

Classes of linear automata

ROBERT H. GILMAN

Department of Mathematics, Stevens Institute of Technology, Hoboken, NJ 07030, USA and School of Mathematics, Institute for Advanced Study, Princeton NJ 08540, USA

(Received 23 July 1985 and revised 2 March 1986)

Abstract. This paper discusses a division of one-dimensional cellular automata into classes which seem to correspond to different types of behaviour.

1. Introduction

We are interested in classifying cellular automata according to their behaviour, and we suggest an approach for linear automata (i.e. one-dimensional cellular automata).

The preface to [3] contains a general discussion of cellular automata and their applications. Linear automata are defined as follows: Let S be a finite alphabet of size $s > 1$ and $X = S^{\mathbb{Z}}$ the set of all maps from the integers to S . X is a compact metric space with metric $d(x, y) = 2^{-n}$, $n = \inf \{|i|: x(i) \neq y(i)\}$. Linear automata are induced by arbitrary maps $F: S^{2r+1} \rightarrow S$ where r is any non-negative integer. The automaton f induced by F is defined by $f(x) = y$ with $y(i) = F(x(i-r), \dots, x(i+r))$. In this situation we say f has range r . For example the (right) shift, σ , is the automaton of range 1 induced by $F(a, b, c) = a$. Linear automata are the continuous self-maps of X commuting with σ [7, theorem 3.4].

Let f be a linear automaton. Pick a probability distribution on S such that each $a \in S$ has positive probability, and let μ be the corresponding product measure on X . Wolfram [9], [10] has observed on the basis of extensive computer simulation that f may behave differently with different arguments; but if the argument is chosen at random, then the probability is high that the behaviour of f will be confined to one of four classes. We are thus led to consider linear automata in conjunction with a fixed measure, μ . In contrast to the usual situation in ergodic theory, f need not preserve μ .

We divide linear automata into three classes. Automata in class \mathcal{A} are equicontinuous on a set of measure 1. Automata in class \mathcal{B} satisfy a stochastic analogue of equicontinuity, and those in class \mathcal{C} satisfy a stochastic analogue of expansiveness. Class \mathcal{A} is independent of the choice of μ , but \mathcal{B} and \mathcal{C} may vary if μ changes.

The definitions of our classes are given in § 3. § 4 shows that the classes are invariant under certain transformations, while § 5 analyzes the dynamic behaviour of linear automata and makes a connection with Wolfram's classes. The connection might be stronger if we knew a way to decide membership of \mathcal{A} , \mathcal{B} , \mathcal{C} .

I would like to thank the Institute for Advanced Study for its hospitality while this paper was being written.

2. Notation and preliminary results

Keep the notation of the introduction: S is a finite alphabet, f is a linear automaton on $X = S^{\mathbb{Z}}$, d is the metric on X , and μ is a product measure induced by assigning each $a \in S$ a positive probability. $F : S^{2r+1} \rightarrow S$ is a map inducing f , and f has range r . Also σ is the right shift on X .

\mathbb{Z} is the integers, \mathbb{N} the positive integers, and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Intervals of integers are denoted $[i, j]$, $\langle -\infty, k \rangle$, etc.

A word is a sequence of letters in S . A word may be finite, infinite to either the left or right, or infinite in both directions. We distinguish between a word and a map from an interval of \mathbb{Z} to S . The words corresponding to the restriction of $x \in X$ to an interval are written $x|_{[i, j]}$, $x|_{\langle -\infty, k \rangle}$, etc. The maps obtained by restricting x are $x|_{[i, j]}$, $x|_{\langle -\infty, k \rangle}$, etc. For any $a \in S$ define $x_a \in X$ by $x_a(i) = a$, $i \in \mathbb{Z}$. Thus $x_a|_{[i, j]}$ is a sequence of $j - i + 1$ a 's. The length of any word w is $|w|$. The product of two words w and v is wv , their concatenation (as long as w is not infinite to the right nor v to the left).

Let T_σ be the set of all $x \in X$ which have dense forward and backward orbits under σ . An element x is in T_σ if and only if every finite word is obtained as $x|_{[i, j]}$ for some $i, j \in \mathbb{Z}$ with $0 \leq i \leq j$ and also for some $i, j \in \mathbb{Z}$ with $i \leq j \leq 0$. $X - T_\sigma$ is first category and $\mu(T_\sigma) = 1$.

A cylinder is a set $C(i, j, w) = \{x \in X : x|_{[i, j]} = w\}$ where $|w| = j - i + 1$. Cylinders are open and closed. The open disk of radius 2^{-n} around x is $C_n(x) = C(-n, n, x|_{[-n, n]})$, $n \in \mathbb{N}_0$.

The metric d is non-archimedean; that is, $d(x, y) \leq \max\{d(x, z), d(z, y)\}$. Consequently of any two disks $C_m(x)$, $C_n(y)$, either one contains the other or they intersect trivially. For fixed $n \in \mathbb{N}_0$ the relation $x \sim y$ if $d(x, y) < 2^{-n}$ is an equivalence relation with equivalence classes $\{C_n(x)\}$. Likewise

$$x \approx y \text{ iff } \forall i \in \mathbb{N}_0, d(f^i(x), f^i(y)) < 2^{-n}$$

is an equivalence relation, and we write $B_n(x)$ for the equivalence class of x .

$B_n(x)$ is the set of y for which $(f^i(y))|_{[-n, n]} = (f^i(x))|_{[-n, n]}$, $i \in \mathbb{N}_0$. If we visualize the behaviour of $(f$ with argument) x as an array $(a_{i,j})$ with entry $a_{i,j} = (f^i(x))(j)$ in row i and column j , then $B_n(x)$ is the set of all y whose behaviour agrees with that of x on the infinite vertical strip under the interval $[-n, n]$. More generally we will say that y has behaviour $B_n(x)$ on $[-n+k, n+k]$ if $\sigma^{-k}(y) \in B_n(x)$.

Insight into linear automata has been gained by studying with a computer finite portions of the array $(a_{i,j})$ defined above. Our division of linear automata into classes is based on the analysis of the sets $B_n(x)$, which correspond to certain infinite subarrays of $(a_{i,j})$. $B_n(x)$ is *ultimately periodic* if the sequence of words $x|_{[-n, n]}$, $(f(x))|_{[-n, n]}$, $(f^2(x))|_{[-n, n]}$, ... is ultimately periodic. For any $x \in X$, O_x is the (forward) orbit of x under f , $\overline{O_x}$ is the closure of O_x , and $\omega(x)$ is the set of limit points of O_x . For any $Y \subseteq X$, Y° is the interior of Y .

LEMMA 2.1. *If $x \in X$ and $n \in \mathbb{N}_0$, then:*

- (i) $B_n(x)$ is closed;
- (ii) $f(B_n(x)) \subseteq B_n(f(x))$;

(iii) f is equicontinuous at x if and only if $x \in B_m(x)^\circ$ for all $m \in \mathbb{N}_0$;

(iv) the restriction of f to $\overline{O_x}$ is equicontinuous if and only if $B_m(x)$ is ultimately periodic for all $m \in \mathbb{N}_0$.

Proof. Parts (i)–(iii) are immediate from the continuity of f and the relevant definitions.

If $f|_{\overline{O_x}}$ is equicontinuous, pick $\delta > 0$ such that for $y, z \in \overline{O_x}$, $d(y, z) < \delta$ implies $d(f^i(y), f^i(z)) < 2^{-m}$, $i \in \mathbb{N}_0$. As $\overline{O_x}$ is compact, there exist $j, k \in \mathbb{N}_0$, $j < k$, with $d(f^j(x), f^k(x)) < \delta$; and taking $y = f^j(x)$, $z = f^k(x)$ yields that $B_m(x)$ is ultimately periodic.

Conversely suppose that for all $m \in \mathbb{N}_0$, $B_m(x)$ is ultimately periodic. The sequence of words $\dots (f^i(x))[-m, m] \dots$ has period $p = p(m)$ when $i \geq i(m)$, and it suffices to show that every $y \in \overline{O_x}$ satisfies the same condition. It is enough to prove $(f^i(y), f^{i+p}(y)) < 2^{-m}$ if $i \geq i(m)$. We are done if $y \in O_x$, so assume $y \in \omega(x)$ and let $z = f^i(y)$. Pick $k \geq i(m)$ with $d(f^k(x), z) < 2^{-m-pr}$. From the way in which f is induced by F it follows that $d(f^{k+p}(x), f^p(z)) < 2^{-m}$ and, by choice of k , $d(f^k(x), f^{k+p}(x)) < 2^{-m}$. As d is non-archimedean, $d(f^p(z), z) < 2^{-m}$ as desired. \square

Recall the array $(a_{i,j})$ which we used above to represent the behaviour of x . If f has range r and if we know $(a_{i,j})$ on two vertical strips of width at least r and if we know the top row of $(a_{i,j})$ between the two strips, then the part of $(a_{i,j})$ between the two strips is uniquely determined. We have the following consequence.

LEMMA 2.2. *If f has range r and:*

(i) $x \in \sigma^{-i}(B_n(y)) \cap \sigma^j(B_n(z))$;

(ii) $2n + 1 \geq r$;

(iii) $0 \leq m \leq i, j \leq p$;

then

$$\sigma^{-i}(B_n(y)) \cap \sigma^j(B_n(z)) \cap C_p(x) \subseteq B_m(x).$$

An arbitrary array $(a_{i,j})$ with entries in S will represent the behaviour of its first row if each $a_{i,j}$ with $i \geq 1$ is the value of F on $(a_{i-1,j-r}, \dots, a_{i-1,j+r})$. Thus if $n \geq r$ and $(a_{i,j})$ agrees with the array for some element y when $j \leq n$ and if it agrees with the array for z when $j \geq -n$, then it is the array representing the behaviour of its top row. We use this observation below.

LEMMA 2.3. *Suppose $n \geq r$ and $i \leq -n, n \leq j$. Let $X = X_1 X_2 X_3$ be the product decomposition with factors corresponding to intervals $\langle -\infty, i \rangle, [i, j], \langle j, \infty \rangle$. Use μ to denote the product measure on the factors as well as on X . For any word w of length $j - i + 1$, let w_1 be the initial segment of length $n - i + 1$ and let w_2 be the terminal segment of length $j + n + 1$. For any $x \in X$ there is a product decomposition $C(i, j, w) \cap B_n(x) = D_1 D_2 D_3$. Here D_1 is the projection of $C(i, n, w_1) \cap B_n(x)$ to X_1 and D_3 is the projection of $C(-n, j, w_2) \cap B_n(x)$ to X_3 . D_2 is the singleton corresponding to w ; i.e. $D_2 = \{y|_{[i,j]}\}$ for any $y \in C(i, j, w)$. Further $\mu(D_2) = \mu(C(i, j, w))$ and $\mu(C(i, j, w) \cap B_n(x)) = \mu(D_1)\mu(D_2)\mu(D_3)$.*

Proof. Let $Y = C(i, j, w) \cap B_n(x)$, and define D_i as above. If $y \in Y$, then $y|_{[i,j]} = w$, and it is evident that the projection of y to X_i lies in D_i . Consequently $Y \subseteq D_1 D_2 D_3$.

Conversely pick any $y_1 \in C(i, n, w_1) \cap B_n(x)$ and $y_3 \in C(-n, j, w_2) \cap B_n(x)$. Define $z \in X$ by $z(i) = y_1(i)$ if $i \leq n$ and $z(i) = y_2(i)$ if $i \geq -n$. As $y_1, y_2 \in B_n(x)$, z is well defined; and by the discussion before the lemma, $z \in B_n(x)$. By construction $z[i, j] = w$, so $z \in Y$ and the projection of z to X_2 is in D_2 . Likewise z and y_i have the same projection on X_i , $i = 1, 3$. Thus $Y \supseteq D_1 D_2 D_3$. The part of lemma 2.3 involving measure follows easily from $Y = D_1 D_2 D_3$. \square

Next we develop a measure-theoretic analogue of the interior of a set. For any measurable set E define

$$\rho_E(x) = \lambda \quad \text{if } \lim_{n \rightarrow \infty} \frac{\mu(C_n(x) \cap E)}{\mu(C_n(x))} = \lambda. \tag{2.1}$$

LEMMA 2.4 (Lebesgue). *For almost all $x \in E$, $\rho_E(x) = 1$.*

Proof. Let E_m be the set of x in E for which $\liminf_{n \rightarrow \infty} \mu(C_n(x) \cap E) / \mu(C_n(x)) \geq 1 - 1/m$. If $\mu(E_m) = \mu(E)$ for all m , we are done; so assume $\mu(E_m) < \mu(E)$ for some m . Replacing E by $E - E_m$ yields that $\mu(E) > 0$ and for each $x \in E$

$$\mu(C_n(x) \cap E) < (1 - 1/m)\mu(C_n(x)) \tag{2.2}$$

for infinitely many $n \in \mathbb{N}_0$. As μ is regular [8, theorem 6.1] we may replace E by a closed subset and assume E is compact. For any $\varepsilon > 0$, E is covered by a finite union of disks, $C_i = C_{k_i}(y_i)$, with $\mu(E) \geq \sum \mu(C_i) - \varepsilon\mu(E)$. Thus

$$\mu(E) \geq \left(\frac{1}{1 + \varepsilon}\right) \sum \mu(C_i). \tag{2.3}$$

Pick ε so that $1/(1 + \varepsilon) > 1 - 1/m$. Since if two C_i 's intersect non-trivially, one contains the other, we may assume $C_i \cap C_j = \emptyset$, $i \neq j$. For each i and $x \in C_i \cap E$ pick $n \geq k_i$ so that (2.2) holds. A finite number of these $C_n(x)$'s are pairwise disjoint and cover E . As each $C_n(x)$ lies in a C_i , (2.2) yields $\mu(E) < (1 - 1/m)\sum \mu(C_i)$ contrary to (2.3). \square

LEMMA 2.5. *For any measurable set $E \subseteq X$ let*

$$E' = \{x \in E : \rho_{\sigma^i(E)}(\sigma^i(x)) = 1, i \in \mathbb{Z}\},$$

where ρ is defined by (2.1). For any measurable subsets $E, F \subseteq X$:

- (i) $\mu(E') = \mu(E)$;
- (ii) $\sigma^i(E') = [\sigma^i(E)]'$, $i \in \mathbb{Z}$;
- (iii) if $E \subseteq F$, then $E' \subseteq F'$;
- (iv) $E' \cap F' \subseteq (E \cap F)'$;
- (v) if $E = C(i, j, w)$, then $E' = E$.

Proof. As σ is measure preserving, lemma 2.4 yields (i); and (ii) and (v) are clear from the definition of E' . For (iii) note that $\rho_E(x) = 1$ and $E \subseteq F$ imply $\rho_F(x) = 1$, and likewise for (iv) $\rho_E(x) = \rho_F(x) = 1$ implies $\rho_{E \cap F}(x) = 1$. \square

Our final result concerns equicontinuous maps. It is well-known that if $h : Y \rightarrow Y$ is an equicontinuous homeomorphism of the compact metric space Y to itself, then the closure of each $\langle h \rangle$ -orbit can be given the structure of a compact topological group in such a way that the action of h becomes rotation by a group element. In

general the restriction of h to $\Omega(h)$, the set of non-wandering points, is always an equicontinuous homeomorphism [2, theorems 4 and 8] and there is a useful projection from Y to $\Omega(h)$.

LEMMA 2.6. *Let $h: Y \rightarrow Y$ be an equicontinuous map of a compact metric space.*

(i) *The restriction of h to $\Omega(h)$ is an equicontinuous homeomorphism of $\Omega(h)$ onto itself;*

(ii) *there is a unique continuous projection $p: Y \rightarrow \Omega(h)$ such that the distance between $h^n(y)$ and $h^n(p(y))$ goes to zero uniformly in y as n goes to infinity. This projection commutes with every continuous map commuting with h .*

Proof. From the reference given above we have that $\Omega(h) = \bigcup_{y \in Y} \omega(y)$ and that (i) holds. By [5, theorem 8.7] y is proximal to some $y' \in \omega(y)$. Because h is equicontinuous, y is forwardly asymptotic to y' ; and since an equicontinuous homeomorphism is distal, y' is the unique point in $\Omega(h)$ to which y is forwardly asymptotic.

Define $p(y) = y'$. Our conditions imply that p is the identity on $\Omega(h)$. Likewise it follows in a straightforward way that p is continuous and that the distance between $h^n(y)$ and $h^n(p(y))$ goes to zero uniformly in y . If p_1 is another projection with these properties, then for every $y \in Y$, y is forwardly asymptotic to $p_1(y)$ whence $p_1(y) = p(y)$. Also if g is continuous and commutes with h , then the distance between $g(h^n(y))$ and $g(h^n(p(y)))$ goes to zero as n goes to infinity; and $g(y)$ is forwardly asymptotic to $g(p(y))$. As g must map $\Omega(h)$ to itself, $g(p(y)) \in \Omega(h)$, and we conclude $g(p(y)) = p(g(y))$ as desired. \square

3. Definitions and examples

In this section we divide linear automata into three classes and give examples of automata in each class. Recall the notation and results of § 2 and let f be a fixed linear automaton on $X = S^Z$. For $\varepsilon > 0$ and $x \in X$ let $D(x, \varepsilon)$ be the set of all y such that $d(f^i(x), f^i(y)) < \varepsilon$ for all $i \in \mathbb{N}_0$. Recall that f is equicontinuous at x if for all $\varepsilon > 0$ there is $n \in \mathbb{N}_0$ such that $C_n(x)$, the open ball of radius 2^{-n} , lies in $D(x, \varepsilon)$; and f is expansive if there is $\delta > 0$ such that for all x $D(x, \delta) = \{x\}$. We use the product measure, μ , to define stochastic analogues of equicontinuity and expansiveness, and then we define the classes of automata.

Definition 3.1. f is *almost equicontinuous at x* if for all $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \frac{\mu(C_n(x) \cap D(x, \varepsilon))}{\mu(C_n(x))} = 1.$$

Definition 3.2. f is *almost expansive* if there is $\delta > 0$ such that for all $x \in X$, $\mu(D(x, \delta)) = 0$.

Definition 3.3. Define classes of linear automata as follows:

- (i) $f \in \mathcal{A}$ if f is equicontinuous at some $x \in X$;
- (ii) $f \in \mathcal{B}$ if f is almost equicontinuous at some $x \in X$ but $f \notin \mathcal{A}$;
- (iii) $f \in \mathcal{C}$ if f is almost expansive.

The following propositions give some information about the classes \mathcal{A} , \mathcal{B} , \mathcal{C} , and in particular imply that these classes form a partition of linear automata.

Let f have range r . Recall that the behaviour of (f with argument) x may be visualized as an array $(a_{i,j})$ with $a_{i,j} = (f^i(x))(j)$ and that $B_n(x)$ is the class of all $y \in X$ whose behaviour agrees with that of x on the infinite subarray given by $|j| \leq n$. Thus $B_n(x) = D(x, 2^{-n})$, and there are equivalent versions of definitions 3.1 and 3.2 with $B_n(x)$ in place of $D(x, \epsilon)$ and $D(x, \delta)$. T_σ is the set of $x \in X$ with dense forward and backward orbit under the shift σ .

PROPOSITION 3.4. *The following are equivalent:*

- (i) $f \in \mathcal{A}$;
- (ii) f is equicontinuous at some $x \in X$;
- (iii) f is equicontinuous on a set of measure 1;
- (iv) f is equicontinuous on T_σ ;
- (v) for some $n \geq (r-1)/2$ there is a class $B_n(x)$ with $B_n(x)^\circ \neq \emptyset$;
- (vi) for all $n \geq 0$ there is a class $B_n(x)$ with $B_n(x)^\circ \neq \emptyset$.

Proof. By definition of \mathcal{A} (i) is equivalent to (ii), and the implications (iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (vi) \Rightarrow (v) are immediate from § 2. To show (v) \Rightarrow (iv) pick $B_n(x)$ as in (v). If $y \in T_\sigma$, then for any $m \in \mathbb{N}$ there are $i, j \geq m$ with $\sigma^i(y) \in B_n(x)^\circ$, $\sigma^{-j}(y) \in B_n(x)^\circ$. As σ is a homeomorphism, $y \in \sigma^{-i}(B_n(x)^\circ) \cap \sigma^j(B_n(x)^\circ)$. By lemma 2.2, $y \in B_m(y)^\circ$; and as m is arbitrary, f is equicontinuous at y . □

PROPOSITION 3.5. *The following are equivalent:*

- (i) $f \in \mathcal{A} \cup \mathcal{B}$;
- (ii) f is almost equicontinuous at some $x \in X$;
- (iii) f is almost equicontinuous on a set of measure 1;
- (iv) for some $n \geq (r-1)/2$ there is a class $B_n(x)$ with $\mu(B_n(x)) > 0$;
- (v) for all $n \geq 0$ there is a class $B_n(x)$ with $\mu(B_n(x)) > 0$.

Proof. It is easy to check that (i) and (ii) are equivalent and that (iii) \Rightarrow (ii) \Rightarrow (v) \Rightarrow (iv). Pick $B_n(x)$ as in (iv) and let Y be the set of $y \in X$ whose forward and backward orbits intersect $B_n(x)'$ infinitely often. By lemma 2.5, $\mu(B_n(x)') = \mu(B_n(x)) > 0$. As Y is $\langle \sigma \rangle$ -invariant, $\mu(Y) = 1$ by ergodicity. To prove that (iii) holds it suffices to show that for all $y \in Y$ and $m \in \mathbb{N}$, $y \in B_m(y)'$. Use lemma 2.5 and argue as in the preceding proof with $B_n(x)'$, $B_m(y)'$ and Y in place of $B_n(x)^\circ$, $B_m(y)^\circ$ and T_σ . □

PROPOSITION 3.6. *The following are equivalent:*

- (i) $f \in \mathcal{C}$;
- (ii) f is almost expansive;
- (iii) f is almost expansive with $\delta = 2^{(1-r)/2}$;
- (iv) there is $n \in \mathbb{N}_0$ such that $\mu(B_n(x)) = 0$ for all $x \in X$;
- (v) for all $n \geq (r-1)/2$ and all $x \in X$, $\mu(B_n(x)) = 0$.

Proof. By definitions 3.2 and 3.3, (i), (ii) and (iv) are equivalent. Also (v) \Rightarrow (iii) \Rightarrow (ii). Finally (iv) \Rightarrow (v) by proposition 3.5. □

It might seem that $\mathcal{A} \cup \mathcal{B}$ is a natural class to study, but automata in \mathcal{A} and \mathcal{B} behave differently. For example if $f \in \mathcal{A}$ and $x \in T_\sigma$, then $f^i(y)$ will stay arbitrarily

close to $f^i(x)$ if y is close enough to x , but the following proposition shows that if $f \in \mathcal{B}$, then the probability that $f^i(y)$ will diverge a distance at least $2^{(1-r)/2}$ from $f^i(x)$ is always positive.

PROPOSITION 3.7. *If $f \in \mathcal{B}$, then for all $x \in X$ and $n \in \mathbb{N}_0$, $\mu(C_n(x) \cap D(x, 2^{(1-r)/2})) < \mu(C_n(x))$.*

Proof. Let m be the least integer not less than $(r-1)/2$; then $D(x, 2^{(1-r)/2}) = D(x, 2^{-m}) = B_m(x)$. If the inequality above fails, then $\mu(C_n(x) \cap B_m(x)) = \mu(C_n(x))$ for some $n \in \mathbb{N}_0$ and $x \in X$. As $C_n(x)$ is open and $B_m(x)$ closed, $C_n(x) - B_m(x)$ is open of measure 0 and hence empty. By proposition 3.4(v), $f \in \mathcal{A}$. □

We know that if $f \in \mathcal{A}$ and $x \in T_\sigma$, then $x \in B_n(x)^\circ$ for all $n \in \mathbb{N}_0$. Thus $C_k(x) \subseteq B_n(x)$ for some $k \in \mathbb{N}_0$. In other words the approximate behaviour of x (i.e. the behaviour on the interval $[-n, n]$) is determined by a finite amount of information about x , namely $x(i), |i| \leq k$. Proposition 3.7 says that if $f \in \mathcal{B}$ and $x \in X$, no finite amount of information about x determines $B_n(x)$ when $n \geq (r-1)/2$. Examples show that $B_n(x)^\circ \neq \emptyset$ is possible for $f \notin \mathcal{A}$ if $n < (r-1)/2$.

The rest of this section is devoted to examples. By definition 3.3 every expansive linear automaton, for example the shift, lies in \mathcal{C} . An automaton f has finite order if $f^m = f^n$ for some $m < n$. Linear automata of finite order are easily seen to be equicontinuous at all $x \in X$ and so lie in \mathcal{A} . (Using lemma 2.1 one can show that linear automata equicontinuous at all $x \in X$ have finite order.) By [7, theorem 6.13] every finite group is embeddable in the monoid of linear automata, so there are many linear automata of finite order.

Take an alphabet $S = \{0, 1\}$ and let f be induced by $F: S^3 \rightarrow S$, $F(1, 1, 1) = 1$, $F(a, b, c) = 0$ otherwise. In the notation of § 2, $C_1(x_0) \subseteq B_0(x_0)$ whence $f \in \mathcal{A}$ by proposition 3.4(v). It is easy to see that $\lim_{n \rightarrow \infty} f^n(x) = x_0$ except for $x = x_1$. Thus \mathcal{A} has members which are not equicontinuous at all $x \in X$.

It is somewhat more difficult to find $f \in \mathcal{B}$. Take $S = \{0, 1, 2\}$ and define $F: S^3 \rightarrow S$ as indicated below.

*00	*01	*02	*10	*11	*12	*20	*21	*22
0	1	0	0	1	0	2	0	2

Here $*$ stands for any $a \in S$. Visualize the array $(a_{i,j})$, $a_{i,j} = (f^i(x))(j)$ and consider 0 as a background element. In passing from one row of $(a_{i,j})$ to the row below 1's move to the left, 2's move straight down, and a 1 and 2 which collide annihilate each other. It is straightforward to show that if $x(j) = 2$, this 2 is annihilated at generation i (i.e. $a_{i,j} = 0$, $a_{i-1,j} = 2$) if and only if the word $x[j, j+k]$ has more 2's than 1's if $0 \leq k < i$ and an equal number when $k = i$.

We connect the behaviour of f with a random walk on the integers (for a similar connection in the analysis of another linear automaton see [6]). Let $p(a)$ be the probability assigned to $a \in S$ in defining μ , and consider a random walk with $p(1)$, $p(2)$, $p(0)$ the probabilities of moving left, moving right, and remaining stationary respectively. If $x(0) = 2$, then the probability that this 2 is never annihilated equals the probability of no return to 0 starting at 1 in the random walk. It can be shown

[4, chapter 14.9, problem 1] that this probability is positive if and only if $p(2) > p(1)$. Thus $p(2) > p(1)$ implies $\mu(B_0(x_2)) > 0$ whence $f \in \mathcal{A} \cup \mathcal{B}$ by proposition 3.5(iv). On the other hand for any $n \in \mathbb{N}_0$ if $x(j) = 0$ for all $j > n$, then $(f^i(x))(0)$ is eventually 0 or 2, while $x(j) = 1, j > n$, implies that $(f^i(x))(0)$ is eventually 1. From the discussion after proposition 3.7, we conclude $f \notin \mathcal{A}$.

By pursuing the connection with the random walk one can show $f \in \mathcal{C}$ if $p(1) \geq p(2)$ and so obtain $f \in \mathcal{C}$ with f not expansive.

4. Factors

We show that the classes $\mathcal{A}, \mathcal{B}, \mathcal{C}$ defined in § 3 are invariant in certain ways. Keep the notation of § 2 and let $f_i, i = 1, 2$ be linear automata defined on $X_i = S_i^{\mathbb{Z}}$. Use σ to denote the shift on both X_1 and X_2 and d for both metrics. Say that f_2 is a factor of f_1 if there is a continuous surjection $\pi : X_1 \rightarrow X_2$ such that $f_2\pi = \pi f_1$, and $\sigma\pi = \pi\sigma$. We will show that if $f_1 \in \mathcal{A}$, then $f_2 \in \mathcal{A}$. The key to this result is that π is semi-open by which we mean that the image of every open set has non-empty interior. Also every linear automaton is semi-open relative to its image (i.e. as a map $f : X \rightarrow f(X)$). The term semi-open is due to R. Adler and B. Marcus [1, page 7].

In preparation for the next lemma let $g : X_1 \rightarrow X_2$ be continuous with $g\sigma = \sigma g$. By uniform continuity there is $r \in \mathbb{N}_0$ such that $d(x, y) < 2^{-r}$ implies $d(g(x), g(y)) < 1$. Thus $x[-r, r]$ determines $(g(x))(0)$ according to some map $G : S_1^{2r+1} \rightarrow S_2$. As $g\sigma = \sigma g$,

$$(g(x))(i) = G(x[i-r, i+r]) \quad x \in X_1, i \in \mathbb{Z}. \tag{4.1}$$

For any word w of length $|w| \geq 2r+1$ in the letters S_1 , applying G to the succession of subwords of length $2r+1$ yields a word $G(w)$ of length $|w| - 2r$ in the letters S_2 . Let the head of $w, |w| \geq 2r$, be the initial subword of length $2r$, and let the tail of w be the terminal subword of the same length. If the tail of w_1 matches the head of w_2 and w is obtained by identifying the head and tail, then $G(w) = G(w_1)G(w_2)$. Recall that $C(i, j, w) = \{x : x[i, j] = w\}$. From (4.1)

$$g(C(i, j, w)) \subseteq C(i+r, j-r, G(w)) \quad |w| \geq 2r+1. \tag{4.2}$$

LEMMA 4.1. Let $g : X_1 \rightarrow X_2$ be as above, then:

- (i) g is semi-open with respect to $g(X_1)$;
- (ii) if $x \in T_\sigma \cap U, U$ open in X_1 , then $g(x) \in g(U)^\circ$, the interior of $g(U)$.

Proof. Pick a word w_1 in the letters S_2 so that $\{v : G(v) = w_1\}$ is non-empty and has minimum number of distinct tails. Let t be one of the tails which occur.

$$\text{If } G(u) = w'_1 w_1, \text{ then } u \text{ may be chosen to have tail } t. \tag{4.3}$$

Indeed if (4.3) failed, then replacing w_1 by $w'_1 w_1$ would contradict our choice of w_1 . Likewise choose w_2 so that $\{v : G(v) = w_2\}$ is non-empty and has minimum number of distinct heads. Let h be one of the heads.

Clearly (ii) implies (i). Suppose x is as in (ii). Every finite word occurs as $x[i, j]$ for $i > 0$ and for $j < 0$. It follows that we can choose i and j so that $x[i, j] = v_1 v v_2, |v| \geq 1, v_1$ has tail $t, G(v_1) = w_1, v_2$ has head h , and $G(v_2) = w_2$. Further we can pick i and j so that $C(i_1, j_1, tvh) \subseteq U$ where $i \leq i_1 < j_1 \leq j$ and $x[i_1, j_1] = tvh$.

Let $Y = g(X_1)$ and $w = G(tvh)$; we have $G(v_1vv_2) = w_1ww_2$. It suffices to show

$$\bigcup g(C(i, j, v'_1vv'_2)) = C(i+r, j-r, w_1ww_2) \cap Y, \tag{4.4}$$

where the union is over all v'_1 with tail t and $G(v'_1) = w_1$ and all v'_2 with head h and $G(v'_2) = w_2$. (4.2) yields inclusion from left to right, so pick y in the righthand set of (4.4). We have $y[i+r, j-r] = w_1ww_2$ and $y = g(z)$ for some $z \in X_1$. We need to show that z can be chosen with $z[i, j] = v'_1vv'_2$. Using (4.3) construct a word u_1 infinite to the left, ending in some v'_1 , and such that $G(u_1) = y\langle -\infty, i+r+|w_1| \rangle$. Likewise construct u_2 infinite to the right, beginning with some v'_2 , and such that $G(u_2) = y\langle j-r-|w_2|, \infty \rangle$. Now u_1, v , and u_2 may be used to define the desired z . □

COROLLARY 4.2. *A linear automaton f is semi-open as a map from X to $f(X)$.*

Proof. Take $X_1 = X_2 = X$ and $g = f$ in lemma 4.1. □

PROPOSITION 4.3. *Let f_2 be a factor of f_1 . If $f_1 \in \mathcal{A}$, then $f_2 \in \mathcal{A}$.*

Proof. Let $\pi : X_1 \rightarrow X_2$ be as in the definition of factor at the beginning of this section, and let f_2 have range r . By uniform continuity $d(\pi(x), \pi(y)) < 2^{-r}$ if $d(x, y) < 2^{-n}$ for some $n \in \mathbb{N}_0$. It follows from $\pi f_1^i = f_2^i \pi, i \in \mathbb{N}_0$, that $\pi(B_n(x)) \subseteq B_r(\pi(x))$. Take $B_n(x)^\circ \neq \emptyset$ by proposition 3.4(vi). By lemma 4.1 $B_r(\pi(x))^\circ \neq \emptyset$ whence $f_2 \in \mathcal{A}$ by proposition 3.4(v). □

There is an analogue of proposition 4.3 which applies to the union of classes $\mathcal{A} \cup \mathcal{B}$. Recall the notation at the beginning of this section and let μ_i be the product measure for X_i . Say that (f_2, μ_2) is a factor of (f_1, μ_1) if f_2 is a factor of f_1 and $\mu_1 \pi^{-1} = \mu_2$.

PROPOSITION 4.4. *Let (f_2, μ_2) be a factor of (f_1, μ_1) . If $f_1 \in \mathcal{A} \cup \mathcal{B}$, then $f_2 \in \mathcal{A} \cup \mathcal{B}$.*

Proof. Choose r and n as in the proof of proposition 4.3. Choose x such that $\mu_1(B_n(x)) > 0$ by proposition 3.5(v). Let $y = \pi(x)$; we have $\mu_1(\pi^{-1}(B_r(y))) > 0$ whence $\mu_2(B_r(y)) > 0$ and $f_2 \in \mathcal{A} \cup \mathcal{B}$ by proposition 3.5(iv). □

Propositions 4.3 and 4.4 have the following consequence.

COROLLARY 4.5. *If μ is the equiprobable measure on X and g is a bijective linear automaton, then g preserves the classes $\mathcal{A}, \mathcal{B}, \mathcal{C}$. I.e. for any linear automaton f on X , f and gfg^{-1} are in the same class.*

Proof. It follows from [7, theorem 5.4] that g is measure-preserving. □

5. Video dynamics

We are interested in what properties of the array $(a_{i,j}), a_{i,j} = (f^i(x))(j)$ are determined by the class of a linear automaton f . If $f \in \mathcal{A} \cup \mathcal{B}$, then certain parts of the array will repeat with fixed frequency independent of x for almost all x , while if $f \in \mathcal{C}$, they will not. If x is changed on an interval, then only a finite number of additional columns of $(a_{i,j})$ will be affected when $f \in \mathcal{A}$, but if $f \in \mathcal{B} \cup \mathcal{C}$, infinitely many columns may change. We will also show that if $f \in \mathcal{A}$, then almost all x behave asymptotically like rotations of a group.

Keep the notation of § 2 and suppose $f \in \mathcal{A} \cup \mathcal{B}$. By proposition 3.5 for any $n \in \mathbb{N}_0$ there is $B_n(x)$ with $\mu(B_n(x)) > 0$.

PROPOSITION 5.1. *Suppose $f \in \mathcal{A} \cup \mathcal{B}$ and for each $n \in \mathbb{N}_0$, $B_n(x_{1,n}), B_n(x_{2,n}), \dots$ are the countably many classes $B_n(x)$ with $\mu(B_n(x)) > 0$; then:*

(i) *the probability that x has behaviour $B_n(x_{i,n})$ on a given interval $[j - n, j + n]$ is $\mu(B_n(x_{i,n}))$;*

(ii) *there is a $\langle \sigma \rangle$ -invariant set Y of measure 1 such that for $x \in Y$ the behaviour of x on every interval $[j - n, j + n]$ is some $B_n(x_{i,n})$; and the frequency with which a given $B_n(x_{i,n})$ appears is asymptotic to $\mu(B_n(x_{i,n}))$.*

Proof. Part (i) is clear. For (ii) use proposition 3.5 to pick Y_0 of measure 1 such that f is almost equicontinuous on Y_0 . For each $B_n(x_{i,n})$ the ergodic theorem yields $Y_{i,n}$, $\mu(Y_{i,n}) = 1$ such that for $x \in Y_{i,n}$, $B_n(x_{i,n})$ occurs with limiting frequency $\mu(B_n(x_{i,n}))$ in both directions. Choose Y to be a $\langle \sigma \rangle$ -invariant subset of measure 1 of $\bigcap Y_{i,n}$. Because f commutes with σ , it suffices to prove the first assertion of (ii) for $j = 0$; and this case follows from the definition of almost equicontinuity. \square

The next proposition gives more information for $f \in \mathcal{A}$ and in particular implies that when $f \in \mathcal{A}$ the classes with $\mu(B_n(x)) > 0$ are the ones with non-empty interior.

PROPOSITION 5.2. *Suppose $f \in \mathcal{A}$ and for each $n \in \mathbb{N}_0$, $B_n(y_{1,n}), B_n(y_{2,n}), \dots$ are the countably many classes $B_n(x)$ with $B_n(x)^\circ \neq \emptyset$. Then:*

(i) *each $B_n(y_{i,n})$ is ultimately periodic;*

(ii) *if $x \in T_\sigma$, then the behaviour of x on every interval $[j - n, j + n]$ is some $B_n(y_{i,n})$; and each $B_n(y_{i,n})$ appears infinitely often;*

(iii) *if $x \in T_\sigma$, then the restriction of f to $\overline{O_x}$ is equicontinuous, and f behaves asymptotically like a rotation of a compact group.*

Proof. To prove (i) note that since periodic points (periodic with respect to the shift) are dense, $B_n(y_{i,n})$ contains a periodic point z . If z has period p , then so does $f^j(z)$, and it follows that for some $j < k$, $f^j(z) = f^k(z)$. Hence, $z, f(z), f^2(z), \dots$ is ultimately periodic, and so $B_n(z) = B_n(y_{i,n})$ is too.

The second part of (ii) holds because the forward and backward orbits of x under $\langle \sigma \rangle$ are dense. It suffices to prove the first part of (ii) for $j = 0$. By proposition 3.4(iv) f is equicontinuous at x whence $x \in B_n(x)^\circ$. But then $B_n(x) = B_n(y_{i,n})$ for some i .

The first part of (iii) follows from (i), (ii), and lemma 2.1(iv). The last part is lemma 2.6. \square

We pause to make some observations. First we note that in the situation of proposition 5.2(iii) f is equicontinuous at x by proposition 3.4, but examples show that f need not be equicontinuous at $f(x)$.

Suppose $f \in \mathcal{C}$ and $n \geq (r - 1)/2$ (recall that f has range r). By proposition 3.6, $\mu(B_n(x)) = 0$ for all x . As there are only countably many ultimately periodic behaviours, the probability that a random x has ultimately periodic behaviour on any interval $[j - n, j + n]$ is 0. However, ultimately periodic behaviour can occur

with positive probability if $n < (r - 1)/2$. This fact may account for the reducible automata of [9, § 3].

As we mentioned in the introduction, we do not have an algorithm for deciding membership in the classes \mathcal{A} , \mathcal{B} , \mathcal{C} . We might try to distinguish between $\mathcal{A} \cup \mathcal{B}$ and \mathcal{C} as follows: Choose x at random and view a finite vertical strip $(a_{i,j}), 0 \leq i \leq i_1, -r \leq j \leq r$. Repeat this procedure several times and count the frequency with which different strips occur. If $f \in \mathcal{A} \cup \mathcal{B}$ and $p = \sup_x \mu(B_r(x))$ then, by proposition 5.1, $p > 0$. In this case we expect to see some strips occurring with frequency about p or more no matter how large i_1 is. On the other hand if $f \in \mathcal{C}$, then $\mu(B_r(x)) = 0$ for all x and we expect the frequencies to go to zero as i_1 increases. Thus we may guess whether or not $f \in \mathcal{A} \cup \mathcal{B}$ based on the observed frequencies. Since there are only finitely many linear automata with range r and fixed alphabet of size s , we may let $p(r, s)$ be the minimum of p as f ranges over all such automata in $\mathcal{A} \cup \mathcal{B}$. We have $p(r, s) > 0$, and if $p(r, s)$ were computable one might be able to make the heuristic test above reliable.

It is worth noting that to view a vertical strip of $(a_{i,j})$ we must compute a triangle containing the strip. The reason for this is that if we know

$$x(j) \quad -j_1 \leq j \leq j_1,$$

we can compute only

$$(f(x))(j) \quad -j_1 + r \leq j \leq j_1 - r$$

and

$$(f^2(x))(j) \quad -j_1 + 2r \leq j \leq j_1 - 2r$$

etc. Most computer simulations of linear automata seem to avoid this difficulty by assuming that x has period $2j_1 + 1$. This procedure changes the sample space from which the inputs are drawn, but on the basis of a few computer experiments it does not seem to affect the appearance of the computer display as long as j_1 is suitably large.

Now we resume our study of $(a_{i,j})$ and consider what happens when x is changed at random on an interval I . We will find that propagation of changes is different for each of the classes \mathcal{A} , \mathcal{B} , \mathcal{C} . These differences can be used to create a heuristic test for membership like the one discussed above.

We assume that I is never equal to all of \mathbb{Z} . If a change on I produces a change in $a_{i,j}$ for some i , we say that site j is affected. It is clear from the way f is induced by F that if no site in an interval J of length r is affected, then the change does not propagate beyond (i.e. affect any site beyond) J . By a random change on I we mean random with respect to the product measure on the factor of X corresponding to I .

For a given x and interval I let p_t be the probability that a random change on I affects a site a distance greater than t from I . $1 - p_0$ is the probability that the change does not propagate beyond I . As $p_t \geq p_{t+1}$, $\lim_{t \rightarrow \infty} p_t$ exists and is the probability of infinite propagation.

Suppose that I is infinite to the right. Shift x so that I begins at $j + 1$ for some $j \geq r$. Since σ preserves μ and commutes with f , the probabilities p_t are not changed

by the shift. Replace x by its shift, apply lemma 2.3, and obtain:

$$\begin{aligned}
 I &= \langle j, \infty \rangle \quad \text{and} \quad j \geq r \\
 C(i, j, w) \cap B_r(x) &= B^- \{x|_{[i, j]}\} B^+ \quad \text{for } i \leq -r, w = x[i, j] \\
 \mu(C(i, j, w) \cap B_r(x)) &= \mu(B^-) \mu(B^+) \mu(C(i, j, w)) \\
 \mu(B^+) &= 1 - p_{j-r}.
 \end{aligned}
 \tag{5.1}$$

PROPOSITION 5.3. *If $f \in \mathcal{B}$, there is a $\langle \sigma \rangle$ -invariant set $Y \subseteq X$ of measure 1 and a direction (left or right) such that for all $x \in Y$:*

- (i) $p_t < 1$ for all t ;
- (ii) for any infinite interval $\lim_{t \rightarrow \infty} p_t = 0$;
- (iii) for any interval infinite in the distinguished direction, $p_t > 0$ for all t .

Proof. Pick Y to be the same as in proposition 5.1(ii). If I is finite, then with positive probability a random change on I leaves x unchanged whence (i) holds. If I is infinite (say to the right), then because Y is $\langle \sigma \rangle$ -invariant, we may shift x so that (5.1) holds with $i = -r, j = r. 1 - p_0 = \mu(B^+) > 0$; and as $p_t \geq p_{t+1}$, (i) is proved.

Let $E = B_r(x_{1,n})$ in the notation of proposition 5.1 and set

$$q_m(y) = \frac{\mu(C_m(y) \cap E)}{\mu(C_m(y))}.$$

By lemma 2.4, $\lim_{m \rightarrow \infty} q_m(y) = 1$ for almost all $y \in E$. By Egoroff's theorem $q_m(y) \rightarrow 1$ uniformly on some $E_0 \subseteq E$ with $\mu(E_0) > 0$. Replace Y by its subset of elements whose positive and negative orbits under $\langle \sigma \rangle$ intersect E_0 infinitely often. With this change Y still satisfies the hypothesis and (i) still holds.

To prove (ii) let I be an infinite interval. Assume I is infinite to the right (the other case is similar). Given $\varepsilon > 0$ we need only find t with $p_t < \varepsilon$. By choice of Y we can shift x so (5.1) holds with $i = -j, j$ arbitrarily large, and $x \in E_0$. As $q_m(y) \rightarrow 1$ uniformly on E_0 , we pick j large enough so that $q_j(y) > 1 - \varepsilon, y \in E_0$. In particular, $q_j(x) > 1 - \varepsilon$. As $x \in E_0, x \in B_r(x_1)$ whence $B_r(x) = B_r(x_1)$. Now as $C(-j, j, x[-j, j]) = C_j(x)$, we have by (5.1)

$$\mu(B^-) \mu(B^+) \mu(C_j(x)) > (1 - \varepsilon) \mu(C_j(x))$$

whence $p_{j-r} = 1 - \mu(B^+) < \varepsilon$ as desired.

Suppose (iii) fails; then we can find $x \in Y$ and a right-infinite interval of x with $p_t = 0$ for all $t \geq t_0$. Shift x so that $x \in E_0$ and (5.1) holds with $i = -r$, and $j - r \geq t_0$. In particular $B_r(x) = B_r(x_{1,n})$. Conclude from lemma 2.3 that the projection of $B_r(x_{1,n}) \cap C(-r, j, x[-r, j])$ onto the factor of X corresponding to the interval $\langle j, \infty \rangle$ has measure 1. Repeat the argument with right replaced by left and obtain the corresponding result for the projection of $B_r(x_{1,n}) \cap C(i, r, y[i, r])$ to the factor of X corresponding to $\langle -\infty, i \rangle$. Here $i + r \leq -t_0$ and $y \in Y \cap E_0$. Form a word w from $y[i, r]$ and $x[-r, j]$ by identifying $y[-r, r]$ and $x[-r, r]$. Lemma 2.3 now yields that $C(i, j, w) - B_r(x_{1,n})$ is an open set of measure 0 whence $C(i, j, w) \subseteq B_r(x_{1,n})$. But then $f \in \mathcal{A}$ by lemma 3.4(v), not the case. □

Notice that the choice of direction in proposition 5.3(iii) is necessary. The example $f \in \mathcal{B}$ of § 3 has no propagation of changes to the right at all.

Next we analyze propagation of changes for $f \in \mathcal{A}$.

PROPOSITION 5.4. *If $f \in \mathcal{A}$ and $x \in T_\sigma$, then*

- (i) $p_t < 1$ for all t ;
- (ii) *there is t_0 such that $p_t = 0$, $t \geq t_0$, and in fact no change propagates a distance greater than t_0 from the interval I .*

Proof. The proof of proposition 5.3(i) is valid for $f \in \mathcal{A}$ so (i) holds. For (ii) assume I is finite; the other cases are done similarly. Pick a cylinder $C = C(i, j, w)$ contained in the class $B_r(y_{1,n})$ of proposition 5.2. Note that $i \leq -r$ and $j \geq r$. As $x \in T_\sigma$, there are translates $C_1 = C(i_1, j_1, w)$ and $C_2 = C(i_2, j_2, w)$ of C such that $x \in C_1 \cap C_2$ and the intervals $[i_1, j_1]$, $[i_2, j_2]$ lie on each side of I and do not intersect it. It follows that any change of x on I does not affect the behaviour of x on some subinterval of width $2r+1$ in $[i_1, j_1]$ and likewise for $[i_2, j_2]$. Thus no change on I propagates beyond these intervals and (ii) is proved. \square

PROPOSITION 5.5. *If $f \in \mathcal{C}$, then there is a set $Y \subseteq X$ of measure 1 and a direction such that for any interval infinite in the distinguished direction $p_t = 1$ for all t . In particular the probability that a random change will propagate forever is 1.*

Proof. Consider X_R , the set of all $x \in X$ such that for every interval $[i-r, i+r]$ shifting x so that this interval moves to $[-r, r]$ and applying the decomposition of (5.1) with $i = -r, j = r$ yields $\mu(B^+) = 0$. We claim $p_t = 1$ for any $t \in \mathbb{N}_0, x \in X_R$, and right-infinite interval in x . Indeed $p_0 = 1$ by (5.1). If $j > r$ in (5.1), lemma 2.3 says that B^+ is a projection of a subset of the B^+ obtained for $j = r$, and it follows that $p_{j-r} = 1$ for all $j \geq r$.

If $\mu(X_R) = 1$ or if the analogue of X_R for left-infinite intervals has measure 1, then we are done. As both these sets are $\langle \sigma \rangle$ -invariant, we may assume they have measure 0. Thus for almost every $x \in X$ there is an interval $[i-r, i+r]$ which gives $\mu(B^-) > 0$ and another interval $[j-r, j+r]$ which yields $\mu(B^+) > 0$. As there are only countably many possibilities for the pair (i, j) , there is a choice $i = i_0, j = j_0$ for which the set Y_0 of x satisfying the above condition has positive measure. As almost all x have infinitely many translates under $\langle \sigma \rangle$ in Y_0 , we may assume $i+r < 0 < j-r$. Using lemma 2.2 we conclude $\mu(B_r(x)) > 0$ for $x \in Y_0$ whence $f \in \mathcal{B}$, not the case. \square

We end with a discussion of the connection between our classes and the four classes of Wolfram. Wolfram's classes have the following qualitative characterizations for a random input [10, Introduction]:

- (1) Tends to a spatially homogeneous state.
- (2) Yields a sequence of simple stable or periodic structures.
- (3) Exhibits chaotic aperiodic behaviour.
- (4) Yields complicated localized structures, some propagating.

They are also distinguished by the effect of changing the input on a small interval.

- (1) No changes in final state.
- (2) Changes only over a region of finite size.

(3) Changes over a region of ever-increasing size.

(4) Irregular changes.

Classes 1 and 2 have ultimately periodic behaviour and so lie in $\mathcal{A} \cup \mathcal{B}$. We raise the question of whether every $f \in \mathcal{B}$ has ultimately periodic behaviour on every interval $[-n, n]$ for x in a set of measure 1. If so, then class 3 would be in \mathcal{C} . This correspondence is consistent with the propagation of changes in class 3 automata if we assume that changes on finite and infinite intervals have similar effects. (This assumption is supported by some computer experiments.) Likewise the way in which changes are propagated in class 4 automata suggests that they cannot be in \mathcal{A} or \mathcal{C} and so must be in \mathcal{B} .

REFERENCES

- [1] R. L. Adler & B. Marcus. *Topological Entropy and Equivalence of Dynamical Systems*. Mem. Amer. Math. Soc. **20** (# 219), 1979.
- [2] G. Della Riccia. Equicontinuous semi-flows (one-parameter semi-groups) on locally compact or complete metric spaces. *Math. Systems Theory* **4** (1970), 29–34.
- [3] D. Farmer, T. Toffoli & S. Wolfram eds. *Cellular Automata*. North-Holland Pub. Co., Amsterdam 1984.
- [4] W. Feller. *An Introduction to Probability Theory and its Applications, Vol. 1, first edition*. J. Wiley, New York, 1950.
- [5] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton Univ. Press, Princeton, 1981.
- [6] P. Grassberger. Chaos and diffusion in deterministic cellular automata. *Physica* **10D** (1984), 52–58 (reprinted in [3]).
- [7] G. A. Hedlund. Endomorphisms and automorphisms of the shift in dynamical systems. *Math. Systems Theory* **3** (1969), 320–375.
- [8] P. Walters. *An Introduction to Ergodic Theory*. Springer Verlag, New York, 1982.
- [9] S. Wolfram. Universality and complexity in cellular automata. *Physica* **10D** (1984), 1–35 (reprinted in [3]).
- [10] S. Wolfram. Computation theory of cellular automata. *Comm. Math. Phys.* **96** (1984), 15–57.