J. Math. Soc. Japan
Vol. 52, No. 4, 2000

Expansive invertible onesided cellular automata

By Mike Boyle and Alejandro Maass

(Received Mar. 8. 1999)

Abstract. We study expansive invertible onesided cellular automata (i.e., expansive automorphisms of onesided full shifts) and find severe dynamical and arithmetic constraints which provide partial answers to questions raised by M. Nasu **[N2]**. We employ the images and bilateral dimension groups, measure multipliers, and constructive combinatorial characterizations for two classes of cellular automata.

1. Introduction.

Let $\mathscr{A}$ be a set of cardinality N, let $\boldsymbol{N}$ denote the nonnegative integers, and denote an element of $\mathscr{A}^{\boldsymbol{N}}$ as $x = x_0x_1x_2\cdots$. An invertible onesided cellular automaton (c.a.) is a bijection $F : \mathscr{A}^{\boldsymbol{N}} \to \mathscr{A}^{\boldsymbol{N}}$ given by some local rule $f : \mathscr{A}^r \to \mathscr{A}$—for all n, $(Fx)_n = f(x_n \cdots x_{n+r-1})$. The onesided full shift on N symbols, S_N, is the local homeomorphism $\mathscr{A}^{\boldsymbol{N}} \to \mathscr{A}^{\boldsymbol{N}}$ defined by setting $(S_Nx)_i = x_{i+1}$ for all i. F and S_N commute; in the language of symbolic dynamics, F is an automorphism of S_N.

We prove that if F is assumed to be a shift of finite type (which we show follows from weaker assumptions), then F must be shift equivalent to some twosided full shift on J symbols, where the same primes divide J and N, and the maps F and S_N have a common measure of maximal entropy. These results are proved in Section 4 by studying the relationship between the images dimension group of S_N (introduced in [**BFF**]) and the bilateral dimension group of F (introduced in [**Kr1**]), which are reviewed in Section 3.

In Section 6 we prove that if F is assumed to be a shift of finite type and N is a power of a prime p, then the number J above satisfies $J \geq p^2$. The proof uses "measure multipliers" (reviewed in Section 5), developed in [**B**] to generalize Welch's theory [**H**] of compatible extension numbers to shifts of finite type.

In Section 7, we make three conjectures about the possible dynamics of an expansive automorphism of S_N. Two of these were originally introduced by M. Nasu in the form of questions, to which our results give partial answers.

In Section 8 we give a constructive combinatorial characterization of the invertible onesided cellular automata F such that the shortest local rules for F and F^{-1} have radius 1. (Any invertible onesided c.a. is in an obvious way topologically conjugate to such an F.) In Section 9 we use this characterization to develop a constructive combinatorial

---

2000 *Mathematics Subject Classification*. Primary: 37B15; Secondary: 37B10.

Key Words and Phrases. Cellular automata, shift of finite type, expansive, symbolic dynamics, dimension group.

The research of the first author was supported by NSF Grant DMS9706852.

The research of the second author was supported by Cátedra Presidencial fellowship and proyecto fondecyt 1980657.

characterization of a certain class of expansive c.a., and for this class we verify all our conjectures.

In our situation, F is expansive and $S = S_N$ is positively expansive (corresponding to F and not S being invertible). For the case that F and S are both positively expansive, see [**N2**], [**N3**], [**Ku**], [**BM**] and [**BFF**]. For the case that F and S are both expansive, see [**N2**] and [**BL**]. This last case is much less rigid and at present much more mysterious than the others (see Remark 4.7).

The first-named author thanks the Departamento de Ingeniería Matemática of the Universidad de Chile for financial support and warm hospitality in his December 1998 visit to Santiago, which made this collaboration possible.

2. Symbolic background.

In this section we recall some elementary facts about symbolic dynamics. For a thorough introduction to the symbolic dynamics, see [**K2**] or [**LM**].

For a positive integer J, let $\mathscr{J}$ be a set of cardinality J; our default choice will be $\{0, 1, \ldots, J-1\}$. Let Σ_J denote the space $\prod_{n \in \boldsymbol{Z}} \mathscr{J}$. We view a point x in Σ_J as a doubly infinite sequence of symbols from $\mathscr{J}$, so $x = \cdots x_{-1}x_0x_1 \cdots$. The space Σ_J is compact metrizable; one metric compatible with the topology is $\mathrm{dist}(x, y) = 1/(|n| + 1)$, where $|n|$ is the minimum nonnegative integer such that $x_n \neq y_n$. The shift map $\sigma : \Sigma_J \to \Sigma_J$ is the homeomorphism defined by the rule $(\sigma x)_i = x_{i+1}$. The topological dynamical system (Σ_J, σ) is called the *full shift on J symbols* ($\mathscr{J}$ is the symbol set). If Σ is a nonempty compact subset of Σ_J and the restriction of σ to Σ is a homeomorphism, then $(\Sigma, \sigma|_\Sigma)$ is a *subshift*. (We may also refer to either Σ or $\sigma|_\Sigma$ as a subshift, also we may suppress restrictions from the notation.) Equivalently, there is some countable set $\mathscr{W}$ of finite words such that Σ equals the subset of Σ_J in which no element of $\mathscr{W}$ occurs. A subshift (Σ, σ) is a *shift of finite type* (SFT) if it is possible to choose a finite set to be a defining set $\mathscr{W}$ of excluded words.

A homomorphism φ of subshifts is a continuous map between their domains which commutes with the shifts. A homomorphism is N-to-1 if every point in the range space has exactly N preimages. The map is constant-to-1 if it is N-to-1 for some integer N. An endomorphism is a homomorphism from a subshift to itself. (Thus a one-dimensional cellular automaton map is the same thing as an endomorphism of some full shift on J symbols.)

Two continuous maps F and G are topologically conjugate, or isomorphic, if there exists a homeomorphism h such that $Fh = hG$. In this case the map h is a topological conjugacy. A topological conjugacy or isomorphism of subshifts is a bijective homomorphism between them.

Now suppose that X and Y are subshifts, m and a are nonnegative integers (standing for memory and anticipation), Φ is a function from the set of X-words of length $m + a + 1$ into the symbol set for Y, and φ is a homomorphism from X to Y defined by the local rule $\varphi(x)_i = \Phi(x_{i-m} \cdots x_{i+a})$. The homomorphism φ is called a block code (a k-block code if $k = m + a + 1$). The "Curtis-Hedlund-Lyndon Theorem" is that every homomorphism of subshifts is a block code.

If A is an $m \times m$ matrix with nonnegative integral entries, let Graph(A) be a directed graph with vertex set $\{1, \dots, m\}$ and with $A(i, j)$ edges from i to j. Let E_A be the edge set of Graph(A). Let Σ_A be the subset of $(E_A)^{\boldsymbol{Z}}$ obtained from doubly infinite walks through Graph(A); that is, a bisequence x on symbol set E_A is in Σ_A if and only if for every i in $\boldsymbol{Z}$, the terminal vertex of the edge x_i equals the initial vertex of the edge x_{i+1}. Let $\sigma_A = \sigma|_{\Sigma_A}$. The SFT (Σ_A, σ_A) (or Σ_A or σ_A) is called an *edge shift*. The edge shift $\sigma_{[J]}$ is a full shift σ_J. Any SFT is topologically conjugate to some edge SFT.

Let X_A be the space of onesided sequences obtained by erasing negative coordinates in Σ_A: that is, if a point x is in Σ_A, then the onesided sequence $x_0 x_1 x_2 \cdots$ is in X_A, and X_A contains only such points. The shift map rule $(\sigma x)_i = x_{i+1}$ defines a surjective local homeomorphism $S_A : X_A \to X_A$. The system (X_A, S_A) is a *onesided shift of finite type*. We will write a onesided full shift on N symbols as $(X_{[N]} = X_N, S_N)$, where $X_N = \mathscr{A}^{\boldsymbol{N}}$ for some alphabet $\mathscr{A}$ with N symbols. If X is a nonempty compact subset of X_N and the restriction of S_N to X_N is an endomorphism, then (X, S_N) is a onesided subshift. Every homomorphism of onesided subshifts is given by a $r + 1$-block code or local rule with memory $m = 0$.

Williams [**W**] explained how to associate to a onesided SFT S_A an essentially canonical matrix, the "total amalgamation" of A. Correspondingly, onesided SFTs are much more rigid than twosided SFTs ([**W**], [**BFK**], [**N2**], [**BFF**]). One striking result in this direction is due to Nasu: if F is a totally transitive (every power of F has a dense orbit) automorphism of a onesided SFT S, then S must be topologically conjugate to a onesided full shift ([**N2**], Thm. 3.12). [This is essentially because F induces a map on the vertices of the total amalgamation (by Lemma 3.10 of [**N1**], or (2.23) and (2.25) of [**BFK**]), so some power of F fixes those vertices.] This result of Nasu justifies a focus on the dynamics of automorphisms of S_N vs. other onesided SFTs.

An SFT is called irreducible if it has a dense forward orbit, and it is mixing if whenever words U and W occur in points of the SFT, for all but finitely many positive n there is a word V of length n such that UVW occurs. A nonnegative matrix A is irreducible if for every i, j there exists $n > 0$ such that $A^n(i, j) > 0$, and it is primitive if n can be chosen independent of (i, j). An irreducible matrix A defines an edge shift which is an irreducible SFT, and a primitive matrix A defines an edge shift which is a mixing SFT.

The (topological) entropy $h(\sigma)$ of a subshift (X, σ) (twosided or onesided) is the growth rate of its words, that is, $\lim (1/n) \log \#\{x_1 x_2 \cdots x_n : x \in X\}$. For an SFT σ_A, the entropy is $\log(\lambda_A)$, where λ_A is the spectral radius of A. An irreducible SFT σ_A has a unique measure μ of maximal entropy; that is, μ is a σ_A-invariant Borel measure, its measure theoretic entropy equals $h(\sigma_A)$, and there is no other such measure.

Twosided SFTs σ_A, σ_B are *shift equivalent* if their defining matrices satisfy certain equations which are equivalent to the following condition: for all sufficiently large k, $(\sigma_A)^k$ and $(\sigma_B)^k$ are topologically conjugate. An SFT σ_A is shift equivalent to a full shift on J symbols iff for some k the characteristic polynomial of A equals $x^k(x - J)$. It is still unknown whether such SFTs must be conjugate to full shifts (contrast [**KR**]).

A continuous map φ from a compact metric space X to itself is *positively expansive* if there exists $\varepsilon > 0$ such that whenever x and x' are distinct points in X, there is a nonnegative integer k such that $\mathrm{dist}(\varphi^k(x), \varphi^k(x')) > \varepsilon$. This property does not depend on the choice of metric compatible with the topology. Now if φ is an endomorphism of a twosided subshift Σ and $k \in \boldsymbol{Z}_+$, then let $\hat{x}^{(k)}$ denote the sequence of words $([\varphi^i(x)_{-k} \cdots \varphi^i(x)_k] : i = 0, 1, 2 \ldots)$. It is easy to check that φ is positively expansive iff there exists $k \in \boldsymbol{Z}_+$ such that the map $x \mapsto \hat{x}^{(k)}$ is injective iff φ is conjugate to a onesided subshift.

Similarly, a homeomorphism F from a compact metric space to itself is *expansive* if there is some $\varepsilon > 0$ such that for all distinct x, x', there is an integer k such that $\mathrm{dist}(F^k(x), F^k(x')) > \varepsilon$. Expansiveness is an important ([**Hi**], [**AH**]) and multifaceted ([**BL**], Sec. 5) dynamical property. If F is an automorphism of a onesided subshift X and $k \in \boldsymbol{N}$, let $\tilde{x}^{(k)}$ denote the sequence of words $([F^i(x)_0 \cdots F^i(x)_k] : i \in \boldsymbol{Z})$. Then F is expansive iff there exists $k \in \boldsymbol{N}$ such that the map $\eta_k : x \mapsto \tilde{x}^{(k)}$ is injective iff F is conjugate to a twosided subshift. If F is expansive with local rule $f : x_0 \cdots x_r \mapsto (Fx)_0$, $r > 0$, then the map η_{r-1} must be injective. (If η_{r-1} collapses x and y, and $W = W_0 \cdots W_j$, then one easily checks that η_{r+j} collapses Wx and Wy.)

3. Two dimension groups.

In this section we review the information we will need on two dimension groups arising in symbolic dynamics.

Let S be a local homeomorphism of a compact zero dimensional metrizable space X. Let $CO(X)$ be the collection of clopen subsets of X. Let $\boldsymbol{Z}CO(X)$ denote the free abelian group with generators $CO(X)$. Let $H(S)$ be the subgroup of $\boldsymbol{Z}CO(X)$ generated by the following relations:

(i) $\Sigma C_i \sim C$ if C is the disjoint union of the clopen sets C_i.

(ii) $C \sim D$ if C and D are clopen sets and there exists $n > 0$ such that $\varphi^n|_C$ and $\varphi^n|_D$ are injective and $\varphi^n C = \varphi^n D$.

The *images group* $\mathrm{Im}(S)$ defined in [**BFF**] is the quotient $\boldsymbol{Z}CO(X)/H(S)$. To an $n \times n$ integral matrix A, associate the direct limit group

$$G(A) = \varinjlim_{A} \boldsymbol{Z}^n.$$

The group $G(A)$ can be presented concretely as a subgroup of a finite dimensional vector space (see pp. 14–15 of [**BMT**] and Sec. 7.5 of [**LM**]). For a onesided SFT S_A, we have $\mathrm{Im}(S_A) \cong G(A)$ ([**BFF**], Thm. 4.5). For $S = S_N$, let μ be the uniform measure (the measure of maximal entropy), i.e. if $x[0,k]$ denotes $\{y \in X_N : y_i = x_i, 0 \le i \le k\}$, then $\mu(x[0,k]) = N^{-(k+1)}$. Then there is an isomorphism $\mathrm{Im}(S_N) \to \boldsymbol{Z}[1/N]$ given by $[\Sigma n_i C_i] \mapsto \Sigma n_i \mu(C_i)$.

Next, let $F : X \to X$ be a subshift. We will similarly define the *bilateral dimension group* $\mathrm{Bilat}(F)$ as a quotient of $\boldsymbol{Z}CO(X)$. Let $K(F)$ be the subgroup of $\boldsymbol{Z}CO(X)$ generated by the following relations:

(i) $\Sigma C_i \sim C$ if C is the disjoint union of the clopen sets C_i.

(ii) $x[i,j] \sim y[i,j]$ if for all sequences $w = \cdots w_{i-2} w_{i-1}$ and $z = z_{i+1} z_{j+2} \ldots$, the point $w x_i \cdots x_j z$ is in X if and only if the point $w y_i \cdots y_j z$ is in X.

Then $\mathrm{Bilat}(F)$ is the quotient group $\boldsymbol{Z}CO(X)/K(F)$. This is one of the dimension groups introduced to symbolic dynamics by Krieger [**Kr1**]. The definition of Bilat(F) in ([**Kr1**], Section 2) appeals to a larger theory; we have given an equivalent but more direct definition suitable to our needs. Following Krieger (personal communication), we use the adjective "bilateral" to distinguish this dimension group from the past and future dimension groups of [**Kr1**], which involve onesided splittings.

The groups Im(S) and Bilat(F) carry order structures which make them dimension groups. We will not need to consider these order structures in the current paper.

If F is SFT and S is a local homeomorphism such that $SF = FS$, then there is an induced homomorphism $S_* : \mathrm{Bilat}(F) \to \mathrm{Bilat}(F)$, given by $[\Sigma n_i C_i] \mapsto [\Sigma n_i S(C_i)]$ when the restriction of S to each C_i is injective. If F is an SFT σ_A, with A $n \times n$, then from Prop. 3.1 of [**Kr1**] we have

$$\mathrm{Bilat}(F) \cong G(A \otimes A^t). \tag{3.1}$$

Here the tensor product $A \otimes A^t$ is an $n^2 \times n^2$ matrix with $(A \otimes A^t)([i,j],[i',j']) = A(i,i')A^t(j,j')$. When σ_A is shift equivalent to a full shift σ_J, this means

$$\mathrm{Bilat}(F) \cong \boldsymbol{Z}[1/J]. \tag{3.2}$$

Let G be a torsion free abelian group. We will say G has finite rank if it is isomorphic to a subgroup of $\boldsymbol{Q}^k$ for some $k < \infty$. In this case, the rank of G is the minimal such k. If $\alpha : H \to G$ is a group homomorphism, where H has rank k and G has rank ℓ, then (after identifying G and H with subgroups of $\boldsymbol{Q}^k$ and $\boldsymbol{Q}^\ell$) it is easily checked that the map α is the restriction of a unique rational vector space homeomorphism $\tilde{\alpha} : \boldsymbol{Q}^k \to \boldsymbol{Q}^\ell$. Consequently, we have the following well known

FACT 3.1. *Suppose H and G are countable torsion free abelian groups of equal finite rank and $\alpha : H \to G$ is a surjective group homomorphism. Then α is injective, and therefore an isomorphism.*

For any subshift (Σ, σ), let $\mathscr{W}(\sigma)$ denote the set of words $\{x_i \cdots x_j : x \in \Sigma\}$. Define an equivalence relation $\approx$ on $\mathscr{W}(\sigma)$ by setting $V \approx V'$ iff for all words U and W,

$$UVW \in \mathscr{W}(\sigma) \quad \Leftrightarrow \quad UV'W \in \mathscr{W}(\sigma).$$

A subshift is *sofic* iff the set of $\approx$ equivalence classes is finite. It is easy to see that for F sofic, the rank of Bilat(F) is finite.

PROPOSITION 3.2. *Suppose σ is sofic and rank* Bilat$(\sigma) = 1$. *Then σ is* SFT *and σ is shift equivalent to some full shift.*

PROOF. If σ is isomorphic to an SFT σ_A, then by equation (3.1) the matrix A has just one nonzero eigenvalue, and therefore σ_A is shift equivalent to a full shift.

To deduce that σ must be SFT we will sketch an argument which requires some familiarity with sofic systems. There is an SFT σ_B which is a topologically canonical "follower set cover" of the sofic shift σ, and here $G(B)$ is isomorphic to the future dimension group of σ ([**Kr2**], Thm. 3.5). It is easy to check that the future dimension group of σ must have rank 1 if Bilat(σ) has rank 1. This forces the SFT σ_B to be irreducible. However, for a nonSFT sofic shift, this SFT σ_B must be reducible, because a nonSFT sofic shift has "non-F-finitary" points ([**Kr2**], Prop. 4.3). □

OPEN PROBLEM 3.3. *Suppose σ is a subshift and* Bilat(σ) *has finite rank. Must σ be sofic*?

OPEN PROBLEM 3.4. *Suppose σ is a subshift and the rank of* Bilat(σ) *is* 1. *Must σ be* SFT?

Of course an answer yes to the former problem implies an answer yes to the latter.

4. Full shifts and primes.

Throughout this section, let $S = S_N$ denote the onesided full shift on N symbols with domain $X = X_N = \prod_{i=0}^{\infty}\{0, 1, \dots, N-1\}$, and let F be an expansive automorphism of S. We will use some dimension group techniques to prove that if F is SFT, then F is shift equivalent to some twosided full shift on J symbols, σ_J, where N and J are divisible by the same primes.

Let $[i_0 i_1 \cdots i_r]_S$ denote $\{x \in X : x_j = i_j, 0 \le j \le r\}$. Any clopen set in X is a finite union of sets of this form. Let S_* denote the homomorphism $\mathrm{Bilat}(F) \to \mathrm{Bilat}(F)$ induced by $[C] \to [SC]$, when $S|_C$ is injective, as described in the previous section.

PROPOSITION 4.1. (i) *The homomorphism* $S_* : \mathrm{Bilat}(F) \to \mathrm{Bilat}(F)$ *is surjective.*

(ii) *If S_* is injective then* $\mathrm{Bilat}(F) \cong \boldsymbol{Z}[1/N]$.

(iii) *If* Bilat(F) *has finite rank* (*in particular, if F is* SFT *or sofic*), *then S_* is an isomorphism.*

PROOF. (i) Every clopen set is a disjoint union of sets of the form $C = [i_0 \cdots i_r]_S$, so every element of Bilat(F) has the form $[\Sigma n_C C]$. Surjectivity of S_* then follows from the observation

$$S_* : [[0 i_0 \cdots i_r]_S] \mapsto [[i_0 \cdots i_r]_S].$$

(ii) Suppose S_* is injective. Then any two S-cylinders of equal length define equivalent elements of Bilat(F) (that is, $[[i_0 \cdots i_{r-1}]_S] = [[i'_0 \cdots i'_{r-1}]_S]$), because

$$(S_*)^r[[i_0 \cdots i_{r-1}]_S] = [X] = (S_*)^r[[i'_0 \cdots i'_{r-1}]_S].$$

For every positive integer r, X is the disjoint union of N^r S-cylinders of length r. Now there is a unique homomorphism $\tau : \mathrm{Bilat}(F) \to \boldsymbol{Q}$ such that $\tau([X]) = 1$; this injective map τ sends each $[[x_0 \cdots x_{r-1}]]$ to N^{-r}. The image of this map is $\boldsymbol{Z}[1/N]$.

(iii) If Bilat(F) has finite rank, then by the Fact 3.1, the surjection S_* must be an isomorphism. □

COROLLARY 4.2. *If F is sofic* (*in particular, if F is* SFT), *then F is an* SFT *which is shift equivalent to some two-sided full shift σ_J.*

PROOF. If F is sofic, then Bilat(F) has finite rank. Then by Proposition 4.1, the rank of Bilat(F) is 1. It follows from Proposition 3.2 that F is an SFT shift equivalent to a full shift. □

REMARK 4.3. Let J be the integer such that F is shift equivalent to σ_J. Then S_N may be viewed as an N-to-1 endomorphism of some power of σ_J, and it follows from

Welch's theorem ([**H**], Theorem 14.9) or its generalizations ([**B**], [**T1**], [**T2**]) that every prime dividing N must also divide J. But to show the *same* primes divide N and J requires more, since (for example) the rule $(\varphi x)_n = 3x_n + x_{n+1} \pmod 6$ defines a 2-to-1 but not positively expansive endomorphism φ of σ_6.

PROPOSITION 4.4. (i) *The identity map on clopen subsets of X induces a group epimorphism* $\mathrm{Bilat}(F) \to \mathrm{Im}(S)$.

(ii) *This epimorphism is an isomorphism if the rank of* $\mathrm{Bilat}(F)$ *is finite.*

PROOF. Recall from Section 3 the definitions

$$\mathrm{Bilat}(F) = \boldsymbol{Z}CO(X)/K(F),$$

$$\mathrm{Im}(S) = \boldsymbol{Z}CO(X)/H(S).$$

To prove (i), we will prove that $H(S)$ contains $K(F)$. Suppose not. Then there is a formal sum $\Sigma n_i C_i$ in $\boldsymbol{Z}CO(X)$ which lies in $K(F)$ but not in $H(S)$. Using the subdivision relation common to both $K(F)$ and $H(S)$, after passing to a different sum we may assume each C_i is an S-cylinder of the same length, r. Now $\Sigma n_i \neq 0$ because $\Sigma n_i C_i \notin H(S)$. But then $(S_*)^r : [\Sigma n_i C_i] \mapsto \Sigma n_i [X] \neq 0$. This contradicts $\Sigma n_i C_i \in K(F)$, and finishes the proof of (i).

If the rank of $\mathrm{Bilat}(F)$ is finite, then by Proposition 4.1 this rank is 1, and the surjective homomorphism of rank 1 groups $\mathrm{Bilat}(F) \twoheadrightarrow \mathrm{Im}(S)$ must be an isomorphism. □

THEOREM 4.5. *If F is sofic, then F is* SFT *and F is shift equivalent to some σ_J, a full shift on J symbols, where J and N are divisible by the same primes.*

PROOF. After Corollary 4.2, it remains to show that J and N are divisible by the same primes. However, $\mathrm{Bilat}(F) \cong \boldsymbol{Z}[1/J]$ and $\mathrm{Im}(S) \cong \boldsymbol{Z}[1/N]$. Now the result follows from Proposition 4.4. □

THEOREM 4.6. *Suppose F is* SFT. *Then F and S have the same measure of maximal entropy.*

PROOF. In this case we have $\mathrm{Bilat}(F) \cong \mathrm{Im}(S)$ as quotients of $\boldsymbol{Z}CO(X)$, and $\mathrm{Im}(S) \cong \boldsymbol{Z}[1/N]$. There is a unique homomorphism τ from this group into $\boldsymbol{R}$ which sends $[X]$ to 1. Let μ_S and μ_F denote the measures of maximal entropy for S and F. For $[C]$ in $\mathrm{Im}(S)$, it is known that $\tau : [C] \mapsto \mu_S(C)$ ([**BFF**], Sec. 9). Likewise for $[C]$ in $\mathrm{Bilat}(F)$, since F is irreducible SFT, $\tau : [C] \mapsto \mu_F(C)$ ([**Kr1**], Theorem 3.2). Therefore $\mu_S = \mu_F$.

REMARK 4.7. A similar dimension group proof scheme was used ([**BFF**], Theorem 9.1) to show that commuting onesided mixing SFT's have the same measure of maximal entropy. In contrast, Nasu ([**N2**], Sec. 10) has given an example of commuting twosided mixing SFTs σ_A and σ_B such that $\boldsymbol{Q}(\lambda_A) \neq \boldsymbol{Q}(\lambda_B)$. This implies that the measures of maximal entropy for σ_A and σ_B do not assume the same set of values on clopen sets, and therefore are not equal. For this example, in addition $G(A)$ and $G(B)$ do not even have the same rank.

5. Multipliers.

In this section we give background for the "multipliers" we use for the entropy constraints of the next section.

Let A be an irreducible nonnegative integral matrix with spectral radius $\lambda > 1$. Let $\sigma_A : \Sigma_A \to \Sigma_A$ be the associated edge SFT. Let u and v be positive vectors such that $uA = \lambda u$, $Av = \lambda v$ and $uv = 1$. For a symbol (edge) e, we let $u(e)$ denote u_i where i is the initial vertex of e, and similarly $v(e) = v_j$ where j is the terminal vertex of e.

Let $x[a,b] = \{y \in \Sigma_A : y_i = x_i, a \le i \le b\}$. Then the measure μ of maximal entropy for the edge SFT σ_A is determined by

$$\mu(x[a,b]) = u(x_a)\lambda^{-(b-a+1)}v(x_b).$$

For $x = \cdots x_{-1}x_0x_1\cdots$ in Σ_A, let $\pi_-(x) = \cdots x_{-2}x_{-1}$ and let $\pi_+(x) = x_0x_1\cdots$. Let $[e]$ denote $\{x \in \Sigma_A : x_0 = e\}$. Let $[e]_+ = \pi_+[e]$ and $[e]_- = \pi_-[e]$. Then $[e] = [e]_- \times [e]_+$.

On $[e]_-$ determine a Borel measure μ_-^e by setting

$$\mu_-^e(x[-n,-1]) = u(x_{-n})\lambda^{-n}$$

and on $[e]_+$ determine a Borel measure μ_+^e by

$$\mu_+^e(x[0,n]) = \lambda^{-(n+1)}v(x_n).$$

Define the measure μ^e as the restriction of μ to $[e]$. (Abusing notation: $\mu = \sum_e \mu^e$). For any Borel set C in X_A, define

$$\mu_-(C) = \sum_e \mu_-^e(C \cap [e]),$$

$$\mu_+(C) = \sum_e \mu_+^e(C \cap [e]).$$

Then for any Borel set C,

$$\mu_-(\sigma_A C) = \frac{1}{\lambda}\mu_-(C),$$

$$\mu_+(\sigma_A C) = \lambda\mu_+(C).$$

We will call (μ_-, μ_+) a *conditional decomposition* of μ. (The conditional decomposition is a simplified version of ideas from the ergodic theory of smooth hyperbolic systems [**Ma**], [**S**]. The measures μ_-^e and μ_+^e can be viewed as conditional measures obtained from μ^e, related closely to the conditional measures on stable and unstable sets discussed in Sec. 3 of [**B**].)

We made a concrete choice of (μ_-, μ_+) with respect to a particular Markov partition $\{[e]\}$ and choice of associated eigenvectors u, v. We could have obtained another conditional decomposition of μ with another Markov partition. We could also have obtained a conditional decomposition above using the vertex sets $\{x \in X_A :$ the initial vertex of x_0 is $i\}$ in place of the edge sets $[e]$. In any of these cases, for the resulting conditional decomposition $(\tilde\mu_-, \tilde\mu_+)$ of μ, one can see from Prop. 3.2 of [**B**] that there will be a constant $c > 0$ such that

$$\tilde{\mu}_- = c\mu_-, \quad \tilde{\mu}_+ = \left(\frac{1}{c}\right)\mu_+.$$

For the special case that $\sigma_A = \sigma_J$, a full shift on J symbols, we have an especially simple choice of (μ_-, μ_+):

$$\mu_-(x[-n,-1]) = J^{-n},$$

$$\mu_+(x[0,n]) = J^{-(n+1)}.$$

We can now summarize the background we need.

THEOREM 5.1. *Suppose σ_A is an irreducible* SFT, *A has spectral radius $\lambda > 1$, and φ is an N-to-*1 *local homeomorphism commuting with σ_A. Let (μ_-, μ_+) be a conditional decomposition of the measure of maximal entropy μ of σ_A. Let C be any nonempty clopen set such that the restriction of φ to C is injective.*

(i) *The ratios $\ell_\varphi = (\mu_-(\varphi C))/(\mu_-(C))$, $r_\varphi = (\mu_+(\varphi C))/(\mu_+(C))$ do not depend on C or the particular choice of conditional decomposition.*

(ii) $\ell_\varphi r_\varphi = N$.

(iii) $\mu(\varphi C) = N\mu(C)$.

(iv) *The numbers ℓ_φ and r_φ are units in the ring $\mathcal{O}_\lambda[1/\lambda]$. (If λ is an integer J, then this means the numbers ℓ_φ and r_φ are products of integral powers of primes dividing J.)*

(v) *If ψ is also a local homeomorphism commuting with σ_A, then $\ell_{\varphi\psi} = \ell_\varphi \ell_\psi$ and $r_{\varphi\psi} = r_\varphi r_\psi$.*

(vi) *Suppose V is homeomorphism such that $V\varphi V^{-1} = \tilde{\varphi}$ and $V\sigma_A V^{-1} = \sigma_{\tilde{A}}$, so $\ell_{\tilde{\varphi}}$ and $r_{\tilde{\varphi}}$ may be computed w.r.t. $\sigma_{\tilde{A}}$. Then $\ell_\varphi = \ell_{\tilde{\varphi}}$ and $r_\varphi = r_{\tilde{\varphi}}$.*

REMARK 5.2. If σ_A is an irreducible SFT and φ is a continuous map such that $\varphi\sigma_A = \sigma_A\varphi$, then φ is open iff φ is constant to 1 iff φ is a local homeomorphism (this is contained in Nasu's Theorem 6.5 in [**N1**]).

EXAMPLE 5.3. Let $\sigma_A = \sigma_J$, the full shift on J symbols. If N is a positive integer dividing a power of J, and ℓ, r are units on $\boldsymbol{Z}[1/J]$ such that $\ell r = N$, then there exists an N-to-1 endomorphism φ of σ_J with $\ell_\varphi = \ell$, $r_\varphi = r$ (this can be proved by the method of Proposition 7.4). Here are a few examples.

(1) $\varphi = \sigma$. Then $\ell_\varphi = 1/J$, $r_\varphi = J$, $N = 1$.

(2) $(\varphi x)_n = x_{n-1} + x_{n+2} \pmod J$. Then $\ell_\varphi = J$, $r_\varphi = J^2$, $N = J^3$.

(3) $(\varphi x)_n = x_{n+1} + x_{n+3} \pmod J$. Then $\ell_\varphi = 1/J$, $r_\varphi = J^3$, $N = J^2$.

(4) $J = 6$ and $(\varphi x)_n = 3x_n + x_{n+1} \pmod 6$. Then $\ell_\varphi = 1/3$, $r_\varphi = 6$, $N = 2$.

REMARK 5.4. For σ_A a full shift, everything in Theorem (5.1) is a translation of (some of) Welch's results as reported by Hedlund (Sections 14–15 of [**H**]). Those results have combinatorial statements which do not generalize to arbitrary irreducible SFT's. The generalizations, using measures, are in Sec. 3 of [**B**]. The results in [**B**] like Welch's are more general than the consequences collected in Theorem 5.1, but proofs for Theorem 5.1 can be obtained from the results of [**B**]. Caveat: our numbers ℓ_φ, r_φ are the reciprocals of the numbers $L(\varphi)$, $R(\varphi)$ used in [**B**].

Remark 5.5. In the next section we only need consider endomorphisms φ of σ_A in the case that σ_A is shift equivalent to some full shift. This means that for large enough $k, (\sigma_A)^k$ is conjugate to a full shift. The multipliers ℓ_φ, r_φ obtained by considering φ an endomorphism of σ_A are the same as those obtained by considering φ as an endomorphism of $(\sigma_A)^k$, so here one may recover the full shift description of the multipliers by passing to a power of σ_A.

6. Entropy and primes.

Below, S is the full onesided shift on N symbols and F is an automorphism of S such that F is conjugate to an irreducible SFT σ_A; that is, there is a homeomorphism U such that $UFU^{-1} = \sigma_A$. Then $USU^{-1} = \varphi$ is an N-to-1, positively expansive local homeomorphism commuting with σ_A. We regard (σ_A, φ) as simply another presentation of (F, S). By Theorem 5.1(vi), the numbers ℓ_φ, r_φ do not depend on the choice of σ_A and U, so we may define $\ell_S = \ell_\varphi$, $r_S = r_\varphi$.

Theorem 6.1. *Let F be an automorphism of S, the onesided full shift on N symbols with $N > 1$. Suppose F is conjugate to some* SFT *σ_A. Then $\ell_S > 1$ and $r_S > 1$. If N is a power of a prime p, then so are ℓ_S and r_S.*

Proof. Let $X_N = \prod_{n\geq 0}\{0, 1, \dots, N-1\}$, so F and S act on X_N. Let $[i_0 \cdots i_r]_S$ denote $\{x \in X_N : x_j = i_j, 0 \leq j \leq r\}$. Then $[i_0 \cdots i_r]_S \mapsto [0 i_0 \cdots i_r]_S$ induces a map γ_0 on clopen sets, $\gamma_0 C = (S^{-1}C) \cap [0]_S$. Now view γ_0 in the (σ_A, φ) presentation and consider C any nonempty clopen set. For $k > 0$, the restriction of φ^k to $\gamma_0^k C$ is injective, and $\varphi^k(\gamma_0^k C) = C$. Because the diameter of $\gamma_0^k C$ goes to zero as k goes to ∞, for large k we must have

$$\frac{\mu_+(\gamma_0^k C)}{\mu_+(C)} < 1,$$

and therefore

$$(r_\varphi)^k = \frac{\mu_+(\varphi^k(\gamma_0^k C))}{\mu_+(\gamma_0^k C)} = \frac{\mu_+(C)}{\mu_+(\gamma_0^k C)} > 1,$$

so $r_\varphi > 1$. Similarly $\ell_\varphi > 1$. For N a power of a prime p, Theorem 5.1(iv) now implies that ℓ_φ and r_φ must be positive integers divisible by p. □

Corollary 6.2. *Suppose S is the onesided full shift on N symbols, such that N is a power of a prime p and S has some automorphism F which is conjugate to a sofic shift. Then p^2 divides N.*

Proof. By Corollary 4.2, F must be conjugate to an SFT σ_A such that σ_A is shift equivalent to a full shift on J symbols, where J is also a power of p. By Theorem 6.1, the numbers ℓ_S, r_S must be positive powers of p. By Theorem 5.1(iii), $N = \ell_S r_S$, so p^2 divides N. □

So for example, no automorphism of the onesided full shift on 5 symbols can be conjugate to an SFT (or sofic shift).

7. Conjectures.

We will make three conjectures. Let S_N be the onesided full shift on N symbols.

Conjecture 7.1. Suppose F is an expansive automorphism of S_N. Then F is SFT.

Conjecture 7.2. Suppose F is an automorphism of S_N and F is conjugate to an SFT. Then F is conjugate to a full shift.

Conjecture 7.3. Suppose there exists an automorphism F of $S_N = S$, such that F is conjugate to an SFT and p is a prime dividing N. Then p divides both l_S and r_S, and in particular p^2 divides N.

The conjectures 7.1, 7.2 are possibilities introduced by Nasu as questions (Question 3.a(2), p. 46, and Question 3.b, p. 50, in [**N2**]). Independently, Nasu [**N2**] and Shereshevsky and Afraimovich [**SA**] proved Conjecture 7.2 in the case that (F, S) is conjugate to a pair (σ_J, ϕ) such that ϕ is given by a local rule $x_i \cdots x_j \mapsto (\phi x)_0$ such that i, j are positive integers and the local rule is bipermutive. Our Theorems 4.5 and 6.1 support 7.1 and 7.2. From Theorem 9.2 it follows that all the conjectures hold for the class of expansive automata we consider in Section 9 (those satisfying $r(F) = r(F^{-1}) = \tilde{r}(F) = 1$). The rigid combinatorics underlying Theorem 9.2 suggest rigid combinatorics in general, as in the algebraic expansive situation considered by Kitchens [**K1**]. Therefore we elevate Nasu's questions to conjectures, and add the third conjecture.

The conjectures together with our results imply the following: if F is an expansive homeomorphism of S_N, then F is conjugate to a full shift σ_J, such that

(i) J and N are divisible by the same primes, and

(ii) if p is a prime dividing N, then p^2 divides N.

Proposition 7.4. *Suppose J and N satisfy the conditions* (i) *and* (ii) *above. Then there is an automorphism F of S_N such that F is conjugate to σ_J.*

Proof. First suppose $N = p^k$ and $J = p^\ell$, with $\ell \geq 1$ and $k \geq 2$. Define an N-to-1 endomorphism φ of σ_p by the rule $(\varphi_x)_n = x_{-1} + x_{k-1} \pmod p$. Then φ is conjugate to S_N and it commutes with $(\sigma_p)^\ell$ which is isomorphic to σ_J.

Next suppose $N = p_1^{k(1)} \cdots p_t^{k(t)}$, with each $k(i) \geq 2$, and $J = p_1^{\ell(1)} \cdots p_t^{\ell(t)}$ with each $\ell(i) \geq 1$. Let $N(i) = p_i^{k(i)}$ and $J(i) = p_i^{\ell(i)}$. Construct F_i as in the preceding paragraph, with F_i an automorphism of $S_{N(i)}$ and F_i conjugate to $\sigma_{J(i)}$. Now $F = \prod_{i=1}^t F_i$ is an automorphism of $S = \prod_{i=1}^t S_{N(i)}$. Here S is conjugate to S_N, and F is conjugate to σ_J. □

8. Construction of one-sided invertible cellular automata.

Let $F : \mathscr{A}^{\mathbf{N}} \to \mathscr{A}^{\mathbf{N}}$ be a onesided invertible cellular automaton. Let $r(F)$ denote the radius of the shortest local rule defining F: that is, $r(F)$ is the minimal nonnegative integer r such that for all x, the symbol $(Fx)_0$ is determined by $x_0 \cdots x_r$; analogously we define $r(F^{-1})$. In this section we give a constructive combinatorial characterization of the invertible F such that $r(F) = r(F^{-1}) = 1$. We remark, it is possible to have $r(F^{-1})$ larger than $r(F)$.

For any positive integer k, by grouping symbols into k-blocks we can view F as a map $F^{(k)} : (\mathscr{A}^k)^{\mathbf{N}} \to (\mathscr{A}^k)^{\mathbf{N}}$; and if $0 < r(F) \le k$, then $r(F^k) = 1$. So any onesided invertible c.a. is topologically conjugate to a c.a. F such that $r(F) = r(F^{-1}) = 1$.

Let $\mathscr{A}^*$ be the set of words appearing in $\mathscr{A}^{\mathbf{N}}$. We say that a word $w' \in \mathscr{A}^*$ is an F-successor of $w \in \mathscr{A}^*$, and we use the shorthand $w \xrightarrow{F} w'$, if there are $x, x' \in \mathscr{A}^{\mathbf{N}}$ such that $F(wx) = w'x'$. From F we define an equivalence relation on the alphabet $\mathscr{A}$: $a \sim_F b$ if and only if $\exists n \in \mathbf{N}$, $\exists a_0, a_1, \dots, a_n \in \mathscr{A}$, $\exists b_0, \dots, b_{n-1} \in \mathscr{A}$, such that $a_0 = a$, $a_n = b$, $a_i \xrightarrow{F} b_i$ and $a_{i+1} \xrightarrow{F} b_i$, for $i \in \{0, \dots, n-1\}$. For $a \in \mathscr{A}$, we denote by $c_F(a)$ its equivalence class with respect to $\sim_F$. This gives a partition of $\mathscr{A}$ into equivalence classes, $\mathscr{P}_F = \{c_F(a) : a \in \mathscr{A}\}$. Given F with $r(F) = 1$, define $\pi_F : \mathscr{A} \to \mathscr{A}$ by $\pi_F(a) = F(aa)$.

With $r(F) = 1$, the map F is *left permutive* if for all α and b in $\mathscr{A}$, there is a unique a in $\mathscr{A}$ such that $F(ab) = \alpha$. In this case we let $\pi_{F,b}$ denote the permutation $a \mapsto F(ab)$. For simplicity we will abbreviate $\pi_{F,b}$ as π_b when the context is clear.

Lemma 8.1. *Suppose F is a onesided invertible c.a. such that $r(F) = r(F^{-1}) = 1$. Then the following hold.*

1. *F is left permutive and π_F is a permutation.*
2. If $c_F(a) = c_F(b)$, *then* $\pi_a = \pi_b$.
3. *$\pi_b(c_F(a)) = \pi_F(c_F(a))$, for all a, b.*

Proof. (1) F is left permutive because F is surjective. The map π_F describes the map on fixed points of the shift, so the invertibility of F implies π_F is a permutation.

(2) It suffices to show $F(ca) = F(cb)$ under the assumption $a \xrightarrow{F} \alpha$ and $b \xrightarrow{F} \alpha$, where α is some element of $\mathscr{A}$. Let $x, x', y, y' \in \mathscr{A}^{\mathbf{N}}$ such that $F(ax) = \alpha y$ and $F(bx') = \alpha y'$. Since $r(F) = r(F^{-1}) = 1$ and $F(cax) = \beta\alpha y$, we deduce that $F^{-1}(\beta\alpha) = c$. Therefore, $F^{-1}(\beta\alpha y') = cbx'$ and $F(cb) = \beta$, which proves the lemma.

(3) Put $\alpha = F(ab)$. Since F is invertible there is a unique $a' \in \mathscr{A}$ such that $\pi_F(a') = \alpha$. Then α is a common successor of a and a', which implies that $a \sim_F a'$. This fact proves that $\alpha \in \{\pi_F(a') : a' \in c_F(a)\}$. □

Theorem 8.2. *Let $\mathscr{P}$ be a partition of $\mathscr{A}$ into equivalence classes, with $c(a)$ denoting the class containing a. Let $F : \mathscr{A}^N \to \mathscr{A}^N$ be a cellular automaton with $r(F) = 1$ satisfying the following conditions*:

1. *F is left permutive and π_F is a permutation.*
2. *If $c(b') = c(b)$, then $\pi_b = \pi_{b'}$.*
3. *$\pi_b(c(a)) = \pi_F(c(a))$, for all a, b.*

Then F is an invertible cellular automaton with $r(F^{-1}) = 1$.

Conversely, if F is a onesided invertible cellular automaton with $r(F) = r(F^{-1}) = 1$, then the properties (1)–(3) *hold for the partition $\mathscr{P}_F$.*

Proof. Let $a_0, a_1 \in \mathscr{A}$. Using the assumption that π_F is a permutation, we let b be the symbol such that $\pi_F(b) = a_1$. Using the left permutivity, we let a be the unique symbol such that $F(ab) = a_0$.

Now if b' is any symbol in $\mathscr{A}$ such that $b' \xrightarrow{F} a_1$, then by property (3) we have that a_1 is in $\pi_F(b')$. Then $\pi_F(c(b))$ and $\pi_F(c(b'))$ have nonempty intersection. Since π_F is a permutation, it follows that $c(b') = c(b)$. Then by property (2), $F(ab) = F(ab') = a_0$.

Because F is left permutive, it follows that a is the unique element in $\mathscr{A}$ such that $a \xrightarrow{F} a_0a_1$. This exhibits the local rule $F^{-1}(a_0a_1) = a$.

The converse claim is the preceding lemma. $\square$

In the preceding theorem, the partition $\mathscr{P}_F$ refines $\mathscr{P}$. This refinement can be proper (consider the identity map). The construction is practical. To construct, freely pick any partition of $\mathscr{A}$ to be $\mathscr{P}$, and freely pick any permutation of $\mathscr{A}$ to be π_F. Then for each class $c(a) = [a]$, freely pick as $\pi_a = \pi_{[a]}$ any permutation p satisfying (i) $p(b) = b$ if $b \in [a]$ and (ii) for each class $[a']$, $p : [a'] \mapsto \pi[a']$.

The following proposition relates the partitions $\mathscr{P}_F$ and $\mathscr{P}_{F^{-1}}$.

Proposition 8.3. *Let $F : \mathscr{A}^{\mathbf{N}} \to \mathscr{A}^{\mathbf{N}}$ be an invertible cellular automaton with $r(F) = r(F^{-1}) = 1$. Then for all a, b* in $\mathscr{A}$,

$$a \sim_F b \quad \Leftrightarrow \quad \pi_F(a) \sim_{F^{-1}} \pi_F(b).$$

In particular $c_{F^{-1}}(\pi_F(a)) = \pi_F(c_F(a))$, and the rule $c_F(a) \mapsto c_{F^{-1}}(\pi_F(a))$ induces a bijection $\mathscr{P}_F \to \mathscr{P}_{F^{-1}}$.

Proof. Let $a, b \in \mathscr{A}$ such that $a \sim_F b$. By definition, there are $b_0, \ldots, b_N \in \mathscr{A}$ and $c_0, \ldots, c_{N-1} \in \mathscr{A}$ verifying: $a = b_0$, $b = b_N$ and for $i \in \{0, \ldots, N-1\}$ $b_i \xrightarrow{F} c_i$, $b_{i+1} \xrightarrow{F} c_i$. On the other hand, since for $\alpha, \beta \in \mathscr{A}$, $\alpha \xrightarrow{F} \beta$ if and only if $\beta \xrightarrow{F^{-1}} \alpha$, we can deduce that $\pi_F(b_0) \xrightarrow{F^{-1}} b_0$, $c_0 \xrightarrow{F^{-1}} b_0$, $c_0 \xrightarrow{F^{-1}} b_1$, $c_1 \xrightarrow{F^{-1}} b_1, \ldots, c_{N-1} \xrightarrow{F^{-1}} b_{N-1}$, $c_{N-1} \xrightarrow{F^{-1}} b_N$, $\pi_F(b_N) \xrightarrow{F^{-1}} b_N$. This fact implies that $\pi_F(b_0) = \pi_F(a) \sim_{F^{-1}} \pi_F(b_N) = \pi_F(b)$. Using the same arguments with respect to F^{-1} we conclude that $a \sim_F b$ if and only if $\pi_F(a) \sim_{F^{-1}} \pi_F(b)$.

The concluding sentence of the proposition follows directly. $\square$

9. A class of expansive examples.

In this section, F will denote an invertible onesided cellular automaton, $F : \mathscr{A}^{\mathbf{N}} \to \mathscr{A}^{\mathbf{N}}$, such that $r(F) = r(F^{-1}) = 1$. We write a point of $\mathscr{A}^{\mathbf{N}}$ as $x = x_0x_1\cdots$. We let $\tilde{x}_i = (F^i x)_0$, $\tilde{x} = (\tilde{x}_i)_{i \in \mathbf{Z}}$, and $\tilde{X} = \{\tilde{x} : x \in \mathscr{A}^{\mathbf{N}}\}$. Here the map F is expansive if and only if the map $x \mapsto \tilde{x}$ is injective, and if the map is injective then it defines a topological conjugacy from $(\mathscr{A}^{\mathbf{N}}, F)$ to the subshift $(\tilde{X}, \sigma)$. In this case the onesided shift on $\mathscr{A}^{\mathbf{N}}$ can be presented by some block code of radius $\tilde{r} = \tilde{r}(F)$, so $\tilde{x}_{-\tilde{r}} \cdots \tilde{x}_{\tilde{r}}$ determines x_1.

We will give a constructive combinatorial characterization of those F for which $\tilde{r}(F) = 1$, and we will describe their dynamics. (For an interesting example in this class giving an idealized physical model, see [**C**]).

Theorem 9.1. *Let $F : \mathscr{A}^{\mathbf{N}} \to \mathscr{A}^{\mathbf{N}}$ be an invertible cellular automaton with $r(F) = r(F^{-1}) = 1$. The cellular automaton F is expansive with $\tilde{r}(F) = 1$ if and only if*

1. $\forall a, b \in \mathscr{A}$, $|c_F(a) \cap c_{F^{-1}}(b)| \leq 1$, *and*
2. $\forall a, b, b' \in \mathscr{A}$, $(F(ab) = F(ab') \Rightarrow b \sim_F b')$.

In this case,

1*. $\forall a, b \in \mathscr{A}$, $|c_F(a) \cap c_{F^{-1}}(b)| = 1$,
2*. $\forall a, b, b' \in \mathscr{A}$, $(F^{-1}(ab) = F^{-1}(ab') \Rightarrow b \sim_{F^{-1}} b')$, *and*
3*. $|\mathscr{A}| = J^2$ *for some J in $\mathbf{N}$ such that $|c_F(a)| = J$ for all a in $\mathscr{A}$.*

Proof. First let us see that property (2) implies property (2*). Suppose $F^{-1}(ab) = F^{-1}(ab')$. Then there exist symbols $\alpha, \beta, \gamma, \beta', \gamma'$ such that $F(\alpha\beta\gamma) = ab$ and $F(\alpha\beta'\gamma') = ab'$. Now property (2) implies $\beta \sim_F \beta'$. Also $\beta \xrightarrow{F} b$ and $\beta' \xrightarrow{F} b'$, and therefore $b \sim_{F^{-1}} b'$, proving (2*).

For the sufficient conditions let us assume F is an invertible cellular automaton with $r(F) = r(F^{-1}) = 1$ satisfying properties (1) and (2) (and therefore also (2*)). To prove that F is expansive it is enough to show $\tilde{x}_{-1}\tilde{x}_0\tilde{x}_1$ determines x_1. Suppose $x \in \mathscr{A}^{\mathbf{N}}$. From properties (2) and (2*) we deduce there are unique classes $c_F \in \mathscr{P}_F$ and $c_{F^{-1}} \in \mathscr{P}_{F^{-1}}$, depending on $\tilde{x}_{-1}\tilde{x}_0\tilde{x}_1$, such that $x_1 \in c_F \cap c_{F^{-1}}$. Therefore, by property (1), x_1 is the unique element of $c_F \cap c_{F^{-1}}$. This procedure provides the required block map $\tilde{x}_{-1}\tilde{x}_0\tilde{x}_1 \mapsto x_1$.

Now we turn to the necessary conditions. Suppose F is expansive with $\tilde{r}(F) = 1$. Take $a^* \in \mathscr{A}$ such that $|c_F(a^*)| = \max\{|c_F(a)| : a \in \mathscr{A}\}$. It follows from Proposition 8.3 that $|c_F(a^*)| = \max\{|c_{F^{-1}}(a)| : a \in \mathscr{A}\}$. On the other hand, from Theorem 8.2 we have $F(ba') = F(ba'')$ for any $b \in \mathscr{A}$ and $a', a'' \in c_F(a^*)$, and since $\tilde{r}(F) = 1$, $F^{-1}(ba') \neq F^{-1}(ba'')$. Therefore, $|c_{F^{-1}}(b)| \geq |\{b' \in \mathscr{A} : b \xrightarrow{F^{-1}} b'\}| \geq |c_F(a^*)| = \max\{|c_{F^{-1}}(a)| : a \in \mathscr{A}\}$. We conclude there is a positive integer J such that for any a in $\mathscr{A}$, $|c_F(a)| = |c_{F^{-1}}(a)| = J$.

To prove property (2) holds, suppose it does not. Then there are b, b', $a \in \mathscr{A}$ such that $F(ab) = F(ab')$ and $c_F(b) \neq c_F(b')$. In this case, by using the same arguments as in the last paragraph we deduce that $|c_{F^{-1}}(a)| \geq 2J$. This is a contradiction. This verifies property (2), and therefore also (2*).

Next note that for any $x = x_0x_1\ldots$, the triple $(x_0, c_F(x_1), c_{F^{-1}}(x_1))$ determines $\tilde{x}_{-1}\tilde{x}_0\tilde{x}_1$, which determines x_1. Therefore $c_F(x_1) \cap c_{F^{-1}}(x_1) = \{x_1\}$. This proves (1).

Now put $m = |\mathscr{P}_F| = |\mathscr{P}_{F^{-1}}|$. By property (1), every element of $\mathscr{P}_F$ intersects J distinct elements of $\mathscr{P}_{F^{-1}}$, so $m \geq J$. On the other hand, from property (2) we deduce that for each symbol a, there are at least J distinct symbols b such that $a \xrightarrow{F} b$; these F-successors of a must lie in the same element of $\mathscr{P}_{F^{-1}}$, so $m \leq J$. Therefore $m = J$. Because $\mathscr{A}$ is the disjoint union of the J members of $\mathscr{P}_F$, and each member contains exactly J symbols, we have $|\mathscr{A}| = J^2$, and (3*) holds. Finally, by (1) each member of $\mathscr{P}_F$ must intersect at least J members of $\mathscr{P}_{F^{-1}}$, so each member of $\mathscr{P}_F$ must intersect every member of $\mathscr{P}_{F^{-1}}$. This verifies (1*) and finishes the proof. □

Theorem 9.2. *Let $F : \mathscr{A}^{\mathbf{N}} \to \mathscr{A}^{\mathbf{N}}$ be an expansive invertible onesided cellular automaton such that $r(F) = r(F^{-1}) = \tilde{r}(F) = 1$. Then S_F is topologically conjugate to a full shift on $\sqrt{|\mathscr{A}|}$ symbols.*

Proof. By Theorem 9.1(3*), $\sqrt{|\mathscr{A}|}$ is a positive integer $J = |\mathscr{P}_F|$. Let Σ_J denote the full shift on J symbols, with alphabet $\mathscr{P}_F$. We already have $(\mathscr{A}^{\mathbf{N}}, F)$ topologically conjugate to $(\tilde{X}, \sigma)$, so it suffices to define a shift commuting bijection $c : \tilde{X} \to \Sigma_J$. We define $c(\tilde{x}) = c\tilde{x}$ by setting $(c\tilde{x})_i = c_F(\tilde{x}_i)$. By Theorem 9.1(2) and Lemma 8.1(2), we have for all a, b in $\mathscr{A}$ that

$$c_F(a) = c_F(b) \quad \Leftrightarrow \quad \pi_{a,F} = \pi_{b,F}. \tag{9.3}$$

If $a \xrightarrow{F} b$, then $\pi_{b,F^{-1}} = (\pi_{a,F})^{-1}$. It follows that if $a \xrightarrow{F} b$, then $c_F(a)$ determines $c_{F^{-1}}(b)$, and therefore the pair $(c_F(a), c_F(b))$ determines $c_F(b) \cap c_{F^{-1}}(b) = \{b\}$. It follows that for all $\tilde{x}$, the word $c_F(\tilde{x}_{-1})c_F(\tilde{x}_0)$ determines x_0. This proves that the map c is injective.

By induction we also observe that for all $n > 0$ and for all $\tilde{x}$,

$$\tilde{x}_0 \quad \text{and} \quad (c\tilde{x})_1 \cdots (c\tilde{x})_n \quad \text{determine } \tilde{x}_0 \cdots \tilde{x}_n. \tag{9.4}$$

To prove that c is surjective, it suffices to prove the claim: for all $n \geq 0$, there are J^{n+1} distinct words $(c\tilde{x})_0 \cdots (c\tilde{x})_n$. The claim is obvious for $n = 0$. Suppose $n > 0$ and the claim holds for $n - 1$. Using the induction hypothesis, let E be a set of J^n points x such that the restriction to E of the map $x \mapsto (c\tilde{x})_0 \cdots (c\tilde{x})_{n-1}$ is injective. Let $E' = \{ax : a \in \mathscr{A}, x \in E\}$, so $|E'| = J^{n+1}$. Using 9.3 above, one can see that the restriction to E' of the map $x \mapsto \tilde{x}_0 \cdots \tilde{x}_n$ is injective. This shows there are J^{n+1} distinct words $\tilde{x}_0 \cdots \tilde{x}_n$.

On the other hand, by 9.4 the number of words $\tilde{x}_0 \cdots \tilde{x}_n$ is at most $|\{\tilde{x}_0\}| \cdot |\{(c\tilde{x})_1 \cdots (c\tilde{x})_n\}| \leq (J)(J^n)$, and equality implies that for every $a \in \mathscr{A}$ and $x \in \mathscr{A}^{\mathbf{N}}$ there is some y such that $\tilde{y}_0 = a$ and $(c\tilde{y})_i = (c\tilde{x})_i$, $1 \leq i \leq n$. Thus there are J^{n+1} words $(c\tilde{x})_0 \cdots (c\tilde{x})_n$. $\square$

References

[AH] N. Aoki and K. Hiraide, Topological theory of dynamical systems, recent advances, North-Holand, Amsterdam, 1994.

[BM] F. Blanchard and A. Maass, Dynamical properties of expansive onesided cellular automata, Israel J. Math., **99** (1997), 149–174.

[B] M. Boyle, Constraints on the degree of a sofic homomorphism and the induced multiplication of measures on unstable sets, Israel J. Math., **53** (1986), no. 1, 52–68.

[BFF] M. Boyle, D. Fiebig and U. Fiebig, A dimension group for local homeomorphisms and endomorphisms of onesided shifts of finite type, J. reine angew. Math., **497** (1997), 27–59.

[BFK] M. Boyle, J. Franks and B. Kitchens, Automorphisms of onesided subshifts of finite type, Ergod. Th. & Dynam. Sys., **10** (1990), 421–449.

[BL] M. Boyle and D. Lind, Expansive subdynamics, Trans. AMS, **349** (1997), no. 1, 55–102.

[BMT] M. Boyle, B. Marcus and P. Trow, Resolving maps and the dimension group for shifts of finite type, Memoirs AMS, **377**, 1987.

[C] M. Courbage, On the approach to statistical equilibrium in an infinite particle lattice dynamic model, J. Math. Phys., **30** (1989), no. 8, 1840–1850.

[H] G. A. Hedlund, Endomorphisms and automorphisms of the shift dynamical system, Math. Systems Th., **3** (1969), 320–375.

[Hi] K. Hiraide, Dynamical systems of expansive maps on compact manifolds, Sugaku Expositions, **5** (1992), no. 2, 133–154.

[KR] K. H. Kim and F. W. Roush, Williams Conjecture is false for irreducible subshifts, Annals of Math., **149** (1999), 545–558.

[K1] B. Kitchens, Expansive dynamics on zero dimensional groups, Ergod. Th. & Dynam. Sys., **7** (1987), 249–261.

[K2] B. Kitchens, Symbolic Dynamics, One-sided, two-sided and countable state Markov shifts, Springer-Verlag, 1998.

[Kr1] W. Krieger, On dimension functions and topological Markov chains, Inventiones Math., **56** (1980), 239–250.

[Kr2] W. Krieger, On sofic systems I, Israel J. Math., **48** (1984), no. 4, 305–330.

[Ku] P. Kůrka, Languages, equicontinuity and attractors in cellular automota, Ergod. Th. & Dynam. Sys., **17** (1997), 417–433.

[LM] D. Lind and B. Marcus, Symbolic Dynamics and Coding, Cambridge University, 1995.

[Ma] G. A. Margulis, Certain measures associated with *U*-flows on compact manifolds, Functional Anal. Appl., **4** (1970), No. 1, 55–67.

[MM] G. Manzini and L. Margara, Invertible linear cellular automata over Z_m: algorithmic and dynamical aspects, J. Comput. System Sci., **56** (1998), no. 1, 60–67.

[N1] M. Nasu, Constant-to-one and onto global maps of homomorphisms between strongly connected graphs, Ergod. Th. & Dynam. Sys., **3** (1983), 387–413.

[N2] M. Nasu, Textile systems for endomorphisms and automorphisms of the shift, Memoirs AMS, **546** (1995).

[N3] M. Nasu, Maps in symbolic dynamics, in: Lecture Notes of the Tenth KAIST Mathematics Workshop 1995, ed. Geon Ho Choe, Korea Advanced Institute of Science and Technology Mathematics Research Center, Taejon, Korea (1996), 67 pages.

[SA] M. A. Shereshevsky and V. S. Afraimovich, Bipermutative cellular automata are topologically conjugate to the one-sided Bernoulli shift, Random & Computational Dynamics, **1** (1992), 91–98.

[S] Ya. G. Sinai, Markov partitions and C-diffeomorphisms, Functional Anal. Appl., **2** (1968), 64–89.

[T1] P. Trow, Degrees of constant-to-one factor maps, Proc. Amer. Math. Soc., **103** (1988), no. 1, 184–188.

[T2] P. Trow, Degrees of finite-to-one factor maps, Israel J. Math., **71** (1990), no. 2, 229–238.

[W] R. F. Williams, Classification of subshifts of finite type, Annals of Math., **98** (1973), 120–154; erratum, **99** (1974), 380–381.

Mike BOYLE
Department of Mathematics
University of Maryland
College Park, MD 20742-4015
U.S.A.
E-mail: mmb@math.umd.edu

Alejandro MAASS
Departamento de Ingeniería Matemática
Universidad de Chile
Casilla 170/3 correo 3
Santiago, Chile.
E-mail: amaass@dim.uchile.cl