ISRAEL JOURNAL OF MATHEMATICS 84 (1993), 129–145

THE CENTRALIZER OF A RANK-ONE FLOW

BY

PAUL ZEITZ

Department of Mathematics
University of San Francisco, San Francisco, CA 94117-1080, USA

ABSTRACT

In [5], King proved that the centralizer of a rank-1 transformation equals the "weak closure" of its (positive and negative) powers (see below for a definition of the weak topology). We define rank-1 flows, and then show that simple modifications of King's proof yield an analogous statement for rank-1 flows.

1. Introduction

For all that follows, let $(X, \mathcal{X}, \mu)$ be a Lebesgue probability space. Let $\{T_i\}$ be a sequence of m.p.t.'s (measure-preserving transformations) on X. We say that $T_i \to S$ **weakly** if, for all measurable sets A, $\mu(T_i^{-1}(A) \triangle S^{-1}(A)) \to 0$ as $i \to \infty$. For T a m.p.t. on X, let wcl(T) denote the weak closure of $\{T^n : n \in \mathbf{Z}\}$. Let $C(T)$ denote the **centralizer** of T, i.e., the set of m.p.t.'s which commute with T. King's weak closure theorem [5] states that $C(T) = \text{wcl}(T)$, for rank-1 T.

Note that often wcl(T) is a much larger set than $\{T^n : n \in \mathbf{Z}\}$. For example, the irrational rotation $R_\alpha : [0, 1) \to [0, 1)$ defined by $x \mapsto x + \alpha \pmod 1$, α irrational is rank-1 [2], and it is easy to see that wcl$(R_\alpha) = \{R_\beta : \beta \in [0, 1)\}$. Hence the only maps which commute with a given irrational rotation are all rotations of the circle.

This example has a flow analogue. By a **flow** ϕ we mean a group $\{\phi^t : t \in \mathbf{R}\}$ of measure-preserving transformations on X with $\phi^0 = \text{id}$ (identity on X) and satisfying $\phi^t \phi^s = \phi^{t+s}$ for all $t, s \in \mathbf{R}$. Furthermore, we require that the flow

Received October 14, 1990 and in revised form April 15, 1992

be measurable, i.e. for all $A \in \mathcal{X}$, $\{(x,t) : \phi^t x \in A\}$ is a measurable set in $(X \times \mathbf{R}, \mu \times m)$, where m denotes Lebesque measure on $\mathbf{R}$.

For any $t \in \mathbf{R}$, let $\phi^t : [0,1) \times [0,1) \to [0,1) \times [0,1)$ be defined by $(x,y) \mapsto (x+t, y+t\alpha)$, where α is irrational. Then ϕ is the well-known irrational slope flow on the torus. Again, it is easy to see that the weak closure of $\{\phi^t : t \in \mathbf{R}\}$ contains all translations of the form $(x,y) \mapsto (x+u, y+v)$ where u, v are constants. These translations are not contained in the group $\{\phi^t : t \in \mathbf{R}\}$, yet certainly commute with each group element.

For any flow ϕ, let wcl(ϕ) denote the weak closure of $\{\phi^t : t \in \mathbf{R}\}$ and let $C(\phi)$ denote the set of m.p.t.'s which commute with all group elements $\{\phi^t : t \in \mathbf{R}\}$. The above example suggests that for certain types of flows, wcl(ϕ) may equal $C(\phi)$. In this paper we extend the notion of rank-1 to flows, and show that for such flows, this equality holds.

In the remainder of this section we summarize useful facts about flows, define rank-1 flows, and give some of their properties. In section 2 we extend the notion of coding to flows, allowing us to emulate King's proof in section 3. We place the proof of lemma 1.2 in the appendix, as it is technical and unrelated to the main coding arguments.

A flow is **ergodic** if all measurable flow-invariant sets are trivial, i.e. if A is measurable and $\phi^t A = A$ for all t, then $\mu(A) = 0$ or 1. The entropy of a flow is defined by the formula $h(\phi) = h(\phi^1)$, where "h" in the right-hand side denotes the entropy of a transformation.

LEMMA 1.1: *Let ϕ be a flow. Then*

(a) *For any measurable set A, $\mu(\phi^t A \triangle A) \to 0$ as $t \to 0$.*
(b) *For any real t, $h(\phi^t) = |t| h(\phi)$.*
(c) *If ϕ is an ergodic flow, then there is a t_0 such that ϕ^{t_0} is an ergodic transformation.*
(d) *If $h(\phi)$ is finite, there exists a real number t_0 and a finite partition $\mathcal{P}$ such that $\bigvee_{n=-\infty}^{\infty} \phi^{nt_0}\mathcal{P} = \mathcal{X}$.*

Proof: The measurability of the flow implies (a). See [4, pp. 255,326] for proofs of (b) and (c). Statement (d) follows from (c) and Krieger's theorem [8]. We call the partition $\mathcal{P}$ a **generating** partition for ϕ.

Let $\mathcal{P}$ be a partition of X with k atoms. Each point $x \in X$ has a P-**name**, a function $x : \mathbf{R} \to \{1, \ldots, k\}$, where $x(t) = j$ iff $\phi^t x$ lies in the jth atom of $\mathcal{P}$.

We will often think of a $\mathcal{P}$-name as a measurable "coloring" of $\mathbf{R}$, where each of k colors corresponds to an atom of $\mathcal{P}$. A **word** W of length h is a finite $\mathcal{P}$-name with domain $[0, h)$. Let len$(\cdot)$ denote the length of a word. To indicate substrings of names, the symbol $x|_a^b$ means the restriction of the name to $[a, b)$.

For two words V, W of length h, define the $\overline{d}$-distance to be

$$\overline{d}(V, W) = \frac{1}{h} m(\{t \in [0, h) : V(t) \neq W(t)\}).$$

Similarly, for any $x, y \in X$ define $\overline{d}(x, y) = \limsup_{t\to\infty} \overline{d}(x|_0^t, y|_0^t)$. Note that $\overline{d}$ satisfies the triangle inequality. A word W' is called a **$\overline{d}$-ε copy** of a word W if the two words have the same length and $\overline{d}(W, W') < \varepsilon$.

A $\mathcal{P}$-name $x|_0^\infty$ is said to be **covered with density** β by $\overline{d}$-ε copies of W, if there is a disjoint sequence of intervals $[a_i, a_i + h)$, $0 \leq a_1 < a_2 < \cdots$, where each $x|_{a_i}^{a_i+h}$ is a $\overline{d}$-ε copy and $\liminf_{n\to\infty} nh/(a_n + h) = \beta$.

A flow ϕ is **rank-one** if for any partition $\mathcal{P}$, and any positive ε, H, there is a single word W with length h greater than H such that for almost every $x \in X$, the $\mathcal{P}$-name of x can be covered, with density greater than $(1 - \varepsilon)$, by $\overline{d}$-ε copies of W. We will say that W **ε-covers** and will call it a **covering word.** We will refer to $\overline{d}$-ε copies of W as **ε-nearcopies** of W, and often denote such nearcopies by W', W'', etc. When there is no ambiguity, we will drop the ε. We will always denote the length of a nearcopy by h, and if a substring of a nearcopy has length s, we let $s\%$ denote the value s/h.

Note that this definition is a "natural" flow analogue of the $\overline{d}$-definition of rank-1 transformations. (See [7] for several equivalent definitions.)

It is easy to see that the irrational slope flow mentioned above is a rank-1 flow; this follows via the same argument used to show that irrational rotations of the circle are rank-1 [2]. There is another, equivalent definition for rank-1 flows which uses a flow version of Rokhlin towers. This idea is presented in the appendix. Also, rank-1 flows can be constructed via "cutting and stacking". An example is a weak-mixing rank-1 flow $\theta = \phi_1 \times \phi_2$, where each ϕ_i is weak-mixing. This can be constructed by adapting example (vi) of [5] to flows. Here $C(\theta)$ contains both $\mathrm{id}_1 \times \phi_2$ and $\phi_1 \times \mathrm{id}_2$, so is rather more complex than $\{\theta^t : t \in \mathbf{R}\}$.

The following facts about rank-1 flows will be needed for the proof of the weak closure theorem. Proofs may be found in the appendix.

LEMMA 1.2: *Let ϕ be a rank-1 flow. Then*

(a) ϕ *is ergodic.*
(b) $h(\phi) = 0$.

2. Coding

Let ϕ be a flow which has a generating partition $\mathcal{P}$, as described in lemma 1.1. Let $S \in C(\phi)$ and fix $\varepsilon > 0$. Define $\mathcal{Q} = S^{-1}(\mathcal{P})$. Since $\mathcal{P}$ generates, there is a real number t_0 and a natural number n such that

$$\mathcal{Q} \overset{\varepsilon}{\subset} \bigvee_{i=-n}^{n} \phi^{it_0}(\mathcal{P}).$$

By this we mean the following: If we write $\mathcal{Q}$ as an ordered partition $\{Q_1, \ldots, Q_K\}$, there exists another ordered partition $\mathcal{Q}' = \{Q'_1, \ldots, Q'_K\}$ with $|\mathcal{Q}, \mathcal{Q}'| \overset{\text{def}}{=} \sum_{j=1}^{K} \mu(Q_j \Delta Q'_j) < \varepsilon$ and

$$\mathcal{Q}' \subset \bigvee_{i=-n}^{n} \phi^{it_0}(\mathcal{P}). \tag{1}$$

Let $x \in X$. The $(2n+1)$-tuple of $\mathcal{P}$-colors $(x(-nt_0), x(-(n-1)t_0), \ldots, x(nt_0))$ determines a $\mathcal{Q}'$ atom by equation (1). More precisely, (1) gives us a map $f : \{1, \ldots, K\}^{2n+1} \to \{1, \ldots, K\}$. Thus we can define a **code C** which takes $\mathcal{P}$-names to $\mathcal{P}$-names by

$$(\mathbf{C}x)(t) = f(x(t - nt_0), \ldots, x(t + nt_0)). \tag{2}$$

Since $S \in C(\phi)$, the ergodic theorem for flows implies that for almost all $x \in X$, $\overline{d}(\mathbf{C}x, Sx) < \varepsilon$. Define the **codelength** lenC and the **error rate** errC of the above code to be $2n+1$ and $2 \cdot |\mathcal{Q}, \mathcal{Q}'|$, respectively.

We can also code the substring $x|_a^b$, using (2). The only difficulty is that this equation is meaningless for $t < a + nt_0$ or $t \geq b - nt_0$. For these values we just let $(\mathbf{C}x)(t) = x(t)$. Hence $\mathbf{C}(x|_a^b)$ has the same length as $x|_a^b$. To absorb "end effects", we always assume that $b - a$ is much greater than $t_0 \cdot$ lenC. We say a substring $x|_a^b$ **codes well** (under **C**) if $\overline{d}(\mathbf{C}(x|_a^b), (Sx)|_a^b) <$ errC. We defined errC to be $2 \cdot |\mathcal{Q}, \mathcal{Q}'|$ rather than $|\mathcal{Q}, \mathcal{Q}'|$ so as to allow

LEMMA 2.1: *For any code* **C***, and any postive* δ*, there exists* H *such that: For almost every* $x \in X$*, if we cover at least* $(1 - \delta)$ *of* $x|_0^\infty$ *with disjoint substrings* *with lengths exceeding H, then at least $(1-2\delta)$ of $x|_0^\infty$ is covered by those, of the above substrings, which code well under* $\mathbf{C}$.

See [6] for a proof of the analogous statement for transformations. This proof carries over to the case of flows with no difficulty. The next lemma is a simple consequence of the way the code is constructed:

LEMMA 2.2: *For any two words W, W' of equal length,*

$$\bar{d}(\mathbf{C}W, \mathbf{C}W') \leq (\text{len}\mathbf{C}) \cdot \bar{d}(W, W').$$

3. The weak closure theorem

This section is devoted to proving

THEOREM 1: *For a rank-one flow ϕ, $C(\phi) = \text{wcl}(\phi)$.*

Our proof follows King's closely, employing similar coding arguments. (Actually, we will only prove that $C(\phi) \subset \text{wcl}(\phi)$, since the converse is immediate.) We shall do the proof in several steps. First, we use commutativity, genericity, and the ergodic theorem to restrict our analysis to a single $\mathcal{P}$-name. Then we show that if the theorem were false, this $\mathcal{P}$-name exhibits two forms of "periodicity". Finally, we use this periodicity to produce a contradiction, by showing that our candidate for the centralizer can be approximated by simple shifts of the form ϕ^{-r}.

3.1 Reducing to a single $\mathcal{P}$-name.

Fix $S \in C(\phi)$. By lemmas 1.2b and 1.1d, we are guaranteed the existence of a generating partition $\mathcal{P}$, i.e., for some real t_0, the σ-algebra generated by $\{\phi^{nt_0}(\mathcal{P}) : n \in \mathbf{Z}\}$ equals $\mathcal{X}$ (up to nullsets). Let $\mathcal{C}$ be a countable sequence of codes with error rates going to zero. For $T = \phi^{t_0}$, let $\mathcal{A}$ be the countable algebra of $T, \mathcal{P}$-cylinder sets. By the ergodic theorem for flows and the commuting of S with T, there exists $x, y \in X$ such that $y = Sx$, with both points generic for all sets in the countable collection $\{\phi^t A \Delta S^{-1} A : t \in \mathbf{Q}, A \in \mathcal{A}\}$, and such that lemma 2.1 holds for both points for each code in $\mathcal{C}$. We will keep $\mathcal{P}, T, x, y$ fixed for the remainder of section 3. The following lemma converts statements about weak convergence into statements about $\mathcal{P}$-names.

LEMMA 3.1: *For any sequence $\{t_i\}$ in* $\mathbf{R}$, *if* $\bar{d}(\phi^{t_i}x, y) \to 0$ *then* $\phi^{t_i} \xrightarrow{weak} S$.

Proof: By lemma 1.1, we can assume that each $t_i \in \mathbf{Q}$. Let $A \in \mathcal{A}$ be a cylinder set contained in $\bigvee_{j=-l}^{l} T^j \mathcal{P}$. Then A has a *discrete* $T, \mathcal{P}$-name $A_{-l} \ldots A_l$. (The A_j-th atom of $\mathcal{P}$ contains $T^{-j}A$.) Define

$$\operatorname{freq}(C) = \lim_{T \to \infty} \frac{1}{T} \int_0^T 1_C(t)\, dm(t),$$

where $C \subset [0, \infty)$, 1_C is the indicator function for the set C, m denotes Lebesque measure on $\mathbf{R}$, and C is such that the limit actually exists. (In our context, the ergodic theorem will guarantee this.) Abbreviate $\phi^{t_i}x$ by z. By genericity and the ergodic theorem, we have

$$\begin{aligned}
\mu(\phi^{-t_i} A \Delta S^{-1} A) &= \operatorname{freq}\{t \geq 0 : \phi^t x \in \phi^{-t_i} A \text{ XOR } \phi^t x \in S^{-1} A\} \\
&= \operatorname{freq}\{t \geq 0 : \phi^t z \in A \text{ XOR } \phi^t y \in A\} \\
&= \operatorname{freq}\{t \geq 0 : z(t + jt_0) = A_j \text{ for all } j \in \{-l, \ldots, l\} \text{ XOR} \\
&\qquad y(t + jt_0) = A_j \text{ for all } j \in \{-l, \ldots, l\}\} \\
&\leq \operatorname{freq}\{t \geq 0 : z(t + jt_0) \neq y(t + jt_0) \\
&\qquad \text{for at least one } j \in \{-l, \ldots, l\}\} \\
&\leq (2l + 1)\bar{d}(z, y).
\end{aligned}$$

Letting $i \to \infty$ establishes the lemma. The converse can be proven in a similar way, but is not needed for what follows. ∎

3.2 THE PERIODICITY EXTENSION LEMMA.

From now on we assume that Theorem 1 is false, i.e., assume that there is an S in $C(\phi)$ which is not contained in $\operatorname{wcl}(\phi)$. As T is rank-1, there is a sequence of covering words $\{W_n\}$, each which ε_n-covers, with $\varepsilon_n \searrow 0$ and $\operatorname{len}(W_n) \nearrow \infty$. The phrase "sufficiently accurate cover" below means that we choose a covering word of sufficiently large index n from this sequence.

Informally, the next lemma tells us that if the action of S on nearcopies is "close" in some way to a shift ϕ^{-s}, then s cannot be an arbitrarily small fraction of the length of the covering word. The lemma is almost identical to King's discrete version (lemma 1.4 of [5]), and King's proof carries over with only notational changes.

LEMMA 3.2 (Shift lemma): *There exist positive constants α, σ such that for any code* $\mathbf{C}$ *with* $\mathrm{err}\mathbf{C}$ *less than α, and every sufficiently accurate covering word W of length h, if there is a shift s with $0 < s < h$ such that*

$$\bar{d}(\mathbf{C}W|_s^h, W|_0^{h-s}) < \alpha,$$

then $s\% > 2\sigma$.

Remark: This lemma is also true for negative shifts s, in which case the hypothesis is $\bar{d}(\mathbf{C}W|_0^{h+s}, W|_{-s}^h) < \alpha$, and the conclusion is $|s\%| > 2\sigma$. ∎

Now fix α, σ, and henceforth, we shall assume that all codes are members from $\mathcal{C}$ with error rates less than α. For a word of length h, and $p \in \mathbf{R}$, let $W \oplus p$ be the word of length h defined by $(W \oplus p)(t) = W([t+p] \bmod h)$, $t \in [0, h)$. The word $W \ominus p$ is defined in a similar way. When $\bar{d}(W, W \oplus p)$ is small for some p, we will informally call this **periodicity**. In the sequel we will often use the simple but important fact that for any words V, W of the same length and any real p, $\bar{d}(V, W) = \bar{d}(V \oplus p, W \oplus p)$.

The following lemma tells us when the periodicity on a sufficiently long piece of a word can be extended to the entire word.

LEMMA 3.3 (Periodicity extension lemma): *Given ε, there exist ε', H and codes $\mathbf{C}_1, \dots, \mathbf{C}_N$ (where $N = \lfloor 1/\sigma \rfloor + 1$) such that if there exist real p and r with $0 \le r < h$ and $r\% > \sigma$, as well as a word W of length $h > H$ satisfying:*

(i) $\bar{d}(W|_0^r, (W \oplus p)|_0^r) < \varepsilon'$,

(ii) *for each i,* $\bar{d}(\mathbf{C}_i W, W \oplus r) < 2\mathrm{err}\mathbf{C}_i$,

then $\bar{d}(W, W \oplus p) < \varepsilon/2$.

Proof: Let ε_0 be positive, to be determined later. For $i = 1, \dots, N$, pick a positive $e_i < \varepsilon_{i-1}/2$ and a code $\mathbf{C}_i$ with $\mathrm{err}\mathbf{C}_i < \sigma e_i$. Letting $k_i = \mathrm{len}\mathbf{C}_i$, pick ε_i so small that $k_i\varepsilon_i + 4e_i < \varepsilon_{i-1}$. Let $\varepsilon' = \varepsilon_N$.

Take H much larger than $\max_i k_i$, so that we may ignore the "end effects" of the codes. Using (i), we have

$$\bar{d}(\mathbf{C}_N(W|_0^r), \mathbf{C}_N(W \oplus p|_0^r)) \le k_N \bar{d}(W|_0^r, W \oplus p|_0^r) < k_N \varepsilon_N.$$

From (ii) we get

$$\bar{d}(W|_r^{2r}, \mathbf{C}_N(W|_0^r)) \le 2\mathrm{err}\mathbf{C}_N / r\% < 2e_N$$

and

$$\overline{d}(W \oplus p|_r^{2r}, \mathbf{C}_N(W \oplus p|_0^r)) < 2e_N.$$

The triangle inequality yields

$$\overline{d}(W|_r^{2r}, W \oplus p|_r^{2r}) < k_N \varepsilon_N + 4e_N < \varepsilon_{N-1}.$$

Continuing in this manner, using codes $\mathbf{C}_{N-1}$, $\mathbf{C}_{N-2}$, etc., we can show, for $i = 1, \ldots, M = \lfloor h/r \rfloor$ that

$$\overline{d}(W|_{ir}^{(i+1)r}, W \oplus p|_{ir}^{(i+1)r}) < \varepsilon_{N-i} \leq \varepsilon_0.$$

If $h - Mr = 0$ we are done, by setting $\varepsilon_0 = \varepsilon/2$. Otherwise, we continue the above analysis using code $\mathbf{C}_{N-M-1}$, yielding

$$\overline{d}(W \oplus Mr|_0^r, W \oplus p \oplus Mr|_0^r) < \varepsilon_{N-M-1} \leq \varepsilon_0.$$

This means that the error on the "leftover" substring of length $h - Mr$ can be controlled, for the above inequality implies

$$(h - Mr)\overline{d}(W|_{Mr}^h, W \oplus p|_{Mr}^h) < \varepsilon_0 r.$$

Thus

$$\overline{d}(W, W \oplus p) < \varepsilon_0 \cdot (Mr)\% + \frac{\varepsilon_0 r}{(h - Mr)} \cdot (h - Mr)\% \leq 2\varepsilon_0.$$

The lemma is established, then, if we set $\varepsilon_0 = \varepsilon/4$. ∎

Remark: In hypothesis (ii) of this lemma, r can be replaced with $-r$. ∎

3.3 Inducing "periodicity".

Fix positive $\varepsilon < \sigma$. The periodicity extension lemma produces $N = \lfloor 1/\sigma \rfloor + 1$ codes and ε'. Now pick a code $\mathbf{C}$ with an even smaller error rate e with $6e/\sigma < \varepsilon'$. Pick positive γ less than ε satisfying $\gamma \cdot (\text{len}\mathbf{C})/\sigma + \gamma < e$. Next choose W, an η-covering word of length h, of "sufficient accuracy" (in the sense defined above) so that h is large enough to make the periodicity extension lemma hold, and η is small enough so that

$$2(\text{len}\mathbf{C})\% + \eta \cdot (\text{len}\mathbf{C})/\sigma + \eta < e.$$

We will add further constraints to η and h below.

Let $W' = x|_a^{a+h}$ be a nearcopy on $x|_0^\infty$. Let $W'' = x|_b^{b+h}$ be the next nearcopy, called the **successor**, with a **gap** of $g = b - a - h$. We call $x|_a^{a+h}$ a **good** nearcopy if: $g\% < \gamma$; there is a nearcopy "lying below" on y, i.e. $W''' = y|_c^{c+h}$ is a nearcopy with $a \le c \le a + h$; and each of the substrings $x|_a^{a+h}$, $x|_b^{b+h}$, and $x|_c^{c+h}$ code well under the $N+1$ codes. Call $s = c - a$ the **shift** associated with the good nearcopy. (See Figure 1.)

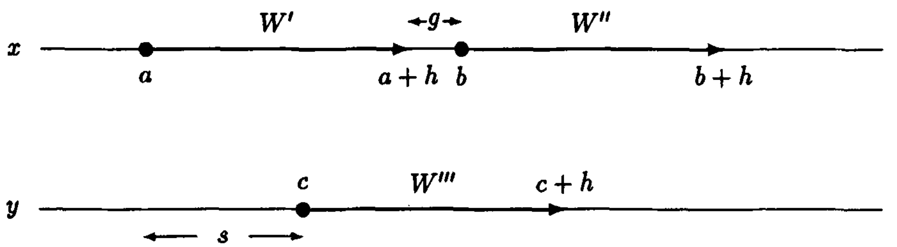

Figure 1: $x|_a^{a+h}$ is a good nearcopy. Nearcopies are shown by thick "arrows", with the large dot indicating the start and the arrowhead indicating the end of the word.

By a Fubini argument, assume that we also choose this cover sufficiently accurately so that at least $(1 - \varepsilon)$ of $x|_0^\infty$ is covered by good nearcopies. (We must make sure that h is large enough so that the hypotheses for lemma 2.1 are satisfied for all $N + 1$ codes.)

Since the good nearcopy codes well, we have

$$\bar{d}(\mathbf{C}W|_s^h, W|_0^{h-s}) \le \bar{d}(\mathbf{C}W|_s^h, \mathbf{C}W'|_s^h) + \bar{d}(\mathbf{C}W'|_s^h, W'''|_0^{h-s}) + \bar{d}(W'''|_0^{h-s}, W|_0^{h-s})$$

$$\le (\eta \cdot \mathrm{len}\mathbf{C} + e + \eta)/(1 - s\%).$$

If η is small enough, the shift lemma yields $2\sigma < s\%$. Since the successor also codes well, a similar argument yields $1 - s\% + g\% > 2\sigma$. Combining this with $g\% < \varepsilon < \sigma$ gives us

$$\sigma < s\% < 1 - \sigma.$$

Next, we assert the existence of a special good nearcopy.

LEMMA 3.4: *There exists a good nearcopy, with shift denoted by r, with no gap between it and the next nearcopy. For all other shifts s associated with other good nearcopies, either $r\% \le s\%$ and $r\% \le 1 - s\%$ or $1 - r\% \le s\%$ and $1 - r\% \le 1 - s\%$.*

We shall defer the proof. Call this special good nearcopy the **reference nearcopy**. Note that $\sigma < r\% < 1-\sigma$.

The existence of the reference nearcopy imposes a periodicity on the covering word. Call the reference nearcopy, its successor nearcopy, and the nearcopy lying below W', W'' and W''' respectively. Since the substring of x "lying above" W''' codes well, we have $\bar{d}(CV, W''') < e$, where V denotes the concatenation $W'|_r^h W''|_0^r$. (See Figure 2.) Since V consists of parts of nearcopies with lengths no shorter than the minimum of r and $h-r$, and $\min(r\%, 1-r\%) > \sigma$, we have

$$\bar{d}(W \oplus r, V) < \eta/\sigma$$

which implies

$$\bar{d}(\mathbf{C}(W \oplus r), \mathbf{C}V) < \eta \cdot (\mathrm{len}\mathbf{C})/\sigma.$$

The triangle inequality yields

$$\bar{d}(\mathbf{C}(W \oplus r), W) \le \bar{d}(\mathbf{C}(W \oplus r), \mathbf{C}V) + \bar{d}(\mathbf{C}V, W''') + \bar{d}(W''', W)$$

$$< \eta \cdot (\mathrm{len}\mathbf{C})/\sigma + e + \eta.$$

Since $\bar{d}(\mathbf{C}W, W \ominus r) \le \bar{d}(\mathbf{C}(W \oplus r), W) +$ "end effects", we have

(1) $$\bar{d}(\mathbf{C}W, W \ominus r) < 2(\mathrm{len}\mathbf{C})\% + \eta \cdot (\mathrm{len}\mathbf{C})/\sigma + e + \eta < 2e.$$

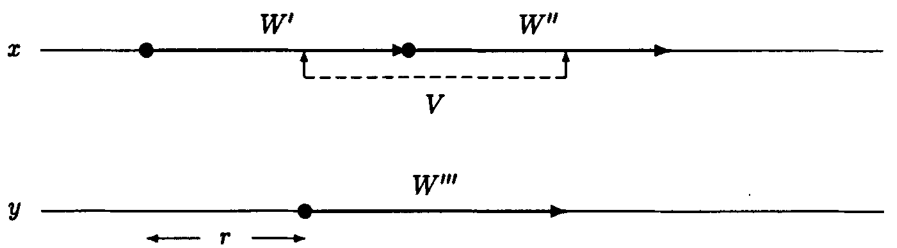

Figure 2: W' is the reference copy with shift r. V denotes the concatenation $W'|_r^h W''|_0^r$.

Next, we will use equation (1) to show that the other good nearcopies impose different periodicity properties on the covering word.

LEMMA 3.5: *Let W' be a good nearcopy with associated gap g. Then*

$$\bar{d}(W, W \oplus g) < \varepsilon.$$

Proof: Let the associated shift be s, and suppose that $r\% \leq s\%$ and $r\% \leq 1-s\%$. A similar argument will handle the case $1 - r\% < s\%$ and $1 - r\% < 1 - s\%$. Use the notation of Figure 1. Consider the substring $x|_c^{c+h}$ "lying above" W'''. Replace each nearcopy with the covering word W. Now $W|_s^h$ "lies above" $W|_0^{h-s}$ and $W|_0^{s-g}$ "lies above" $W|_{h-s+g}^h$. Changing the order of these two pieces of W produces the situation depicted in Figure 3. Since $x|_c^{c+h}$ codes well, the same triangle inequality argument used to derive equation (1) shows that the coded version of the top agrees with $\bar{d}$-$2e$ accuracy with the bottom. Equation (1) and the triangle inequality yield Figure 4, where the top and bottom are $\bar{d}$-$4e$ close.

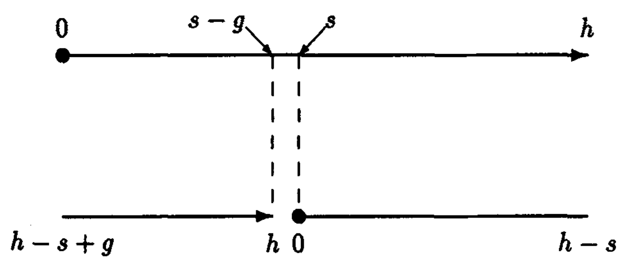

Figure 3: The covering word W is depicted by an "arrow", with the large dot indicating the start and the arrowhead indicating the end of the word. The coded version of the top is $\bar{d}$-$2e$ close to the bottom.

Thus, there are two different regions of periodicity on W. Let $p = r - s + g$. Then we have

$$\bar{d}(W \ominus r|_0^r, W \ominus r \oplus p|_0^r) < 4e/r\% < 4e/\sigma < \varepsilon'.$$

Denote $W \ominus r$ by V. We have

$$\bar{d}(V|_0^r, V \oplus p|_0^r) < \varepsilon'.$$

By the same analysis that produced equation (1) we have

$$\bar{d}(\mathbf{C}_i W, W \ominus r) < 2\mathrm{err}\mathbf{C}_i \quad \text{for } i = 1, \ldots, N.$$

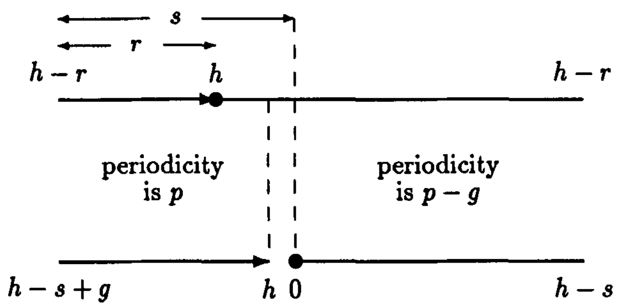

Figure 4: This combines equation (1) with the previous figure. The top and bottom are $\bar{d}$-4ϵ close.

Since "end effects" are negligible (by our choice of h), this yields

$$\bar{d}(\mathbf{C}_i V, V \ominus r) < 2\mathrm{err}\mathbf{C}_i \quad \text{for } i = 1, \ldots, N.$$

The periodicity extension lemma then implies $\bar{d}(V, V \oplus p) = \bar{d}(W, W \oplus p) < \varepsilon/2$. Likewise, using the other region of periodicity, we have

$$\bar{d}(W|_s^{s+r}, W \oplus (p-g)|_s^{s+r}) < 4\epsilon/r\% < \varepsilon',$$

so a similar argument using the periodicity extension lemma yields $\bar{d}(W, W \oplus (p-g)) < \varepsilon/2$. By the triangle inequality, $\bar{d}(W \oplus p, W \oplus (p-g)) < \varepsilon$. The lemma follows, as $\bar{d}(W \oplus p, W \oplus (p-g)) = \bar{d}(W \oplus g, W)$. ∎

It remains to prove lemma 3.4. Let $\rho = \inf(\min(u\%, 1 - u\%))$ as u ranges over all shift values found for good nearcopies on $x|_0^\infty$. If this infimum is attained, simply let the reference nearcopy be that nearcopy with shift r satisfying $\min(r\%, 1 - r\%) = \rho$. Otherwise, there is a positive β such that $(1 - 2\varepsilon)$ of $x|_0^\infty$ is covered by good nearcopies satisfying $\min(s\%, 1 - s\%) \geq \rho + \beta$ for each shift s. Now remove from consideration ("uncover") all but one of the good nearcopies that do not satisfy this condition. Denote this surviving good nearcopy by $x|_a^{a+h}$ and let the shift and gap be r and g respectively. Because we have removed some of the original good nearcopies, r satisfies the required inequalities of lemma 3.4.

However, the gap g may be positive. To solve this problem, we redefine our covering word. If $x|_a^{a+h} = W'$, let the new covering word be the concatenation $V = Wx|_{a+h}^{a+h+g}$ of length $h + g$. We have succeeded in creating the proper reference nearcopy, but now we need to check if lemma 3.5 still holds, because

V is a less accurate covering word and we may have introduced **negative** gap values on some of the other good nearcopies by enlarging W to V.

The new word V is an $(\eta+\gamma)$-cover. Since V is longer than W, the "end effects" are smaller, and since we chose γ so that $\gamma \cdot (\text{lenC})/\sigma + \gamma < e$, the right-hand side of equation (1) will now be $3e$. Likewise, in Figure 4, the top and bottom will now be $\bar{d}$-$6e$ close. Since $6e/\sigma < \varepsilon'$, we can still use the periodicity extension lemma to prove lemma 3.5.

The complexity introduced by negative gap values is only notational, since $|g\%| < \gamma$. Figure 4 would be altered to show an overlap where there is presently a gap, but there would still be two different regions of periodicity, of sufficient length to allow the periodicity extension arguments to work.

Finally, the fact that the newly defined good nearcopies only cover $(1-2\varepsilon)$ of $x|_0^\infty$ (rather than $(1-\varepsilon)$) merely adds one more epsilon to the right hand side of (2) below, which will not affect the argument.

3.4 Using periodicities to arrive at a contradiction.

We will show that S is $o(\varepsilon)$-close, in the weak topology, to ϕ^{-r}. Let us construct a $\mathcal{P}$-name $z|_0^\infty$ as follows: If $x|_a^{a+h}$ is a nearcopy W', let $z|_a^{a+h} = W \ominus r$. Otherwise, let $z(t) = (\mathbf{C}x)(t)$. Thus we have

$$\bar{d}(z|_a^{a+h}, (\mathbf{C}x)|_a^{a+h}) \leq \bar{d}(W \ominus r, \mathbf{C}W')$$

$$\leq \bar{d}(W \ominus r, \mathbf{C}W) + \bar{d}(\mathbf{C}W, \mathbf{C}W') < 2e + \eta \cdot (\text{lenC}) < 3e.$$

Thus $\bar{d}(z, \mathbf{C}x) < 3e$. The triangle inequality yields $\bar{d}(z, y) < 4e < 4\varepsilon' < 4\varepsilon$. Let $x|_a^{a+h}$ be a nearcopy with gap g (possibly not a good nearcopy). If it is a good nearcopy, by the preceding lemma we have $\bar{d}(W, W \oplus g) < \varepsilon$, so

$$\bar{d}(W|_{h-r}^{h}, z|_{a+h}^{a+h+r}) < (\varepsilon + g\%)/r\% < 2\varepsilon/\sigma.$$

Now construct another name $w|_0^\infty$ where $w|_{a+h}^{a+h+r} = W|_{h-r}^{h}$ for each good nearcopy $x|_a^{a+h}$, otherwise, $w(t) = z(t)$. Then we will have

$$\bar{d}(w|_{a+h}^{a+h+r}, z|_{a+h}^{a+h+r}) < \bar{d}(W|_{h-r}^{h}, z|_{a+h}^{a+h+r}) < 2\varepsilon/\sigma,$$

so $\bar{d}(z, w) < 2\varepsilon/\sigma$. Note that when $x|_a^{a+h}$ is good, it equals $w|_{a+r}^{a+r+h}$. In other words, $w|_0^\infty$ looks just like $\phi^{-r}x$, except that w contains exact copies of W, while $\phi^{-r}x$ contains nearcopies, and there is no match when the nearcopies are not good. More precisely, we have $\bar{d}(\phi^{-r}x, w) < \eta + \varepsilon$. Several more applications of the triangle inequality produce

$$\bar{d}(\phi^{-r}x, y) < 5\varepsilon + 2\varepsilon/\sigma + \eta < 8\varepsilon/\sigma. \tag{2}$$

Since ε can be made arbitrarily small, (and r depends on ε) we have $\inf\{\bar{d}(\phi^t x, y) : t \in \mathbf{R}\} = 0$, which contradicts $S \notin \mathrm{wcl}(\phi)$. This establishes the theorem. ∎

4. Appendix

Our goal is to prove lemma 1.2:

LEMMA: *Let ϕ be a rank-1 flow. Then*
(a) *ϕ is ergodic.*
(b) $h(\phi) = 0$.

Proof of (a): Let A be a flow-invariant set of positive measure. Let $\mathcal{P}$ be the partition $\{A, A^c\}$, and fix $0 < \varepsilon < 1/2$. Since the flow is rank-one, there is a $\mathcal{P}$-word W which ε-covers. Pick $x \in A$ such that x is covered by nearcopies of W. W must be an at least $(1-\varepsilon)$ "monochromatic" word, the same color as A, since A is flow-invariant. If A^c had positive measure, we could find an $x \in A^c$ which is ε-covered by W, an impossibility, as all $\mathcal{P}$-names in A^c are monochromatic, colored differently than W. Hence A has full measure. ∎

In order to calculate the entropy of a rank-1 flow, we require the concept of a flow under a function. By a **flow under a function** (f, E, T, ν) we mean the following: Let $(E, \mathcal{E}, \nu)$ be a finite measure space; $f : E \to \mathbf{R}$ an integrable function, bounded above zero, satisfying $\nu(E)\int f dm = 1$; $T : E \to E$ an ergodic, invertible, measure-preserving transformation. Let $Y = \{(e, s) \in E \times \mathbf{R} : 0 \le s < f(e)\}$, with measure space structure inherited from $(E \times \mathbf{R}, \nu \times m)$. We define (f, E, T, ν) to be the flow on Y which takes a point (e, s) vertically upward at unit speed until it reaches the graph of f, then to $(T(e), 0)$, where it continues upward again, etc. Ambrose [1] showed that every ergodic flow is isomorphic to a flow under a function. Rokhlin's theorem can be easily generalized to flows, using Ambrose's representation. A proof may be found in [9].

LEMMA 4.1: *Let ϕ be an ergodic flow. For any positive H and ε, we can represent ϕ as a flow under a function (f, E, T, ν) where $B = \{b \in E : f(b) = H\}$ has ν-measure greater than $(1-\varepsilon)$ and $f(b) < H$ for $b \notin B$.*

We refer to the set $B \times [0, H]$ as an (H, ε)-*tower* with **base** B.

LEMMA 4.2: *Let ϕ be a rank-one flow. Then for every partition $\mathcal{P}$, given positive ε and h_0 there exists an (h, ε)-tower with $h \geq h_0$ and such that at least $(1 - \varepsilon)$ of the horizontal rows are at least $(1 - \varepsilon)$ monochromatic.*

Proof: Let W be an η-covering word of length $h \geq h_0$, where $\eta > 0$ will be determined later. Because rank-one flows are ergodic, we can form a (H, η)-tower with base B, with H chosen (by a Fubini argument) so large that for all $x \in B$ outside a set of measure less than η, we can η-cover $x|_0^H$ by nearcopies of W with density greater than $(1 - \eta)$. Define $f_1 : B \to \mathbf{R}$ by $f_1(x) = \inf\{t \geq 0 : \bar{d}(x|_t^{t+h}, W) < \eta\}$. Define $B_1 = \{\phi^{f_1(x)}x : x \in B\}$. Likewise, for $i = 2, \ldots, \lfloor H/h \rfloor$, define $f_i(x) = \inf\{t \geq f_{i-1}(x) + h : \bar{d}(x|_t^{t+h}, W) < \eta\}$, and $B_i = \{\phi^{f_i(x)}x : x \in B\}$. Now we can create a new tower with base $B' = \bigcup_i B_i$ and height h. (B' is measurable, since the f_i are.) This new tower used up all of the old tower, except for the non-nearcopies and the error between nearcopies, with total measure at most 2η. By another Fubini argument we can choose η small enough so that at least $(1 - \varepsilon)$ of the horizontal rows are at least $(1 - \varepsilon)$ monochromatic. If we also insist that $\eta < \varepsilon/2$, we are done. ∎

Proof of (b): By ergodicity, represent the flow ϕ as a flow under a function. We can make the simplifying assumption that $\mathcal{P}$ is **very good** in the sense of [9, p. 63], i.e., for each x, the $\mathcal{P}$-name for x consists entirely of monochromatic intervals, and the infimum of the lengths of these intervals, taken over all x, is positive, say α. This is possible for two reasons: First, any partition can be approximated arbitrarily well by a very good partition with the same number of atoms (because our measure is normalized $\nu\times$ Lebesgue, and sets measurable with respect to dyadic intervals generate the Lebesgue-measurable sets). Second, the map $\mathcal{P} \mapsto h(\phi^1, \mathcal{P})$ is continuous. Since $\mathcal{P}$ is very good, note that we have $\bar{d}(x|_0^t, x|_u^{t+u}) \leq u/\alpha$ for any $x \in X$ and any $t, u > 0$. (Here x is actually a point (e, s) in the product space.)

We shall use Feldman's *r-entropy* formulation to compute the entropy of ϕ. Call a collection $\mathcal{D}$ of disjoint measurable sets a $(\mathcal{P}, t, r)$-**family** if for all $D \in \mathcal{D}$,

$$x, y \in D \Longrightarrow \bar{d}(x|_0^t, y|_0^t) \leq r.$$

Let $|\mathcal{D}|$ denote the cardinality of $\mathcal{D}$, abbreviate $\bigcup_{D\in\mathcal{D}} D$ with $\bigcup \mathcal{D}$, and define $k_r(\phi, \mathcal{P})$ to be the infimum of numbers b such that for every positive ε, there exists an arbitrarily large t such that there exists a $(\mathcal{P}, t, r)$-family $\mathcal{D}$ with $\mu(\cup\mathcal{D}) > 1-\varepsilon$ and $(1/t)\log|\mathcal{D}| \le b$. By corollary 2.7 of [3], $h(\phi) = \lim_{r\to 0} k_r(\phi, \mathcal{P})$. Thus it will suffice to show that for fixed (small) r and ε, there exists an arbitrarily large t and a $(\mathcal{P}, t, r)$-family $\mathcal{D}$ with $\mu(\cup\mathcal{D})$ greater than $(1-\varepsilon)$ such that $\frac{1}{t}\log|\mathcal{D}|$ can be made arbitrarily small.

Fix (small) positive r, ε and (large) positive h_0. By lemma 4.2 we can produce an $(h, \varepsilon/2)$-tower with $h \ge h_0$ and such that for any two vertical words W', W'' of length h we have $\bar{d}(W', W'') < r\varepsilon/8$. (See Figure 5.) Let $t = h\varepsilon/4$ and set $\lambda = \alpha r/2$. We can certainly assume that h_0 was chosen large enough so that $\lambda < t$. Let B be the base of the tower and define

$$\mathcal{D} = \bigcup_{i=1}^{\lfloor (h-t)/\lambda \rfloor} D_i, \text{ where } D_i = B \times [(i-1)\lambda, i\lambda).$$

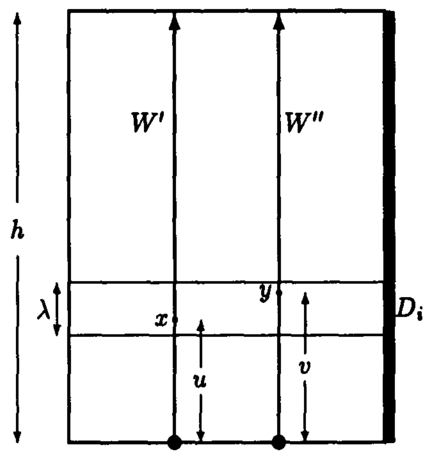

Figure 5: For any vertical words W', W'', $\bar{d}(W', W'') < r\varepsilon/8$.

Let $x, y \in D_i$ for some fixed i. Let the vertical coordinates of x and y respectively be u, v, and let W', W'' respectively be the vertical words (of length h) that x, y live on. Then we have

$$\bar{d}(x|_0^t, y|_0^t) = \bar{d}(W'|_u^{u+t}, W''|_v^{v+t}) \le \bar{d}(W'|_u^{u+t}, W''|_u^{u+t}) + \bar{d}(W''|_u^{u+t}, W''|_v^{v+t})$$
$$< r\varepsilon h/(8t) + \lambda/\alpha = r/2 + r/2 = r.$$

Thus $\mathcal{D}$ is a $(\mathcal{P}, t, r)$-family. Since $\mathcal{D}$ uses up all the vertical space of the tower except at most $(t+\lambda)/h < \varepsilon/4 + \varepsilon/4$, and the tower uses up all but at most a set of measure $\varepsilon/2$, we also have $\mu(\bigcup\mathcal{D}) > 1-\varepsilon$.

Now $|\mathcal{D}| < h/\lambda = 8t/(\varepsilon\alpha r)$, so $\frac{1}{t}\log|\mathcal{D}| \to 0$ as $t \to \infty$. Since we can pick h_0 as large as we want, t can be made arbitrarily large. ∎

ACKNOWLEDGEMENT: I am grateful to Daniel Rudolph and Jacob Feldman for illuminating discussions, and am especially indebted to Jonathan King for much advice, encouragement and inspiration. I also thank the referee for pointing out errors in the initial version and suggesting many improvements. This work was partially supported by an NSF Graduate Fellowship and a Department of Education National Need Fellowship.

References

[1] W. Ambrose, *Representation of ergodic flows*, Ann. Math. **42** (1941), 723–729.

[2] A. del Junco, *Transformations with discrete spectrum are stacking transformations*, Can. J. Math. **14** (1973), 26–38.

[3] J. Feldman, *r-entropy, equipartition, and Ornstein's isomorphism theorem in* $\mathbf{R}^n$, Israel J. Math. **36** (1980), 321–345.

[4] I. Cornfeld, S. Fomin and Ya. Sinai, *Ergodic Theory*, Springer-Verlag, New York, 1982.

[5] J. King, *The commutant is the weak closure of the powers, for rank-1 transformations.*, Ergodic Theory & Dynamical Systems **6** (1986), 363–384.

[6] J. King, *For mixing transformations, rank*$(T^k) = k\cdot$*rank*(T), Israel J. Math. **56** (1986), 102–122.

[7] J. King, *Joining-rank and the structure of finite rank mixing transformations*, J. Analyse Math. **51** (1988), 182–227.

[8] W. Krieger, *On entropy and generators of measure-preserving transformations*, Trans. Amer. Math. Soc. **119** (1970), 453–464.

[9] D. Ornstein, *Ergodic Theory, Randomness, and Dynamical Systems*, Yale University Press, New Haven, 1974.