -bounded shift \vec{e} , there exists a δ - A -word $D = D_{\vec{e}}$ such that the shift-words $D^{\vec{s}}$ and $A^{\vec{e}}$ have nearby frequency vectors:

$$\mathbb{F}(D^{\vec{s}}) \text{ is } \delta\text{-close to } \mathbb{F}(A^{\vec{e}}).$$

Furthermore, D has length \mathcal{D} . \diamond

Proof. The *Blocked-Chameleon Lemma* (29), an enhanced version, will be argued in §D. \blacklozenge

24: Frequency Mixture Corollary. Consider a word A and positive δ . Then there exists a shift \vec{s} and a δ - A -word H for which

$$24.1: \quad \max(\vec{s}), \max(\vec{p}), \max(\vec{q}) < \delta \cdot \text{Len}(H);$$

$$24.2: \quad \mathbb{F}(H^{\vec{s}}) \stackrel{\delta}{\approx} \frac{5}{7}\mathbb{F}(H^{\vec{p}}) + \frac{2}{7}\mathbb{F}(H^{\vec{q}}). \quad \diamond$$

Proof. Shrink δ so that $\max(\vec{p}), \max(\vec{q}) < \frac{1}{5}$, ie. each of \vec{p} and \vec{q} is a $\frac{1}{5}$ -bounded shift.

Replace the input word A by a long concatenation $A^{\otimes \mathfrak{R}}$ to *gratis* yield

$$*: \quad \max(\vec{p}), \max(\vec{q}) < \delta \cdot \text{Len}(A).$$

Replacing the output H by an $H^{\otimes \mathfrak{R}}$ preserves letter-frequencies in all shift-words $H^{\vec{e}}$. Thus (24.2) is unaffected, so inequality (24.1) comes for free.

The *Chameleon Lemma* provides a shift \vec{s} and length \mathcal{D} , as well as two \mathcal{D} -words

$$B := D_{\vec{p}} \quad \text{and} \quad C := D_{\vec{q}}$$

for which

$$\mathbb{F}(B^{\vec{s}}) \stackrel{\delta}{\approx} \mathbb{F}(A^{\vec{p}}) \quad \text{and} \quad \mathbb{F}(C^{\vec{s}}) \stackrel{\delta}{\approx} \mathbb{F}(A^{\vec{q}}).$$

Choose a big repetition \mathfrak{R} and define \mathcal{H} to be the length of this long concatenation

$$H := \overbrace{\text{BBBBB} \dots \text{B}}^{5\mathfrak{R} \text{ copies}} \overbrace{\text{CC} \dots \text{C}}^{2\mathfrak{R} \text{ copies}}$$

of words. Since \mathfrak{R} could have been taken large enough to swamp the BC and CB transitions in H (viewed as a circular word) we *gratis* obtain that

$$**: \quad \mathbb{F}(H^{\vec{s}}) \stackrel{\delta}{\approx} \frac{5}{7}\mathbb{F}(A^{\vec{p}}) + \frac{2}{7}\mathbb{F}(A^{\vec{q}}).$$

Coverage by word A. Inequality (*) implies that

$$\mathbb{F}(A^{\vec{p}}) \stackrel{2\delta}{\approx} \mathbb{F}(H^{\vec{p}}),$$

since H is $[1 - \delta]$ -covered by A . Observing the same for \vec{q} thus evinces

$$\frac{5}{7}\mathbb{F}(A^{\vec{p}}) + \frac{2}{7}\mathbb{F}(A^{\vec{q}}) \stackrel{2\delta}{\approx} \frac{5}{7}\mathbb{F}(H^{\vec{p}}) + \frac{2}{7}\mathbb{F}(H^{\vec{q}}).$$

Together with (**), this delivers (24.2) but with δ replaced by 3δ . \blacklozenge

§D Reformulating the *Chameleon Lemma*

Strengthening the *Chameleon Lemma* to be cognizant of blockings will be central in our proof of it. Intuitively, the “Nested Gym” recipe (28), below, builds the word D demanded by this strengthening, the *Blocked-Chameleon Lemma* (29). For clarity of the idea –at the expense of formality– the argument proceeds via a “mini-calculus of blockings” in the form of four almost-tautologies (the *Joint-word*, *Gym-shift*, *Same-shift*, and *Composition* propositions) whose proofs I sketch or leave to the Reader. After justifying *Blocked-Chameleon* for $\mathfrak{D}=2$, I illustrate how to argue it by induction on dimension \mathfrak{D} , by detailing the $\mathfrak{D}=3 \Rightarrow \mathfrak{D}=4$ implication.

Blockings revisited. Consider two words A and D . Suppose that we produce Υ , a blocking $([\ell_i .. \ell_i + A])_{i=1}^I$ of the interval $[0 .. \mathcal{D}]$, so that those Υ -blocks with

$$D[\ell_i .. \ell_i + A] = A$$

cover at least density $1 - \delta$ of $[0 .. \mathcal{D}]$. Write this relation as

$$D \xrightarrow[\delta]{\Upsilon} A$$

and note that it implies $\mathbb{F}(D) \stackrel{\delta}{\approx} \mathbb{F}(A)$, closeness in frequency.

Below, propositions will usually take a blocking Υ of a P -word by another P -word, then apply Υ to joint-words; that is, words over some alphabet $P^{\times N}$.

25: **Joint-word Prop'n.** Consider words D_1, \dots, D_N each of length \mathcal{D} , and words A_1, \dots, A_N of length \mathcal{A} . Suppose that Υ is a blocking of $[0.. \mathcal{D})$ by \mathcal{A} -blocks. Further imagine that

$$\forall j: D_j \xrightarrow[\varepsilon_j]{\Upsilon} A_j.$$

Letting δ denote $\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_N$, then,

$$\left[\begin{array}{c} D_1 \\ \vdots \\ D_N \end{array} \right] \xrightarrow[\delta]{\Upsilon} \left[\begin{array}{c} A_1 \\ \vdots \\ A_N \end{array} \right].$$

In particular, the joint-words have δ -close letter-frequencies. \diamond

Swimming Pools

Fix a particular P -letter to be used as “spacer”; use \square to denote this letter.

Given an \mathcal{A} -word A and repetition \mathfrak{R} , each integer e determines a “swimming pool”: The word

$$S_e := \square^{\otimes[A-\hat{e}]} A^{\otimes \mathfrak{R}} \square^{\otimes \hat{e}}$$

where $\hat{e} \in [0.. \mathcal{A})$ denotes the residue of e mod \mathcal{A} . We think of concatenation $A^{\otimes \mathfrak{R}}$ as “floating” in the pool, with its location specified by integer e . A pool’s length is

$$\text{Lap}^{\mathfrak{R}}(A) := [1+\mathfrak{R}] \cdot \text{Len}(A).$$

There are only \mathcal{A} many pools, since $S_{e+\mathcal{A}} = S_e$.

A gymnasium often has several swimming pools, so lets use “Gym” to name the following concatenation of pools,

$$\text{Gym}^{\mathfrak{R}}(A, e) := S_0 S_e S_{2e} S_{3e} \dots S_{[A-1]e}.$$

Its length is $\mathcal{A} \cdot \text{Lap}^{\mathfrak{R}}(A) \stackrel{\text{note}}{=} [1+\mathfrak{R}] \mathcal{A}^2$.

Defining a word $D := \text{Gym}^{\mathfrak{R}}(A, e)$ gives a natural blocking of D by A -copies. Call such a blocking “ da ”, using the names of the words to make a name for the list of blocks $[l_i .. l_i + \mathcal{A})$.

26: **Gym-shift Prop'n.** Freeze an $\mathfrak{R} > \frac{1}{\delta}$ and natural number e . Given C , a word, define

$$s := \text{Lap}^{\mathfrak{R}}(C), \quad \text{a natural number};$$

$$D := \text{Gym}^{\mathfrak{R}}(C, e), \quad \text{a long word};$$

and let dc denote the resulting C -blocking of D . Then

$$\left[\begin{array}{c} D \\ D \oplus s \end{array} \right] \xrightarrow[2\delta]{dc} \left[\begin{array}{c} C \\ C \oplus e \end{array} \right].$$

$$\text{Equivalently, } D \oplus s \xrightarrow[2\delta]{dc} C \oplus e. \quad \diamond$$

Commentary. For varying shift-lengths e , the Prop'n allows us to mimic the frequency of letter-pairs in $\left[\begin{array}{c} C \\ C \oplus e \end{array} \right]$ using a shift $s = [1+\mathfrak{R}]C$ and length $\mathcal{D} = [1+\mathfrak{R}]C^2$ which are independent of e . \square

Sketch of Proof. For specificity, suppose that length C

is 100; so word $\left[\begin{array}{c} D \\ D \oplus s \end{array} \right]$ equals

$$\begin{array}{cccccccc} S_0 & S_e & S_{2e} & S_{3e} & \dots & S_{98e} & S_{99e} & \\ S_e & S_{2e} & S_{3e} & S_{4e} & \dots & S_{99e} & S_0 & \dots \end{array}$$

Fixing an n and $m:=n+1$, consider a joint-pool $\left[\begin{array}{c} S_{ne} \\ S_{me} \end{array} \right]$. In the upper word S_{ne} , all but density δ is covered by consecutive C -blocks. Each C -block sees, in $\left[\begin{array}{c} S_{ne} \\ S_{me} \end{array} \right]$, the joint-word $\left[\begin{array}{c} C \\ C \oplus e \end{array} \right]$ —well, except for the leftmost or rightmost C -block. Since $\mathfrak{R} > \frac{1}{\delta}$, this lost block costs us at most density δ . So 2δ bounds the total expense. \blacklozenge

There is an analogous proposition when the shifts of D and C are equal. Below, from a $C^{\otimes \mathfrak{R}}$ -blocking, we examine how the induced C -blocking interacts with shift-words $D^{\vec{s}}$ and $C^{\vec{s}}$.

27: **Same-Shift Prop'n.** Fix an $\mathfrak{R} > 1/\delta$. Suppose that D is a word which is $[1-\delta]$ -covered by copies of $C^{\otimes \mathfrak{R}}$. Let dc denote the resulting C -blocking of D . Then

$$D^{\vec{s}} \xrightarrow[3\delta]{dc} C^{\vec{s}},$$

for each shift \vec{s} with $\max(\vec{s}) \leq \delta \cdot \mathfrak{R} \cdot \text{Len}(C)$. In particular, this applies when $\max(\vec{s}) \leq \text{Len}(C)$. \diamond

Nested Gym Construction

Given δ and word A , pick a large repetition \mathfrak{R} and let $B_0 := A$. Each shift

$$\vec{e} = (e_0, e_1, e_2, \dots, e_{\mathfrak{D}-1})$$

will give us a list of words:

$$B_j := \text{Gym}^{\mathfrak{R}}(B_{j-1}, e_j) \text{ and}$$

$$28: \quad s_j := \text{Lap}^{\mathfrak{R}}(B_{j-1}), \text{ for } j = 1, 2, \dots, \mathfrak{D}-1.$$

Then define $D := B_{\mathfrak{D}-1}$.

Let \vec{s} be the shift $(0, s_1, \dots, s_{\mathfrak{D}-1})$. Our goal is to show that D satisfies a “blocking” strengthening of (23), the *Chameleon Lemma*.

Below, word D adapts to \vec{e} , but neither its length \mathfrak{D} nor its shift \vec{s} , change. Quantities \mathfrak{D} and \vec{s} depend solely on the three integers \mathfrak{D} , A and \mathfrak{R} .

29: **Blocked-Chameleon Lemma.** *Given dimension \mathfrak{D} , length A , positive δ and \mathfrak{D} -shift \vec{e} , then the following holds for all large \mathfrak{R} : For A an arbitrary A -word, define \vec{s} and D as in the foregoing (28). Then*

$$D^{\vec{s}} \xrightarrow[\delta]{\text{da}} A^{\vec{e}}.$$

Here, da is the canonical blocking of D by A -copies, arising from the (28) construction. \diamond

Induction on \mathfrak{D} . The $\mathfrak{D}=2$ case of *Blocked Chameleon* follows from *Gym-shift*. The idea of the general case will be clear by showing one step of the induction. Let us establish the $\mathfrak{D}=4$ case from $\mathfrak{D}=3$, giving explicit names to all the words from (28). There will be no true loss of generality in assuming that the given shift is

$$(e_0, e_1, e_2, e_3) = (0, 17, 29, 38).$$

(The ordering among 17, 29, 38 is irrelevant.)

Construction (28) gives

$$B := \text{Gym}^{\mathfrak{R}}(A, 17), \quad p := \text{Lap}^{\mathfrak{R}}(A);$$

$$28': \quad C := \text{Gym}^{\mathfrak{R}}(B, 29), \quad q := \text{Lap}^{\mathfrak{R}}(B);$$

$$D := \text{Gym}^{\mathfrak{R}}(C, 38), \quad r := \text{Lap}^{\mathfrak{R}}(C).$$

Each pair of words from D, C, B, A produces a canonical blocking; use ca for the blocking of C by A -copies,

etc.. These blockings form a nested family: Blocking da is the “composition” $\text{dc} \circ \text{ca}$ in the sense of putting a copy of blocking ca at the location of *each* block in dc .

30: **Composition Prop'n.** *Consider three lengths $\mathfrak{D} \geq \mathfrak{C} \geq A$ and three blockings. Let $\text{da} := \text{dc} \circ \text{ca}$, where*

dc : *A blocking of $[0 .. \mathfrak{D})$ by length- \mathfrak{C} blocks;*

ca : *A blocking of interval $[0 .. \mathfrak{C})$ by A -blocks.*

Suppose we have words D, C, A with lengths $\mathfrak{D}, \mathfrak{C}, A$. For ε and δ positive, these next two relations

$$\begin{array}{c} D \xrightarrow[\varepsilon]{\text{dc}} C \\ C \xrightarrow[\delta]{\text{ca}} A \end{array} \text{ imply that } D \xrightarrow[\varepsilon+\delta]{\text{da}} A. \quad \diamond$$

Proof of Blocked-Chameleon for $\mathfrak{D}=4$

Take \mathfrak{R} large enough that blocking da from (28') is a δ -blocking. So *all six* of the blockings $\text{ca}, \text{db}, \dots$, are δ -blockings. Enlarge \mathfrak{R} enough that

$$\ddagger: \quad \mathfrak{R} > \frac{1}{\delta} \text{ and } \max(17, 29, 38) < \delta \mathfrak{R}.$$

Further suppose \mathfrak{R} big enough that

$$\left[\begin{array}{c} C \\ C \oplus p \\ C \oplus q \end{array} \right] \xrightarrow[\delta]{\text{ca}} \left[\begin{array}{c} A \\ A \oplus 17 \\ A \oplus 29 \end{array} \right],$$

courtesy of *Blocked-Chameleon* for $\mathfrak{D}=3$.

Since $p, q \leq \text{Len}(C)$, we may apply *Same-Shift* to obtain

$$\left[\begin{array}{c} D \\ D \oplus p \\ D \oplus q \end{array} \right] \xrightarrow[3\delta]{\text{dc}} \left[\begin{array}{c} C \\ C \oplus p \\ C \oplus q \end{array} \right].$$

Composing this with the preceding display, the *Composition Prop'n* gives

$$31: \quad \left[\begin{array}{c} D \\ D \oplus p \\ D \oplus q \end{array} \right] \xrightarrow[4\delta]{\text{da}} \left[\begin{array}{c} A \\ A \oplus 17 \\ A \oplus 29 \end{array} \right].$$

Recall that D equals $\text{Gym}^{\mathfrak{R}}(C, 38)$. So *Gym-shift* applies, because $38 < \delta \mathfrak{R}$, to assure that

$$D \oplus r \xrightarrow[2\delta]{\text{dc}} C \oplus 38.$$

The paragraph above (§) chose \mathfrak{R} big enough that $C \xrightarrow[\delta]{ca} A$. So *Same-Shift* grants us

$$C \oplus 38 \xrightarrow[\delta]{ca} A \oplus 38,$$

courtesy the (§) inequality. Now *Composition* yields

$$32: \quad D \oplus r \xrightarrow[\delta]{da} A \oplus 38.$$

Lastly, applying *Joint-Words* to (31) and (32) gives

$$\left[\begin{array}{c} D \\ D \oplus p \\ D \oplus q \\ D \oplus r \end{array} \right] \xrightarrow[\delta]{da} \left[\begin{array}{c} A \\ A \oplus 17 \\ A \oplus 29 \\ A \oplus 38 \end{array} \right]$$

as required. \blacklozenge

§E End-Notes

My friend and advisor Don Ornstein showed me the following particularly concise and elegant proof of the *Typicality Lemma*.

Proof of (13). We need but establish this for a specified $\delta := \varepsilon > 0$, since we could then painlessly intersect countably many full-mass sets as $\varepsilon = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$

The *Pointwise Ergodic Theorem* provides a length L_0 large enough that

$$\xi(Y \setminus G) < \frac{\delta}{3},$$

where G comprises all points $z \in Y$ such that:

33: For all integers $\ell \leq -L_0$ and $r \geq L_0$: The words $z[\ell..0)$ and $z[0..r)$ are each $[\varepsilon/2]$ -good for the S, Q process.

Enlarging L_0 so that $\frac{1/\varepsilon}{L_0}$ is negligible, we now have: If word $C \stackrel{\text{Len}}{\geq} L_0$ is $[\varepsilon/2]$ -good, likewise word C' , then their concatenation CC' is automatically ε -good.

Take $L := 3L_0$. Then for each length $m \geq L$, remark that

$$34: \quad \frac{m}{m - 2L_0} \leq 3.$$

The Idea. Ergodicity guarantees that a.e y hits G with asymptotic frequency $\xi(G)$. On such a point y , consider an L -block $y[\ell..r)$ which is ε -bad. Courtesy (33), the point $z := S^i(y)$ must be in $Y \setminus G$, for **each** value of index i in $[\ell+L_0..r-L_0)$.

So if $([\ell_i..r_i])_{i=1}^\infty$ is an ε -bad L -blocking of y , then necessarily

$$\overline{\text{Den}}\left(\bigcup_{i=1}^\infty [\ell_i+L_0..r_i-L_0)\right) \leq \xi(Y \setminus G),$$

from the ergodic theorem. By (34), then,

$$\overline{\text{Den}}\left(\bigcup_{i=1}^\infty [\ell_i..r_i)\right) < 3 \cdot \xi(Y \setminus G) < \delta. \quad \blacklozenge$$

Flat stacks. In the early 1980s, I started using the term (*asymptotically*) *flat stacks* for the notion defined in §A. However the concept goes back earlier at least to Katok and Stepin [5] (and perhaps also to Akcoglu and Chacón (?)).

Weak closure. The original proof [6] of the *Weak-Closure thm*, (1), used an involved coding argument. Valery Ryzhikov found, in [12], a shorter “joinings” argument.

§References/Bibliography

- [1] A. del Junco and D. Rudolph, *On ergodic actions whose self-joinings are graphs*, Ergodic Theory Dynamical Systems **7** (1987), no. 4, 531–557. MR 89e:28029
- [2] G. R. Goodson, *A survey of recent results in the spectral theory of ergodic dynamical systems*, J. Dynam. Control Systems **5** (1999), no. 2, 173–226. MR 2000f:28021
- [3] Paul R. Halmos, *Lectures on ergodic theory*, The Mathematical Society of Japan, 1956, Publications of the Mathematical Society of Japan, no. 3. MR 20 #3958
- [4] A. B. Katok, *Monotone equivalence in ergodic theory*, Izv. Akad. Nauk SSSR Ser. Mat. **41** (1977), no. 1, 104–157, 231. MR 56 #581
- [5] A. B. Katok and A. M. Stepin, *Approximation of ergodic dynamical systems by periodic transformations*, Dokl. Akad. Nauk SSSR **171** (1966), 1268–1271. MR 34 #7774a
- [6] Jonathan King, *The commutant is the weak closure of the powers, for rank-1 transformations*, Ergodic Theory Dynamical Systems **6** (1986), no. 3, 363–384. MR 88a:28021

- [7] Jonathan L. King, *Joining-rank and the structure of finite rank mixing transformations*, *J. Analyse Math.* **51** (1988), 182–227. MR89k:28009
- [8] ———, *Ergodic properties where order 4 implies infinite order*, *Israel J. Math.* **80** (1992), no. 1-2, 65–86. MR94i:28014
- [9] Jonathan L. F. King, *The generic transformation has roots of all orders*, *Colloq. Math.* **84/85** (2000), no. , part 2, 521–547, Dedicated to the memory of Anzelm Iwanik. MR1 784 212
- [10] M. Lemańczyk, F. Parreau, and J.-P. Thouvenot, *Gaussian automorphisms whose ergodic self-joinings are Gaussian*, *Fund. Math.* **164** (2000), no. 3, 253–293. MR2001h:37009
- [11] Daniel J. Rudolph, *An example of a measure preserving map with minimal self-joinings, and applications*, *J. Analyse Math.* **35** (1979), 97–122. MR81e:28011
- [12] V. V. Ryzhikov, *Mixing, rank and minimal self-joining of actions with invariant measure*, *Mat. Sb.* **183** (1992), no. 3, 133–160. MR93m:28020
- [13] J.-P. Thouvenot, *Some properties and applications of joinings in ergodic theory*, *Ergodic theory and its connections with harmonic analysis* (Alexandria, 1993), Cambridge Univ. Press, Cambridge, 1995, pp. 207–235. MR96d:28017

Filename: Article/24Flatstacks/flatstacks.latex

As of: Tuesday 05Jun2001. Typeset: 29Jul2001 at 23:06.