Multiplicative Cellular Automata on Nilpotent Groups: Structure, Entropy, and Asymptotics

Marcus Pivato
Department of Mathematics, Trent University
Peterborough, Ontario, Canada
email: mpivato@trentu.ca
Phone: (705) 748 1011 x1293 Fax: (705) 748 1630
March 28, 2022

Abstract

If \( M = \mathbb{Z}^D \), and \( B \) is a finite (nonabelian) group, then \( B^M \) is a compact group; a multiplicative cellular automaton (MCA) is a continuous transformation \( \mathcal{G} : B^M \to B^M \) which commutes with all shift maps, and where nearby coordinates are combined using the multiplication operation of \( B \).

We characterize when MCA are group endomorphisms of \( B^M \), and show that MCA on \( B^M \) inherit a natural structure theory from the structure of \( B \). We apply this structure theory to compute the measurable entropy of MCA, and to study convergence of initial measures to Haar measure.

Keywords: Cellular Automata, Group, Structure, Entropy, Haar

1 Introduction

If \( B \) is a finite set, and \( M \) is some indexing set, then the configuration space \( B^M \) is the set of all \( M \)-indexed sequences of elements on \( B \). If \( B \) is discretely topologised, then the Tychonoff product topology on \( B^M \) is compact, totally disconnected, and metrizable. If \( M \) is an abelian monoid (e.g. \( M = \mathbb{Z}^D, \mathbb{N}^E, \) or \( \mathbb{Z}^D \times \mathbb{N}^E \)), then the action of \( M \) on itself by translation induces a natural shift action of \( M \) on configuration space: for all \( v \in M \), and \( b = \{b_m | m \in M\} \in B^M \), define \( \sigma^v b = \{b'_m | m \in M\} \) where, \( \forall m, \quad b'_m = b_{v+m} \).

A cellular automaton (CA) is a continuous map \( \mathcal{G} : B^M \to B^M \) which commutes with all shifts: for any \( m \in M \), \( \sigma^m \circ \mathcal{G} = \mathcal{G} \circ \sigma^m \). A result of Curtis, Heddlund, and Lyndon [4]
says any CA is determined by a local map \( g : \mathcal{B}^V \rightarrow \mathcal{B} \) (where \( V \subset M \) is some finite subset), so that, for all \( m \in M \), if we define \( m + V = \{ m + v \mid v \in V \} \), and for all \( b \in \mathcal{B}^M \), if we define \( b|_{m+V} \) to be the restriction of \( b \) to an element of \( \mathcal{B}^{m+V} \), then \( \mathcal{G}(b)_m = g \left( b|_{m+V} \right) \).

If \( (\mathcal{B}, \cdot) \) is a finite multiplicative group, let \( \text{End} [\mathcal{B}] \) be the set of group endomorphisms of \( \mathcal{B} \). A \textbf{multiplicative cellular automaton} (MCA) is a CA whose local map is a product of affine endomorphisms of separate coordinates. To be precise, let \( v : [1...I] \rightarrow \mathbb{V} \) be a (possibly noninjective) map, let \( g_1, g_2, \ldots, g_I \in \text{End} [\mathcal{B}] \), and let \( g_0, g_1, \ldots, g_I \in \mathcal{B} \) be constants. If \( b = [b_i]_{i=1}^I \in \mathbb{G}^V \), then the local map \( g \) has the form:

\[
g(b) = g_0 \cdot g_1 (b|_{[1]} ) \cdot g_1 \cdot g_2 (b|_{[2]} ) \cdot g_2 \cdot \ldots \cdot g_{I-1} \cdot g_I (b|_{[I]} ) \cdot g_I \quad (1)
\]

The \textbf{ordering function} \( v \) imposes an order on this product, which is necessary if \( \mathcal{B} \) is nonabelian. The endomorphisms \( \left[ g_i \right]_{i=1}^I \) are called the \textbf{coefficients} of \( \mathcal{G} \). We can rewrite equation (1) as

\[
g(b) = g \cdot g'_1 (b|_{[1]} ) \cdot g'_2 (b|_{[2]} ) \cdot \ldots \cdot g'_I (b|_{[I]} ) \quad = \quad g \cdot \prod_{i=1}^I g'_i (b|_{[i]} ) , \quad (2)
\]

where \( g = g_0 \cdot g_1 \cdot \ldots \cdot g_I \), and, for each \( i \in [1...I] \), \( g'_i (b) = \left( g_I g_{I-1} \ldots g_i \right)^{-1} \cdot g_i (b) \cdot \left( g_I g_{I-1} \ldots g_i \right) \) is an endomorphism. The product “\( \prod_{i=1}^I \)” inherits the obvious order from \( [1...I] \). We assume MCAs are written in the form (2), and call \( g \) the \textbf{bias}. If the bias is trivial (\( \mathcal{G} \) is “unbiased”), then \( g(b) \) is just a product of endomorphic images of the components \( \{ b_i \}_{v \in V} \).

\( \mathcal{B}^M \) is a compact group under componentwise multiplication; an \textbf{endomorphic cellular automaton} (ECA) is a topological group endomorphism \( \mathcal{G} : \mathcal{B}^M \rightarrow \mathcal{B}^M \) which commutes with all shift maps. If \( \mathcal{B} \) is abelian, then all unbiased MCA are ECA, and vice versa; when \( \mathcal{B} \) is nonabelian, however, the ECA form only a small subclass of MCA (see §2).

**Example 1:** Consider the following local maps:

(a) Let \( M = \mathbb{Z} \), \( V = \{0, 1\} \), and let \( g \left( b_0, b_1 \right) = b_0 \cdot b_1 \).

(b) \( M = \mathbb{N} \), \( V = [0...2] \); \( g \left( b_0, b_1, b_2 \right) = b_0 \cdot b_1 \cdot b_2 \).

(c) \( M = \mathbb{N} \), \( V = [0...2] \); \( g \left( b_0, b_1, b_2 \right) = b_0^2 \cdot b_1^2 \cdot b_0 \).

(d) \( M = \mathbb{Z}^2 \), \( V = [-1...1]^2 \); \( g(b) = \left( g \cdot b_{(-1,0)} \cdot g^{-1} \right) \cdot b_{(0,-1)} \cdot b_{(0,0)} \cdot b_{(0,1)} \cdot h \cdot b_{(1,0)} \), where \( g, h \in \mathcal{B} \) are constants.

(e) Suppose \( \mathcal{B} = \text{GL}_n(\mathbb{F}) \) is the group of invertible \( n \times n \) matrices over a finite field \( \mathbb{F} \) and let \( g \left( B_{-1}, B_0, B_1 \right) = \det[B_{-1}] \cdot \det[B_1]^2 \cdot B_0 \).

Example (1a) is the \textbf{nearest-neighbour multiplication} CA [1]. Examples (1a-1d) are unbiased, and all coefficients are the identity map on \( \mathcal{B} \).
In (I), \( v : [1..8] \rightarrow [0..2] \) is defined: \( v[1] = v[2] = v[3] = v[4] = 2, v[5] = v[5] = v[7] = 1, \) and \( v[8] = 0; \) by repeating indices in this way, we can obtain any exponents we want.

In (I), suppose \( card(B) = B; \) then \( v : [1..4+B] \rightarrow V \) is defined: \( v[1] = (-1, 0), v[2] = (0, -1), v[3] = (0, 0), v[4] = (0, 1), \) and \( v[n] = (1, 0) \) for \( n \in [5..4+B]; \) in this way, we obtain the exponent \( b_{(1,0)}^{B-1} = b_{(1,0)}^{-1}. \) All coefficients are the identity map, except for \( g_1(b) = g \cdot b \cdot g^{-1}, \) which is the endomorphism of conjugation-by-\( g. \)

In (I), let \( I \in GL^n(F) \) be the identity matrix. Then \( g_1(B) = det(B) \cdot I \) and \( g_2(B) = det(B)^2 \cdot I \) are endomorphisms of \( GL^n(F), \) and \( g_3 = Id. \) In fact, (I) is an ECA.

When \( B \) is an additive abelian group (e.g., \( B = (Z/p, +)), \) unbiased MCA are called linear CA (or affine CA, when biased). Classical modular arithmetic has been applied to study the entropy [9], and computational complexity [10, 11] of linear CA, while techniques of harmonic analysis yield convergence of initial probability measures on \( B^M \) to the uniformly distributed, or Haar measure under iteration by affine CA [2, 6, 8, 7, 12]. However, the case when \( B \) is nonabelian is poorly understood; “abelian” techniques usually fail to apply.

In §2, we give necessary and sufficient conditions for an MCA to be endomorphic. In §3, we use the structure theory of the group \( B \) to develop a corresponding structure theory for MCA over \( B. \) We apply this structure theory in §4 to compute the measurable entropy of MCA, and in §5 to establish sufficient conditions for convergence of initial measures to Haar measure under iteration of MCA. The major results are Theorems [4, 14, and 21].

\[ 2 \text{ Endomorphic Cellular Automata} \]

Suppose \( \mathcal{G} : B^M \rightarrow B^M \) is an ECA. Since \( \mathcal{G} \in \text{End}[B^M], \) the local map \( g \) must be a group homomorphism from the product group \( B^V \) into \( B. \) This constrains the coefficients \( \{g_v\}_{v \in V} \) and their interactions.

**Lemma 2** Let \( g : B^V \rightarrow B \) be a group homomorphism. Then there are endomorphisms \( g_v \in \text{End}[B] \) for all \( v \in V \) so that, for any \( b = [b_v]_{v \in V} \in B^V, \) \( g(b) = \prod_{v \in V} g_v(b_v), \) where this product is commutative.

**Proof:** For each \( v \in V, \) let \( i_v : B \rightarrow B^V \) be the embedding into the \( v \)th coordinate: for any \( b \in B, \) \( (i_v(b))_v = b, \) and \( (i_v(b))_w = e \) for all \( w \neq v \) in \( V, \) where \( e \in B \) is the identity element. Then define \( g_v = g \circ i_v. \) If \( b = [b_v]_{v \in V}, \) then clearly, \( b = \prod_{v \in V} i_v(b_v), \) where the factors all commute, and thus, \( g(b) = g \left( \prod_{v \in V} i_v(b_v) \right) = \prod_{v \in V} g(i_v(b_v)) = \prod_{v \in V} g_v(b_v), \) where, again, the factors all commute. \( \square \)
We say two endomorphisms $g_w$ and $g_v$ have **commuting images** if, for any $b_w, b_v \in \mathcal{B}$, $g_v(b_v) \cdot g_w(b_w) = g_w(b_w) \cdot g_v(b_v)$. Thus, the coefficients of any ECA $\mathcal{G}$ must all have commuting images; this restricts the structure of $\mathcal{G}$, and the more noncommutative $\mathcal{B}$ itself is, the more severe the restriction becomes. The noncommutativity of $\mathcal{B}$ is measured by two subgroups: the **centre**, $Z(\mathcal{B}) = \{z \in \mathcal{B} : \forall b \in \mathcal{B}, b \cdot z = z \cdot b\}$, and the **commutator subgroup**, $[B, B] = \langle b \cdot h \cdot b^{-1} \cdot h^{-1}; \ b, h \in B \rangle$. If $\phi : B \rightarrow A$ is any homomorphism from $\mathcal{B}$ into an abelian group $\mathcal{A}$, then $[B, B] \subset \ker[\phi]$.

**Corollary 3** Continuing with the previous notation,

1. If $\exists v \in \mathcal{V}$ so that $g_v$ is surjective, then, for all other $w \in \mathcal{V}$, $\text{image}[g_w] \subset Z(\mathcal{B})$. If $Z(\mathcal{B}) = \{e\}$, then all other coefficients of $g$ are trivial.

2. Suppose $v \neq w \in \mathcal{V}$ are such that $g_v = g_w$. Then $\text{image}[g_v]$ is an abelian subgroup of $\mathcal{B}$, and thus, $[B, B] \subset \ker[g_v]$. Thus, if $[B, B] = B$, then $g_v$ and $g_w$ are trivial.

3. If $\mathcal{B}$ is simple but nonabelian, then only one coefficient of $\mathcal{G}$ can be nontrivial; this coefficient is an automorphism.

**Proof:** Part 1 and Part 2 are straightforward. To see Part 3, note that $Z(\mathcal{B})$ is a normal subgroup, so if $\mathcal{B}$ is simple nonabelian, then $Z(\mathcal{B}) = \{e\}$. On the other hand, any endomorphism of $\mathcal{B}$ is either trivial or an automorphism. Hence, if $\mathcal{G}$ is nontrivial, it must have one automorphic coefficient, and then, by Part 1 all other coefficients must be trivial. 

---

### 3 Structure Theory

We now relate the structure of the group $\mathcal{B}$ to the structure of MCA on $\mathcal{B}^\mathcal{M}$. We review the structure theory of dynamical systems in §3.1 and group structure theory in §3.2. In §3.3, we show that, if $\mathcal{A}$ is a fully characteristic subgroup of $\mathcal{B}$, and $\mathcal{C} = \mathcal{B}/\mathcal{A}$, then the decomposition of $\mathcal{B}$ into $\mathcal{A}$ and $\mathcal{C}$ yields a corresponding decomposition of MCA on $\mathcal{B}^\mathcal{M}$.

**Notation:** We will often decompose objects (eg. groups, spaces, measures, functions) into factor and cofactor components. We will use three lexicographically consecutive letters to indicate, respectively, the cofactor, product, and factor (eg. for groups: $\mathcal{A} \hookrightarrow \mathcal{B} \twoheadrightarrow \mathcal{C}$; for measure spaces: $(Y, \mathcal{Y}, \mu) = (X \times Z, \mathcal{X} \otimes \mathcal{Z}, \lambda \otimes \nu)$; for dynamical systems, $\mathcal{G} = F \star H$; for cellular automata, $\mathcal{G} = \mathfrak{F} \star \mathfrak{H}$, and for their local maps, $g = f \star h$, etc.).
3.1 Relative and Nonhomogeneous CA

Let $X$ and $Z$ be a topological spaces. A topological $Z$-relative dynamical system \[ F : X \times Z \to X \] on $X$ is a continuous map $F : X \times Z \to X$. We write the second argument of $F$ as a subscript: for $(x, z) \in X \times Z$, \[ F(x, z) \] is written as “$F_z(x)$”. Thus, $F$ is treated as a $Z$-parameterized family of fibre maps \[ \{ F_z : X \rightarrow X \}_{z \in Z}. \] Let $\mathcal{M}[X]$ be the set of Borel probability measures on $X$; if $\lambda \in \mathcal{M}[X]$, then $F$ is $\lambda$-preserving if $F_z(\lambda) = \lambda$ for all $z \in Z$.

If $H : Z \to Z$ is a topological dynamical system, then the skew product of $F$ and $H$ is the topological dynamical system $G = F \star H$ on $Y = X \times Z$ defined: $G(x, z) = (F_z(x), H(z))$.

Now, suppose $X = A^M$ and $Z = C^M$, where $A$ and $C$ are finite sets. If $B = A \times C$, then there is a natural bijection $A^M \times C^M \cong B^M$. A $C$-relative cellular automaton (RCA) on $A^M$ is a continuous map $\mathfrak{F} : A^M \times C^M \to A^M$ which commutes with all shift maps: $\sigma^m \circ \mathfrak{F} = \mathfrak{F} \circ \sigma^m$, for all $m \in \mathbb{M}$. Like an ordinary CA, $\mathfrak{F}$ is determined by a local map $f : A^V \times C^V \to A$, where $\forall \in \mathbb{M}$ is finite, so that, for all $(a, c) \in B^M$, and $m \in \mathbb{M}$, $\mathfrak{F}(a, c)_m = f_{a,m+V}(c_{m+V})$. For any $c \in C^V$, the local fibre map $f_c : A^V \to A$ is defined by $f_c(a) = f(c, a)$. If $A$ is a group and $f_c$ is a product of affine endomorphisms for every $c \in C^V$, then $\mathfrak{F}$ is called a multiplicative relative cellular automaton (MRCA).

If $\mathfrak{S} : C^M \to C^M$ is a CA with local map $\mathfrak{S} : C^V \to C$, then the skew product $\mathfrak{S} \star \mathfrak{F}$ is a CA on $B^M$, with local map $g : B^V \cong A^V \times C^V \to B$ defined: $g(a, c) = (f_c(a), h(c))$.

A nonhomogeneous cellular automaton (NHCA) is a continuous map $\mathfrak{S} : B^M \to B^M$ which does not necessarily commute with shift maps, but where there is some finite $V \subset \mathbb{M}$, so that, for all $m \in \mathbb{M}$, there is a local map $g_m : B^{(m+V)} \to B$ so that, $\forall b \in B^M$, $\mathfrak{S}(b)_m = g_m(b_{(m+V)})$. Thus, for example, any CA is an NHCA. If $\mathfrak{S} : A^M \times C^M \to A^M$ is an RCA, then, for any $c \in C^M$, the fibre map $\mathfrak{S}_c : A^M \to A^M$ is an NHCA.

3.2 Group Structure Theory

Let $B$ be a group. A subgroup $A \subset B$ is called fully characteristic \[ \text{fully characteristic} \] if, for every $\phi \in \text{End}[B]$, we have $\phi(A) \subset A$. We indicate this: “$A \prec B$”. For example, if $Z(B)$ is the center of $B$, then $Z(B) \prec B$. Observe that any fully characteristic subgroup is normal.

The main result of this section is:

**Theorem 4** Suppose that $A \prec B$, and $B/A = C$. If $\mathfrak{S} : B^M \to B^M$ is an MCA, then there is an MCA $\mathfrak{S} : C^M \to C^M$ and an MRCA $\mathfrak{F} : A^M \times C^M \to A^M$ so that $\mathfrak{S} = \mathfrak{F} \star \mathfrak{S}$.

We will prove this result in §3.3, and also describes the structure of the local maps of $\mathfrak{F}$ and $\mathfrak{S}$ (see Proposition 3). First we introduce the relevant algebraic machinery.

Semidirect Products and and Pseudoproducts: Suppose $A \subset B$ is a normal subgroup, and $C = B/A$, and let $\pi : B \to C$ be the quotient map. Let $\zeta : C \to B$ be a section of
Thus, multiplication using pseudoproduct notation satisfies the equation:

\[ A \times C \ni (a \star c) = a \cdot \varsigma(c). \]

In general, \( \varsigma(b) \) write: “\( B = A \times C \).” We call \( B \) a pseudoproduct of \( A \) and \( C \), and write: “\( \mathcal{B} = \mathcal{A} \rtimes \mathcal{C} \).”

If \( c \in C \), the conjugation automorphism \( c^* \in \text{Aut}[\mathcal{A}] \) is defined:

\[ c^* a = \varsigma(c) \cdot a \cdot \varsigma(c)^{-1}. \]

Thus, multiplication using pseudoproduct notation satisfies the equation:

\[ (a_1 \star c_1) \cdot (a_2 \star c_2) = (a_1 \cdot \varsigma(c_1)) \cdot (a_2 \cdot \varsigma(c_2)) = a_1 \cdot (c_1^* a_2) \cdot (\varsigma(c_1) \cdot \varsigma(c_2)). \] (3)

In general, \( \varsigma(c_1) \cdot c_2 \) does not equal \( \varsigma(c_1 \cdot c_2) \); this is true only \( B \) is a semidirect product of \( A \) and \( C \). In this case, \( \varsigma \) is an isomorphism from \( C \) into an embedded subgroup \( \varsigma(C) \subset B \), and (3) becomes:

\[ (a_1 \star c_1) \cdot (a_2 \star c_2) = (a_1 \cdot (c_1^* a_2)) \star (c_1 \cdot c_2). \] (4)

In this case, we write: “\( \mathcal{B} = \mathcal{A} \rtimes \mathcal{C} \).” We can treat \( C \) as embedded in \( B \), so \( \varsigma \) is just the identity, and \( a \star c = a \cdot c \).

We call \( B \) a polymorph of \( A \) if: (1) \( B = A \rtimes C \); (2) \( A \) and \( \varsigma(C) \) are both fully characteristic in \( B \); and (3) \( c^* \in \text{Aut}(\mathcal{A}) \), for every \( c \in C \).

Example 5:

(a) Let \( B = \mathbb{Q}_8 = \{ \pm 1, \pm i, \pm j, \pm k \} \) be the Quaternion Group, defined by: \( i^2 = j^2 = k^2 = -1 \), and \( q_1 \cdot q_2 = q_3 = -q_2 \cdot q_1 \) for \( (q_1, q_2, q_3) = (i, j, k) \) or any cyclic permutation thereof. Let \( A = \mathbb{Z}(\mathbb{Q}_8) = \{ \pm 1 \} \); then \( C = \mathbb{Q}_8/A \cong \mathbb{Z}/2 \oplus \mathbb{Z}/2 \).

Define homomorphism \( \pi : \mathbb{Q}_8 \rightarrow \mathcal{C} \) by \( \pi(\pm 1) = \mathcal{O} := (0, 0), \pi(\pm i) = \mathcal{I} := (1, 0), \pi(\pm j) = \mathcal{J} := (0, 1), \pi(\pm k) = \mathcal{K} := (1, 1) \), with \( \text{ker} \[ \pi \] = \{ \pm 1 \} = A \). If \( \varsigma : \mathcal{C} \rightarrow \mathcal{B} \) is defined: \( \varsigma(\mathcal{O}) = 1, \varsigma(\mathcal{I}) = i, \varsigma(\mathcal{J}) = j, \varsigma(\mathcal{K}) = k \), then \( \varsigma \) induces a (non-semidirect) pseudoproduct structure \( \mathbb{Q}_8 = \mathcal{A} \rtimes \mathcal{C} \). In this case, \( A = \mathbb{Z}(\mathcal{B}) \), and multiplication satisfies the formula:

\[ (a_1 \star c_1) \cdot (a_2 \star c_2) = (a_1 \cdot a_2 \cdot \varsigma(c_1, c_2)) \star (c_1 \cdot c_2) \quad \text{for all } a_1, a_2 \in \mathcal{A}, c_1, c_2 \in \mathcal{C}. \]

Here, \( \varsigma(c_1, c_2) = \varsigma(c_1 \cdot c_2)^{-1} \cdot \varsigma(c_1) \cdot \varsigma(c_2) = \text{sign}[\varsigma(c_1) \cdot \varsigma(c_2)]. \) For example, \( \varsigma(\mathcal{I}, \mathcal{J}) = \text{sign}[k] = +1 \), while \( \varsigma(\mathcal{J}, \mathcal{I}) = \text{sign}[\mathcal{K}] = -1 \), and \( \varsigma(\mathcal{O}, c) = 1 \) for any \( c \in C \).

---

1...but generally not a homomorphism.
If $p$ is prime and $A = (\mathbb{Z}/p, +)$ is the (additive) cyclic group of order $p$, then $\text{Aut} [A]$ is the (multiplicative) group $(\mathbb{Z}/p^\times, \cdot)$ of nonzero elements of the field $\mathbb{Z}/p$, acting on $\mathbb{Z}/p$ by multiplication, mod $p$. The group $(\mathbb{Z}/p^\times, \cdot)$ is isomorphic to $(\mathbb{Z}/p-1, +)$; thus, $B = \mathbb{Z}/p \times \mathbb{Z}/p^\times \cong \mathbb{Z}/p \times \mathbb{Z}/p-1$ is a group of order $p \cdot (p-1)$, and $\mathbb{Z}/p$ is a characteristic subgroup. Since $\mathbb{Z}/p^\times$ is itself abelian, $B$ is a polymorph of $\mathbb{Z}/p$.

If $q$ divides $p-1$, then there is a cyclic multiplicative subgroup $C_q \subset \mathbb{Z}/p^\times$ of order $q$. The semidirect product $D_{p,q} = \mathbb{Z}/p \rtimes C_q$, is also a polymorph of $\mathbb{Z}/p$. For example, if $p = 7$, then $C_3 = \{1, 2, 4\}$, and $D_{7,3} = \mathbb{Z}/7 \rtimes C_3$ has cardinality 21.

### 3.3 The Induced Decomposition

If $B = A \rtimes C$, then for any $a \in A^M$ and $c \in C^M$, define $b = a \rtimes c \in B^M$ by $b_m = a_m \rtimes c_m$ for all $m \in M$; we will thus identify $A^M \times C^M$ with $B^M$.

Suppose $B = A \rtimes C$, and let $\mathcal{G} : B^2 \rightarrow B^2$ be the nearest neighbour multiplication CA, Example [13]. If $b = a \rtimes c$, then Moore [11] noted that
\[
\mathcal{G}(b_0, b_1) = (a_0 \rtimes c_0) \cdot (a_1 \rtimes c_1) = (a_0 \cdot c_0^a \cdot a_1) \cdot (c_0 \cdot c_1) = f_c(a_0, a_1) \rtimes h(c_0, c_1)
\]
where $h : C^{(0,1)} \rightarrow C$ is defined $h(c_0, c_1) = c_0 \cdot c_1$, and $f : A^{(0,1)} \times C^{(0,1)} \rightarrow A$ is defined $f_c(a_0, a_1) = a_0 \cdot c_0^a \cdot a_1$. Thus, $\mathcal{G} = \mathfrak{G} \rtimes \mathfrak{H}$, where $\mathfrak{G} : C^M \rightarrow C^M$ is the CA with local map $h$, and $\mathfrak{H} : A^M \times C^M \rightarrow A^M$ is the RCA with local map $f$. In other words, the decomposition of $B = A \rtimes C$ induces a decomposition of $\mathcal{G}$. We now generalize this idea to arbitrary MCA.

**Lemma 6** Suppose $A \lt B$, $C = B/A$, and $B = A \rtimes C$. Let $g \in \text{End } [B]$.

1. There exist $f \in \text{End } [A]$ and $h \in \text{End } [C]$ so that the following diagram commutes:
\[
\begin{array}{ccc}
A & \rightarrow & B \\
\downarrow{f} & & \downarrow{g} \\
A & \rightarrow & B \\
\end{array}
\]

We indicate this: “$g = f \rtimes h$.”

2. Define $g' : C \rightarrow A$ by $g'(c) = g(\zeta(c)) \cdot \zeta(h(c))^{-1}$. If $a \in A$ and $c \in C$, then $g(a \rtimes c) = (f(a) \cdot g'(c)) \rtimes h(c)$.

3. If $B$ is a polymorph of $A$, then $g'$ is trivial, so $g(b) = f(a) \rtimes h(c)$.

4. If $B$ is a polymorph of $A$ and $f \in \text{Aut } [A]$, then $g|_{\zeta(c)} = \text{Id}$ and $h = \text{Id}_c$.

**Proof:** Part 1: Define $f = g|_A$. Then $f \in \text{End } [A]$ because $A \rtimes B$. Define $h$ by:

$h(b \cdot A) = g(b) \cdot A$ for any coset $(b \cdot A) \in C$. Then, for any $(b_1 \cdot A)$ and $(b_2 \cdot A)$ in $C$, we have...
\( h(b_1A \cdot b_2A) = h((b_1 \cdot b_2) \cdot A) = g(b_1 \cdot b_2) \cdot A = g(b_1) \cdot g(b_2) \cdot A = g(b_1)A \cdot g(b_2)A = h(b_1A) \cdot h(b_2A), \) so \( h \) is an endomorphism of \( C. \) Clearly, \( h \circ \pi = \pi \circ g. \)

**Part 4:** If \( a_1, a_2 \in A \) and \( d_1, d_2 \in D = \zeta(C), \) then by (7.1),

\[
\begin{align*}
\hat f(a_1) \cdot \left( (g(d_1^\ast) \circ \hat f)(a_2) \cdot (g(d_1) \cdot g(d_2)) \right) &= \left( g(a_1) \cdot (g(d_1)) \cdot (g(a_2) \cdot g(d_2)) \right) = g \left( (a_1 \cdot (d_1) \cdot (a_2 \cdot d_2)) \right) \\
&= g \left( (a_1 \cdot d_1^\ast(a_2)) \cdot (d_1 \cdot d_2) \right) = \hat f(a_1) \cdot (f \circ d_1^\ast) \cdot (a_2) \cdot g(d_1) \cdot g(d_2).
\end{align*}
\]

Cancel \( \hat f(a_1) \) and \( g(d_1) \cdot g(d_2), \) and note that \( a_2 \) is arbitrary to conclude: \( g(d_1^\ast) \circ \hat f = \hat f \circ d_1^\ast. \) Since \( d_1^\ast \in Z(A, \text{Aut}[A]) \), commute these terms to get \( g(d_1^\ast) \circ \hat f = \hat f \circ d_1^\ast; \) cancel \( \hat f \in \text{Aut}[A] \) to conclude that \( g(d_1^\ast) = d_1^\ast. \) Now, \( D \) is fully characteristic, so \( g(d_1) \in D. \) But the map \( D \ni d \rightarrow d^\ast \in \text{Aut}[A] \) is really just the inclusion map \( D \cong C \hookrightarrow \text{Aut}[A], \) and therefore injective, so we conclude \( g(d_1) = d_1. \)

Thus, \( g|_D = \text{Id}_D, \) so \( g \circ \zeta = \zeta. \) Since \( \pi \circ \zeta = \text{Id}_C, \) we conclude: \( h = h \circ \text{Id}_C = h \circ \pi \circ \zeta = \pi \circ g \circ \zeta = \pi \circ \zeta = \text{Id}_C. \)

**Part 2 and Part 3** are straightforward. \hfill \Box

**Example 7:** Recall \( Q_8 = A \ast C \) from Example (Pa). Define \( g_1, g_2 \in \text{Aut}[Q_8], \) by Table 7.1. Then \( f_1 = f_2 = \text{Id}_A, \) while \( h_1, h_2, g_1', \) and \( g_2' \) are defined by Table 7.2.

|   | O | I | J | K |
|---|---|---|---|---|
| \( g_1 \) | 1 | -1 | j | -j |
| \( g_2 \) | 1 | -1 | i | k   |
| \( g_1' \) | 1 | 1  | 1  | 1  |
| \( g_2' \) | 1 | -1 | 1  | 1  |

**Table 7.1**

**Table 7.2**

**Proposition 8**  The statement of Theorem 4 is true. To be specific, if \( G \) has local map

\[
g : B^\forall \ni b \mapsto \left( g \cdot \prod_{i=1}^{I} g_i \left( b_{v_i} \right) \right) \in B, \quad \text{(where } b \in B \text{ and } g_i \in \text{End}[B], \text{ for all } i \in [0..I])
\]

then the local maps \( h : C^\forall \rightarrow C \) and \( \hat f : A^\forall \times C^\forall \rightarrow A \) are defined as follows. Fix a pseudoproduct representation \( B = A \ast C. \) Let \( g = f \ast h \) for some \( f \in A \) and \( h \in C. \) For all \( i \in [0..I], \) let \( g_i = f \ast h_i, \) where \( f_i \in \text{End}[A] \) and \( h_i \in \text{End}[C], \) as in Lemma 2. Then:

8
1. \( h(c) = h \cdot \prod_{i=1}^{l} h_i(c_{v[i]}), \) and \( f \) is defined by expression [8] below.

In particular:

2. Suppose \( B \) is a polymorph of \( A \), and \( f_i \in \text{Aut}[B], \forall i \in [0..l] \). Then \( h(c) = \prod_{i=0}^{l} c_{v[i]} \) and

\[
f_i(a) = f \cdot \prod_{i=0}^{l} f_i^c(a_{v[i]}), \quad \text{where, } \forall i \geq 0, \forall a \in A, \quad f_i^c(a) = h \cdot c_{v[0]}^c c_{v[1]}^c \cdots c_{v[i-1]}^c f_i(a).
\]

3. Suppose \( A \subset Z(B) \). Treat \( A \) as an additive group \((A,+)\). Then \( f_c(a) = \mathcal{L}(a) + \mathcal{P}(c) \),

where \( \mathcal{L} : A^M \rightarrow A^M \) is a linear cellular automaton with local map

\[
\mathcal{L} : A^V \ni a \mapsto \left( \sum_{i=0}^{l} f_i(a_{v[i]}) \right) \in A \quad (5)
\]

and \( \mathcal{P} : C^M \rightarrow A^M \) is a block map with local map \( p : C^V \rightarrow A \) given by [9] below.

Example 9:

(a) Suppose \( A \subset Z(B) \), as in Part 3 of Proposition 8. If \( g(b) = b_{v_1}^{n_1} b_{v_2}^{n_2} \cdots b_{v_j}^{n_j} \), then

\[
l(a) = \sum_{v \in V} \ell_v \cdot a_v, \quad \text{where } \ell_v = \sum_{v_j = v} n_j \text{ for each } v \in V.
\]

Meanwhile, \( h(c) = c_{v_1}^{n_1} c_{v_2}^{n_2} \cdots c_{v_j}^{n_j} \).

(b) Consider Example [10], with \( B = \mathbb{Z}/5 \times \mathbb{Z}/4 \). In this case, \( h_i = \text{Id}_c \) and \( f_i = \text{Id}_a \) for all \( i \).

Thus, \( f : (\mathbb{Z}/4)^{[0..2]} \times (\mathbb{Z}/4)^{[0..2]} \rightarrow \mathbb{Z}/5 \) is the linear CA with local map: \( h(c_0, c_1, c_2) = c_0 + c_1 + c_2 \).

For any \( a \in \mathbb{Z}/5 \) and \( c \in \mathbb{Z}/4 \), \( c^a a = 2^c \cdot a \); thus, Part 2 of Proposition 8 implies that \( f : (\mathbb{Z}/5)^{[0..2]} \times (\mathbb{Z}/4)^{[0..2]} \rightarrow \mathbb{Z}/5 \) is defined for all \((a_0, a_1, a_2) \in (\mathbb{Z}/5)^{[0..2]} \) and \((c_0, c_1, c_2) \in (\mathbb{Z}/4)^{[0..2]} \), by:

\[
f_{(c_0, c_1, c_2)}(a_0, a_1, a_2) = a_0 + 2^{c_0} a_1 + 2^{c_0 + c_1} a_2.
\]

(c) Consider Example [11], with \( B = \mathbb{Z}/5 \times \mathbb{Z}/4 \). Now \( f : (\mathbb{Z}/4)^{Z} \rightarrow (\mathbb{Z}/4)^{Z} \) has local map

\[
h(c_0, c_1, c_2) = c_0 + 3c_1 + 4c_2 \equiv c_0 - c_1 \pmod{4}.
\]

Meanwhile, Part 2 of Proposition 8 implies that

\[
f_{(c_0, c_1, c_2)}(a_0, a_1, a_2)
\]

\[
= a_2 + 2^{c_2} a_2 + 2^{4c_2} a_2 + 2^{3c_2} a_2 + 2^{4c_2} a_1 + 2^{4c_2 + c_1} a_1 + 2^{4c_2 + 2c_1} a_1 + 2^{4c_2 + 3c_1} a_0
\]

\[
= (1 + 2^{c_2} + 2^{2c_2} + 2^{3c_2}) a_2 + (2^{4c_2} + 2^{4c_2 + c_1} + 2^{4c_2 + 2c_1}) a_1 + 2^{4c_2 + 3c_1} a_0.
\]
(d) Recall \( g_1, g_2 \in \text{Aut} [Q_8] \) from Example 1, and define \( g : Q_8^{[0..2]} \rightarrow Q_8 \) by \( g(q_0, q_1, q_2) = q_0 \cdot g_1(q_1) \cdot g_2(q_2) \). Identify \( \{ \pm 1 \} = A \) with \( (\mathbb{Z}/2, +) \), and identify \( C \) with \( \mathbb{Z}/2 \oplus \mathbb{Z}/2 \) as described in Example 5. Then by Part 3 of Proposition 8,

\[
I(a_0, a_1, a_2) = a_0 + f_1(a_1) + f_2(a_2) = a_0 + a_1 + a_2;
\]

\[
h(c_0, c_1, c_2) = c_0 + h_1(c_1) + h_2(c_2) = \begin{bmatrix} c_{0,1} \\ c_{0,2} \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_{1,1} \\ c_{1,2} \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} c_{2,1} \\ c_{2,2} \end{bmatrix}.
\]

(where \( c_i = \begin{bmatrix} c_{i,1} \\ c_{i,2} \end{bmatrix} \in \mathbb{Z}/2 \oplus \mathbb{Z}/2 \), for \( i = 0, 1, 2 \). Also, applying (9) below,

\[
p(c_0, c_1, c_2) = e(c_0, c_1, c_2) + \text{I}d'(c_0) + g_1'(c_1) + g_2'(c_2) = e(c_0, c_1, c_2) + g_1'(c_1) + g_2'(c_2),
\]

where \( e(c_0, c_1, c_2) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \) if \( c_0 \cdot c_1 \cdot c_2 \), \( e(c_0, c_1, c_2) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \) if \( c_0 \cdot c_1 \cdot c_2 \).

(e) \( g : Q_8^{[0..3]} \rightarrow Q_8 \) by \( g(q_0, q_1, q_2, q_3) = q_3 \cdot q_0^3 \cdot q_2^5 \cdot q_1^{-1} \). Then, in additive notation, \( I(a_0, a_1, a_2, a_3) = a_0 + a_1 + a_2 + a_3 \) and \( h(c_0, c_1, c_2, c_3) = c_0 + c_1 + c_2 + c_3 \pmod{2} \).

Proof of Proposition 8: For \( i \in [0..I] \), define \( f_i : C^V \times A \rightarrow A \) by

\[
f_i^a(c) = \left( h \cdot \prod_{j=0}^{i-1} h_j(c_{v[j]}) \right)^* \left( f_i(a) \cdot g_i'(c_{v[i]}) \right); \tag{6}
\]

for example, \( f_i^0(a) = h^* \left( f_0(a) \cdot g_0'(c_{v[0]}) \right) \). Next, define \( e : C^V \rightarrow A \) by:

\[
e(c) = \left( \varsigma(h) \cdot \prod_{i=0}^{l} \varsigma(h_i(c_{v[i]})) \right) \cdot \varsigma \left( h \cdot \prod_{i=0}^{l} h_i(c_{v[i]}) \right)^{-1}, \tag{7}
\]

and define \( f : A^V \times C^V \rightarrow A \) by:

\[
f_c(a) = f \cdot \left( \prod_{i=0}^{l} f_i^a(c_{v[i]}) \right) \cdot e(c); \tag{8}
\]

To prove Part 1, we must show, for any \( a \in A^V \) and \( c \in C^V \), that \( g(a \ast c) = f_c(a) \ast h(c) \).

To see this, let \( c = [c_{v} |_{v \in V}] \in C^V \) and \( a = [a_{v} |_{v \in V}] \in A^V \). Then

\[
g(a \ast c) = g \cdot \prod_{i=0}^{l} g_i(a_{v[i]} \cdot \varsigma(c_{v[i]})) = f \cdot \varsigma(h) \cdot \prod_{i=0}^{l} g_i(a_{v[i]} \cdot g_i(\varsigma(c_{v[i]}))).
\]
In the case $I = 1$, this becomes:

$$g(a \ast c) = f \cdot \varsigma (h) \cdot g_0(a_{v[0]}) \cdot g_0(\varsigma (c_{v[0]})) \cdot g_1(a_{v[1]}) \cdot g_1(\varsigma (c_{v[1]}))$$

$$= f \cdot \varsigma (h) \cdot f_0(a_{v[0]}) \cdot g'_0(c_{v[0]}) \cdot \varsigma (b_0(c_{v[0]})) \cdot f_1(a_{v[1]}) \cdot g'_1(c_{v[1]}) \cdot \varsigma (b_1(c_{v[1]}))$$

$$= f \cdot \varsigma (h) \cdot f_0(a_{v[0]}) \cdot g'_0(c_{v[0]}) \cdot b_0(c_{v[0]})^* \left( f_1(a_{v[1]}) \cdot g'_1(c_{v[1]}) \right) \cdot \varsigma (b_0(c_{v[0]})) \cdot \varsigma (b_1(c_{v[1]}))$$

$$= f \cdot h^* \left( f_0(a_{v[0]}) \cdot g'_0(c_{v[0]}) \right) \cdot (h \cdot b_0(c_{v[0]}))^* \left( f_1(a_{v[1]}) \cdot g'_1(c_{v[1]}) \right) \cdot \varsigma (h) \cdot \varsigma (b_0(c_{v[0]})) \cdot \varsigma (b_1(c_{v[1]}))$$

$$= f_c(a) \cdot \varsigma (b(c)) = f_c(a) \ast b(c).$$

A similar argument clearly works for $I \geq 2$.

It remains to show that $e(c) \in A$, which is equivalent to showing that $\pi (e(c)) = e_c$, where $\pi : B \to C$ is the quotient map and $e_c \in C$ is the identity. But

$$\pi (e(c)) = \pi \left[ \left( \varsigma (h) \cdot \prod_{i=0}^l \varsigma (b_i(c_{v[i]})) \right) \cdot \varsigma \left( h \cdot \prod_{i=0}^l b_i(c_{v[i]}) \right)^{-1} \right]$$

$$= \pi \left( \varsigma (h) \cdot \prod_{i=0}^l \pi \left( \varsigma (b_i(c_{v[i]})) \right) \cdot \pi \left( \varsigma \left( h \cdot \prod_{i=0}^l b_i(c_{v[i]}) \right) \right)^{-1} \right)$$

$$= h \cdot \prod_{i=0}^l b_i(c_{v[i]}) \cdot \left( h \cdot \prod_{i=0}^l b_i(c_{v[i]}) \right)^{-1} = e_c.$$

**Part 2**: $g'_i$ and $e$ are trivial, and **Part 4** of Lemma 3 implies that $b_i = \text{Id}_{c_{v[i]}}, \forall i \in [0...l]$. **Part 3**: All the conjugation automorphisms are trivial, so expression (3) (written additively) simplifies to $f'_c(a) = f_i(a) + g'_i(c)$, for all $i \in [0...l]$. Substitute this into (3), to get

$$f_c(a) = f + \sum_{i=0}^l \left( f_i(a) + g'_i(c) \right) + e(c) = f(a) + p(c),$$

where

$$p(c) = f + e(c) + \sum_{i=0}^l g'_i(b_{v[i]}), \quad (9)$$

and $f(a)$ is as in (5).
Application to Nilpotent groups: A fully characteristic series is an ascending chain of subgroups:

\[ \{e\} = Z_0 < Z_1 < \ldots < Z_K = B. \]  

where each is fully characteristic in the next. For example, the upper central series of \( B \) is the series (10), where \( Z_1 = Z(B) \), and for each \( k \geq 1 \), \( Z_{k+1} \) is the complete preimage in \( B \) of \( Z(C_k) \) under the quotient map \( B \to C_k := B/Z_k \), until we reach \( K > 0 \) so that \( Z_K = Z_{K+1} = Z_{K+2} = \ldots \). Thus, for all \( k \in [1..K] \), the factor groups \( Q_k = Z_k/Z_{k-1} \) are abelian (but \( C_K = B/Z_K \) is not). In general, \( Z_K \neq B \); if they are equal, then \( B \) is called nilpotent, and \( C \) is trivial.

Example 10: Let \( B = Q_8 \) from Example 5. Then \( Z_1 = Z(Q_8) = \{\pm 1\} \), and \( Q_8/Z_1 \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2 \) is abelian, so that \( Z_2 = Q_8 \). Thus, \( Q_8 \) is nilpotent, with upper central series:

\[ \{1\} \prec \{\pm 1\} \prec Q_8. \]

We can apply Theorem 4 recursively to a totally characteristic series like (10). Let \( A_1 = Z_1 \), and \( C_1 = B/A_1 \), and write \( B = A_1 \rtimes C_1 \), so that \( \mathfrak{S} = \mathfrak{S}_1 \rtimes \mathfrak{S}_1 \), where \( \mathfrak{S}_1 : C_1^M \longrightarrow C_1^M \) and \( \mathfrak{S}_1 : A_1^M \times C_1^M \longrightarrow A_1^M \) are multiplicative. The series (10) induces a fully characteristic series

\[ \{e\} = Z_{1/1} < Z_{2/1} < \ldots < Z_{K/1} = C_1. \]  

where for each \( k \in [1..K] \), we let \( Z_{k/1} = Z_k/A_1 \subset C_1 \). Now let \( A_2 = Z_{2/1} \) and \( C_2 = C_1/A_2 \cong B/Z_2 \), and write \( C_1 = A_2 \rtimes C_2 \), so that \( \mathfrak{S}_1 = \mathfrak{S}_2 \rtimes \mathfrak{S}_2 \), where \( \mathfrak{S}_2 : A_2^M \times C_2^M \longrightarrow A_2^M \) and \( \mathfrak{S}_2 : C_2^M \longrightarrow C_2^M \). Proceed inductively. In particular, if \( B \) is nilpotent, apply this to the upper central series to obtain a decomposition: \( \mathfrak{S} = \mathfrak{S}_1 \rtimes (\mathfrak{S}_2 \star [\ldots \star (\mathfrak{S}_{K-1} \rtimes \mathfrak{S}) \ldots]) \), where, for all \( k \in [1..K] \), \( \mathfrak{S}_k : A_k^M \times C_k^M \longrightarrow A_k^M \) is an affine RCA, while \( \mathfrak{S} : A_K^M \longrightarrow A_K^M \) is an affine CA on the final abelian factor \( A_K = B/Z_{K-1} \).

4 Entropy

Throughout this section, \( M = \mathbb{Z} \) or \( \mathbb{N} \), and \( V = [V_0..V_1] \subset M \). Let \( L = -\min\{V_0, 0\} \) and \( R = \max\{0, V_1\} \), and let \( V = R + L \). Let \( B \) be a finite set, with \( B = \text{card } \mathcal{B} \). If \( c \in B^{[J..K]} \) and \( \langle c \rangle = \{ b \in B^Z : b|_{[J..K]} = c \} \) is the corresponding cylinder set, we say \( \langle c \rangle \) is a cylinder set of length \( \ell = K - J \). Let \( \eta_B \) be the uniformly distributed Bernoulli measure on \( B^Z \), which assigns probability \( B^{-\ell} \) to all cylinder sets of length \( \ell \). If \( Y_1, Y_2 \subset Z \) are disjoint, and \( b_k \in B^{Y_k} \), then we define \( b_1 \Delta b_2 \in B^{Y_1 \cup Y_2} \) by: \( (b_1 \Delta b_2)|_{Y_k} = b_k \), for \( k = 1, 2 \). Thus, \( \langle b_1 \Delta b_2 \rangle = \langle b_1 \rangle \cap \langle b_2 \rangle \). Also, if \( Y \subset Z \) and \( X = Y + V \), then we abuse notation by letting \( \mathfrak{S} : B^X \longrightarrow B^Z \) be the “local map” induced by \( \mathfrak{S} \).
4.1 Permutativity and Relative Permutativity

A local map \( g : B^V \to B \) is left permutative if \( L > 0 \) and, for every \( b \in B^{[-L\ldots R]} \) the map \( B \ni a \mapsto g(a \cdot b) \in B \) is a bijection; \( g \) is right permutative if \( R > 0 \) and, for every \( b \in B^{[-L\ldots R]} \) the map \( B \ni c \mapsto g(b \cdot c) \in B \) is a bijection. \( g \) is permutative if it is either left- or right-permutative. A CA on \( B \) sides; a CA on \( B \) is left-permutative, then for all \( \in B \) \( b \cdot c \in B \) is a bijection. \( g \) is permutative if it is either left- or right-permutative. A CA on \( B^2 \) is bipermutative if it is permutative on both sides; a CA on \( B^N \) is called bipermutative if it is right-permutative. If \( V = R + L \) then we say \( g \) is \( V \)-bipermutative. A nonhomogeneous CA \( \mathcal{G} \) is \( V \)-bipermutative if \( g_m \) is \( V \)-bipermutative for every \( m \in M \). A relative CA \( \mathcal{F} : A^M \times C^M \to A^M \) is \( V \)-bipermutative if the NHCA \( \mathcal{F}_c \) is \( V \)-bipermutative for every \( c \in C^M \).

Example 11:

(a) If \( B = (\mathbb{Z}_n, +) \) and \( V_0 < 0 < V_1 \), then a linear CA with local map \( g(b) = \sum_{v=V_0}^{V_1} g_v \cdot b_v \)
is left- (resp. right-) permutative iff \( g_{V_0} \) (resp. \( g_{V_1} \)) is relatively prime to \( n \).

(b) In Example 13, \( \mathcal{F} \) is a 3-bipermutative CA, while \( \mathcal{F} \) is a 3-bipermutative RCA.

(c) In Example 13, \( h(c_0, c_1, c_2) = h(c_0, c_1) = c_0 - c_1 \) is right-permutative, with \( V = \{0, 1\} \),so \( \mathcal{F} \) is 2-bipermutative as a map on \( C^N \). However, \( \mathcal{F} \) is not right-permutative. To see this, write:

\[
\begin{align*}
    f(c_0, c_1, c_2)(a_0, a_1, a_2) &= f_0(c_0, c_1, c_2)(a_0) + f_1(c_0, c_1, c_2)(a_1) + f_2(c_0, c_1, c_2)(a_2).
\end{align*}
\]

Then:

\[
\begin{align*}
    f_2(c_0, c_1, c_2) &= 1 + 2c^2 + 2^{2c^2} + 2^{3c^2} = \begin{cases} 
    4 \pmod{5} & \text{if } c_2 = 0 \text{ or } 4; \\
    0 \pmod{5} & \text{if } c_2 = 1, 2 \text{ or } 3;
\end{cases}
\end{align*}
\]

Thus, \( f(c_0, c_1, c_2) \) is right-permutative if and only if \( c_2 = 0 \) or 4.

The following results extend well-known properties of permutative cellular automata to the nonhomogeneous case; the proofs are similar, and are left to the reader.

Lemma 12 Let \( J < K \), and \( \ell = K - J \). Let \( \mathcal{G} \) be an NHCA, and let \( d \in B^{[J+R\ldots K+R]} \).

1. If \( \mathcal{G} \) is right-permutative, then for all \( b \in B^{[J-L\ldots J+R]} \), there is a unique \( c \in B^{[J+R\ldots K+R]} \) so that \( \mathcal{G}(b \cdot c) = d \).

2. If \( \mathcal{G} \) is left-permutative, then for all \( b \in B^{[K-L\ldots K+R]} \), there is a unique \( a \in B^{[J-L\ldots K-L]} \) so that \( \mathcal{G}(a \cdot b) = d \).

3. If \( \mathcal{G} \) is bipermutative, then for any \( j \in [J\ldots K] \) and \( b \in B^{[j-L\ldots j+R]} \), there are unique \( a \in B^{[j-L\ldots j-L]} \) and \( c \in B^{[j+R\ldots K+R]} \) so that \( \mathcal{G}(a \cdot b \cdot c) = d \).

\[ \square \]

Corollary 13 If \( \mathcal{G} \) is permutative, then \( \eta_\mathcal{G} \) is \( \mathcal{G} \)-invariant.
4.2 Measurable Entropy

Suppose \((Y, \mathcal{Y}, \mu)\) is a probability space, \(Q\) is a finite set, and \(Q : Y \to Q\) is measurable; we say \(Q\) is a partition of \(Y\), indexed by \(Q\). If \(\rho = Q(\mu) \in \mathcal{M}(Q)\), then let
\[
h(Q; \mu) = -\sum_{q \in Q} \rho[q] \cdot \log(\rho[q]).\]
If \(Q_k : Y \to Q_k\) for \(k = 1, 2\), then let \(Q_1 \vee Q_2 : Y \to Q_1 \times Q_2\) be the partition mapping \(y \in Y\) to \((Q_1(y), Q_2(y))\). If \(G : Y \to Y\) is \(\mu\)-preserving, then define \(GQ = Q \circ G\). If \(G\) is invertible (respectively noninvertible), let \(\mathcal{T} = \mathbb{Z}\) (respectively \(\mathcal{T} = \mathbb{N}\)); for any \(I \subset \mathcal{T}\), define \(G^I Q = \bigvee_{i \in I} G^i Q : Y \to Q^I\). The \(G\)-entropy of \(Q\) is the limit:
\[
h(Q, G, \mu) = \lim_{N \to \infty} \frac{1}{N} h(G^{(0, ..., N)}Q; \mu).
\]

Let \(\Sigma(G^T Q)\) be the smallest \(\sigma\)-algebra for which the function \(G^T Q\) is measurable. \(Q\) is a generator for \(G\) if \(\Sigma(G^T Q)\) is \(\mu\)-dense in \(\mathcal{Y}\), meaning that, for all \(V \in \mathcal{Y}\), there is \(W \in \Sigma(G^T Q)\) so that \(\mu(V \Delta W) = 0\); the (measurable) entropy of the system \((Y, \mathcal{Y}, \mu; G)\) is then defined: \(h(Y, \mathcal{Y}, \mu; G) = h(Q, G, \mu)\) (independent of the choice of generator).

Suppose that \(\{R_n\}_{n=0}^{\infty}\) is a collection of NHCA on \(\mathcal{B}^M\). For any \(n \in \mathbb{N}\), define \(R^{(n)} = R_{n-1} \circ R_{n-2} \circ \ldots \odot R_0\). If \(Q\) is a partition, then, for all \(N \in \mathbb{N}\), define \(\mathcal{R}^{(0, ..., N)}Q = \bigvee_{n=0}^{N-1} \mathcal{R}^{(n)}Q\).

Thus, if \(\mathcal{R}_n = \mathcal{G}\) for all \(n \in \mathbb{N}\), then \(\mathcal{R}^{(n)} = \mathcal{G}^n\), and \(\mathcal{R}^{(0, ..., N)}Q = \mathcal{G}^{(0, ..., N)}Q\).

Say that \(Q_k : Y \to Q_k\) \((k = 1, 2)\) are equivalent if each is measurable with respect to the other. Then, for any \(\mu \in \mathcal{M}(Y)\), \(h(Q_1; \mu) = h(Q_2; \mu)\).

**Proposition 14** Let \(\mathcal{R}_n\) be \(V\)-bipermutative, for all \(n \in \mathbb{N}\). Let \(\mu \in \mathcal{M}(\mathcal{B}^M)\) and let \(Q = \text{pr}_{[-L .... R]} : \mathcal{B}^M \to \mathcal{B}^{[-L .... R]}\). Then:

1. \(\mathcal{R}^{(0, ..., N)}Q\) and \(\text{pr}_{[-NL .... NR]}\) are equivalent. Thus, \(h(\mathcal{R}^{(0, ..., N)}Q; \mu) = h(\text{pr}_{[-NL .... NR]}; \mu)\).
2. \(\Sigma(\mathcal{R}^{(N)}Q)\) is the Borel sigma-algebra of \(\mathcal{B}^M\).

In particular, suppose \(\mathcal{R}_n = \mathcal{G}\), for all \(n \in \mathbb{N}\). Then:

3. \(Q\) is a \((\mathcal{G}, \mu)\)-generator.
4. If \(\mu\) is \(\sigma\)-invariant and \(\mathcal{G}\)-invariant, then \(h(\mathcal{G}; \mu) = V \cdot h(\sigma; \mu)\).

In particular, \(h(\mathcal{G}; \eta_\mathcal{B}) = V \cdot \log(B)\).

**Proof:** Part 1 is proved by repeated application of Part 3 from Lemma [12]. The other statements then follow. \(\square\)

**Remark:** By combining Part 4 with Example [12], we recover the previously computed entropy of linear CA on \((\mathbb{Z}/p)^\mathbb{Z}, \eta_\mathcal{B})\).
4.3 Relative Entropy

Let $Y = X \times Z$, $G = F \ast H$, and $\lambda \in M [X]$ be as in §3.1. If $\nu \in M [Z]$ is $H$-invariant, then $\mu = \lambda \otimes \nu$ is $G$-invariant. For any $z \in Z$ and $n \in \mathbb{N}$, define $F^{(n)}_z = F_{H^{n-1}(z)} \circ \cdots \circ F_{H^1(z)} \circ F_z$.

If $P : X \rightarrow \mathcal{P}$ is a partition, then for all $z \in Z$, let $G^{[0..N]}_z P : X \rightarrow \mathcal{P}^N$ be the partition:

$$G^{[0..N]}_z P = \bigvee_{n=0}^{N-1} F^{(n)}_z P,$$

and define $h(P; F, \lambda/ H, \nu) = \lim_{N \to \infty} \frac{1}{N} \int_Z h(G^{[0..N]}_z P; \lambda) \, d\nu[z]$. The relative entropy of $F$ over $H$ is defined: $h(F, \lambda/ H, \nu) = \sup_P h(P; F, \lambda/ H, \nu)$, where the supremum is taken over all measurable partitions $P$ of $X$.

**Theorem 15** $h(G, \lambda \otimes \nu) = h(H, \nu) + h(F, \lambda/ H, \nu)$. — (Abramov & Rokhlin) [3, 13].

**Theorem 16** Let $\mathcal{B} = A \times C$. Let $\lambda \in M [A^M]$, $\nu \in M [C^M]$, and $\mu = \lambda \otimes \nu \in M [B^M]$ be $\sigma$-invariant. Suppose $\mathcal{G} = \mathcal{F} \ast \mathcal{F}$, where $\mathcal{F} : A^M \times C^M \rightarrow A^M$ is $\lambda$-preserving and $V$-bipermutative, while $\mathcal{F} : C^M \rightarrow C^M$ is $\nu$-preserving and $W$-bipermutative. Then:

1. $h(\mathcal{G}, \lambda/ \mathcal{F}, \nu) = V \cdot h(\lambda, \sigma)$;
2. $h(\mathcal{F}, \nu) = W \cdot h(\nu, \sigma)$;
3. $h(\mathcal{G}, \mu) = V \cdot h(\lambda, \sigma) + W \cdot h(\nu, \sigma)$.

**Proof:** To see (1), let $Q = \text{pr}_{[-L..R]} : A^M \rightarrow A^{[-L..R]}$. Fix $c \in C^M$, and, for all $n \in \mathbb{N}$, let $\mathcal{R}_n = \mathcal{F}^{(n)}(c)$. Then $\mathcal{R}^{(n)} = \mathcal{F}^{(n)}(c)$, and $\mathcal{R}^{[0..N]} Q = \mathcal{G}^{[0..N]} Q$, so Part 1 of Proposition [14] says: $h(\mathcal{G}^{[0..N]} Q, \lambda) = h(\text{pr}_{[-NL..NR]}; \lambda)$.

Let $P : A^M \rightarrow \mathcal{P}$ be any other partition of $A^M$, and fix $\epsilon > 0$. Then, by Part 2 of Proposition [14], there is some $M = M(c) > 0$ so that, for all $N \in \mathbb{N}$, $h(\mathcal{G}^{[0..N]} P; \lambda) < N \epsilon + h(\mathcal{G}^{[0..N]} Q; \lambda)$. If $\text{card}[P] = P$, then also, $h(\mathcal{G}^{[0..N]} P; \lambda) < N \log(P)$. Find $M$ so that $\mu[\mathcal{D}] < \epsilon/\log(P)$, where $\mathcal{D} = \{ c \in C^M : M(c) > M \}$. Then

$$\int_{C^M} h(\mathcal{G}^{[0..N]} P; \lambda) \, d\nu[c] = \int_{\mathcal{D}} h(\mathcal{G}^{[0..N]} P; \lambda) \, d\nu[c] + \int_{C^M \setminus \mathcal{D}} h(\mathcal{G}^{[0..N]} P; \lambda) \, d\nu[c] < \left( N \log(P) \frac{\epsilon}{\log(P)} \right) + N \epsilon + h(\text{pr}_{[-(N+M)L..(N+M)R]}; \lambda).$$

Thus, $h(P; \mathcal{F}, \lambda/ \mathcal{F}, \nu) < 2 \epsilon + \lim_{N \to \infty} \frac{1}{N} h(\text{pr}_{[-NL..NR]}; \lambda) = 2 \epsilon + V \cdot h(\sigma; \lambda).$ Take the supremum over all $P$ to conclude: $h(\mathcal{G}, \lambda/ \mathcal{F}, \nu) < 2 \epsilon + V \cdot h(\sigma; \lambda)$. Now let $\epsilon \to 0$. (2) follows from Part 4 of Proposition [14] and (3) follows from Theorem [14].
Note that, in Theorem 16, \( B \) need not be a group, nor \( G \) a multiplicative cellular automaton. However, Theorem 4 provides a natural skew product decomposition in this case.

**Example 17:** In Example (11b), \( V = W = 2, \) \( \text{card } \mathcal{A} = 5 \) and \( \text{card } \mathcal{C} = 4; \) thus \( h(\mathcal{G}, \eta_\mathcal{B}) = 2 \cdot \log_2(5) + 2 \cdot \log_2(4) = 2 \log_2(5) + 4. \)

## 5 Convergence of Measures

Endow \( \mathcal{M} [\mathcal{B}^\mathfrak{M}] \) with the weak* topology induced by \( \mathcal{C} (\mathcal{B}^\mathfrak{M}; \mathbb{C}), \) the space of continuous, complex-valued functions. The uniformly distributed Bernoulli measure \( \eta_\mathcal{B} \in \mathcal{M} [\mathcal{B}^\mathfrak{M}] \) is the Haar measure on \( \mathcal{B}^\mathfrak{M} \) as a compact group, and is invariant under the action of any left- or right-permutative MCA (Lemma 13). Thus, if \( \mu \in \mathcal{M} [\mathcal{B}^\mathfrak{M}] \) is some initial measure, then \( \eta_\mathcal{B} \) is a natural candidate for the (Cesàro) limit of \( \mathcal{G}^n \mu \) as \( n \to \infty. \) Since \( \eta_\mathcal{B} \) is the measure of maximal entropy on \( \mathcal{B}^\mathfrak{M}, \) such limiting behaviour is a sort of “asymptotic randomization” of \( \mathcal{B}^\mathfrak{M}. \) When \( \mathcal{B} \) is abelian, and \( \mathcal{G} \) is an affine CA, the Cesàro convergence of measures to \( \eta_\mathcal{B} \) is somewhat understood [2, 6, 8, 7, 12]; we now extend these results to nonabelian MCA.

**Harmonic Mixing and Diffusion:** The characters of an abelian group \( \mathcal{A}, + \) are the continuous homomorphisms from \( \mathcal{A} \) into the unit circle group \( (\mathbb{T}^1, \cdot) \subset \mathbb{C}. \) The set of characters forms a group, denoted \( \hat{\mathcal{A}}. \) For example, if \( \mathcal{A} = \mathbb{Z}/n, \) then every \( \chi \in \hat{\mathcal{A}} \) has the form \( \chi(a) = \exp \left( \frac{2\pi i}{n} c \cdot a \right), \) where \( c \in \mathbb{Z}/n \) is a constant coefficient, and the product \( c \cdot a \) is computed mod \( n. \)

For any \( \chi \in \hat{\mathcal{A}}^\mathfrak{M} \) there is some finite \( \mathfrak{K} \subset \mathfrak{M} \) and, for each \( k \in \mathfrak{K}, \) a nontrivial \( \chi_k \in \hat{\mathcal{A}}, \) so that, if \( a \in \mathcal{A}^\mathfrak{M}, \) then \( \chi(a) = \prod_{k \in \mathfrak{K}} \chi_k(a_k); \) we indicate this: “\( \chi = \bigotimes_{k \in \mathfrak{K}} \chi_k. \)” For example, if \( \mathcal{A} = \mathbb{Z}/n, \) this means there is a collection of nonzero coefficients \( [c_k]_{k \in \mathfrak{K}} \) so that if \( a \in \mathcal{A}^\mathfrak{M}, \) then \( \chi(a) = \exp \left( \frac{2\pi i}{n} \sum_{k \in \mathfrak{K}} c_k \cdot a_k \right). \)

The **rank** of \( \chi \in \hat{\mathcal{A}}^\mathfrak{M} \) is the cardinality of \( \mathfrak{K}. \) Let \( \mu \in \mathcal{M} [\mathcal{A}^\mathfrak{M}]; \) we will use the notation \( \langle \chi, \mu \rangle = \int_{\mathcal{A}^\mathfrak{M}} \chi \ d\mu. \) We call \( \mu \) **harmonically mixing** (and write “\( \mu \in \mathcal{H} \mathcal{M} [\mathcal{A}^\mathfrak{M}]. \)”) if, for all \( \epsilon > 0, \) there is \( r \in \mathbb{N} \) so that, if \( \chi \in \hat{\mathcal{A}}^\mathfrak{M} \) and \( \text{rank } \chi > r, \) then \( |\langle \mu, \chi \rangle| < \epsilon. \) For example, most Bernoulli measures \( \Box \) and Markov random fields \( \square \) on \( \mathcal{A}^{(Z^D)} \) are harmonically mixing.

If \( \mathcal{G} \in \text{End } [\mathcal{A}^\mathfrak{M}]; \) then for any \( \chi \in \hat{\mathcal{A}}^\mathfrak{M}, \) the map \( \chi \circ \mathcal{G} \) is also a character. If \( \mathcal{J} \subset \mathbb{N}, \) then \( \mathcal{G} \) is **\mathcal{J}-diffusive** if, for every \( \chi \in \hat{\mathcal{A}}^\mathfrak{M}, \) \( \lim_{j \in \mathcal{J}} \text{rank } \left[ \chi \circ \mathcal{G}^j \right] = \infty. \) If \( \mathcal{J} = \mathbb{Z}, \) then we just say \( \mathcal{G} \) is **diffusive**; if \( \mathcal{J} \subset \mathbb{N} \) is a subset of Cesàro density 1, then \( \mathcal{G} \) is **diffusive in density**.
Example 18: Let $(\mathcal{A}, +)$ be a finite abelian group, $\mathbb{M} = \mathbb{Z}^D$, and let $\mathcal{L}$ be a linear CA with local map $I(a) = \sum_{v \in \mathcal{V}} \ell_v a_v$, where $\ell_v \in \mathbb{Z}$ is relatively prime to $\text{card } [\mathcal{A}]$ for all $v \in \mathcal{V}$.

Then $\mathcal{L}$ is diffusive in density \[\eta.\]

If $\mathcal{G}$ is $\mathbb{J}$-diffusive and $\mu \in \mathcal{H} \mathcal{M} [\mathcal{A}^M]$, then Theorem 12 of $\mathcal{G}$ says $\nu^* \lim_{\mathcal{J}, j \to \infty} \mathcal{G}^n \mu = \eta_\mathcal{G}$.

In particular, if $\mathcal{G}$ is diffusive in density, then the Cesàro average weak*-converges to Haar measure:

$$\nu^* \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mathcal{G}^n \mu = \eta_\mathcal{G}. \quad (12)$$

To extend these results to multiplicative cellular automata, we need a version of diffusion applicable to affine relative cellular automata. An affine endomorphism of $\mathcal{A}^M$ is a self-map $\mathcal{G} : \mathcal{A}^M \to \mathcal{A}^M$ of the form $\mathcal{G}(a) = c + \mathcal{L}(a)$, where $\mathcal{L} \in \mathcal{E} \mathcal{N} [\mathcal{A}^M]$ and $c \in \mathcal{A}^M$ is constant. The set of affine endomorphisms is denoted $\mathcal{E} \mathcal{N} [\mathcal{A}^M]$.

An affine character is a function $\alpha : \mathcal{A}^M \to \mathbb{T}^1$ of the form $\alpha(a) = c \cdot \chi(a)$, where $c \in \mathbb{T}^1$ is a constant, and $\chi \in \mathcal{A}^M$. The set of affine characters of $\mathcal{A}^M$ is denoted $\mathcal{A}^M$. For example, if $\zeta \in \mathcal{A}^M$, and $\mathcal{G} \in \mathcal{E} \mathcal{N} [\mathcal{A}^M]$, then $\zeta \circ \mathcal{G} \in \mathcal{A}^M$. The rank of $\alpha = c \cdot \chi$ is the rank of $\chi$. If $\mu \in \mathcal{M} [\mathcal{A}^M]$ is harmonically mixing, with $\epsilon$ and $r$ as before, then it follows that $|\langle \mu, \alpha \rangle| < \epsilon$ for any $\alpha \in \mathcal{A}^M$ with rank $[\alpha] > r$.

Relative Diffusion: Suppose $\mathcal{B} = \mathcal{A} \times \mathcal{C}$, where $\mathcal{A}$ is abelian, and let $\mathcal{G} = \mathfrak{F} \star \mathfrak{H}$ as in Theorem $\mathfrak{H}$. For any $c \in \mathcal{C}^M$, the fibre map $\mathfrak{F}_c$ is an affine endomorphism; we say that $\mathfrak{F}$ is an affine relative cellular automaton (ARCA). For any $j \in \mathbb{N}$, $\mathcal{G}^j = \mathfrak{F}^{(j)} \star \mathfrak{H}^j$, where $\mathfrak{F}^{(j)}$ is another ARCA, so $\alpha \circ \mathfrak{F}^{(j)}_c \in \mathcal{A}^M$ for any $c \in \mathcal{C}^M$ and $\alpha \in \mathcal{A}^M$. We say $\mathcal{G}$ is relatively $\mathbb{J}$-diffusive if $\lim_{J, j \to \infty} \text{rank } [\alpha \circ \mathfrak{F}^{(j)}_c] = \infty$ for every $c \in \mathcal{C}^M$ and $\alpha \in \mathcal{A}^M$. If $\nu \in \mathcal{M} [\mathcal{C}^M]$, and $J \subset \mathbb{N}$, then $\mathcal{G}$ is $\nu$-relatively $\mathbb{J}$-diffusive if,

$$\forall \alpha \in \mathcal{A}^M, \forall r > 0, \lim_{J, j \to \infty} \nu \{ c \in \mathcal{C}^M : \text{rank } [\alpha \circ \mathfrak{F}^{(j)}_c] \leq r \} = 0. \quad (13)$$

Clearly, relative diffusion implies $\nu$-relative diffusion for any $\nu \in \mathcal{M} [\mathcal{C}^M]$.

Proposition 19 If $\mathcal{A} \subset Z(\mathcal{B})$ as in Part 3 of Proposition $\mathfrak{H}$, then $\mathcal{F}$ is relatively $\mathbb{J}$-diffusive if and only if $\mathcal{L}$ is $\mathbb{J}$-diffusive as a linear cellular automaton.

Proof: For any $N \in \mathbb{N}$, define $\mathcal{P}^{(N)} = \sum_{n=0}^{N-1} \mathcal{L}^n \circ \mathcal{F} \circ \mathfrak{H}^{N-n-1}$. If $j \in \mathbb{J}$, then $\mathfrak{F}^{(j)}_c = \mathcal{L}^j + \mathcal{P}^{(j)}(c)$. Thus, for any $\alpha \in \mathcal{A}^M$, $\text{rank } [\alpha \circ \mathfrak{F}^{(j)}_c] = \text{rank } [\alpha \circ \mathcal{L}^j]$.  

\[\Box\]
Proposition 20 Let $\lambda \in \mathcal{H} \mathcal{M}[A^M]$, $\nu \in \mathcal{M}[C^M]$, and $\mu = \lambda \otimes \nu \in \mathcal{M}[B^M]$. Let $\mathcal{V} = \text{wk}^* \lim_{j \to \infty} \mathcal{S}^j \nu$, and let $\eta_A$ be the Haar measure on $A^M$. If $\mathcal{G}$ is $\nu$-relatively $\mathcal{H}$-diffusive, then $\text{wk}^* \lim_{j \to \infty} \mathcal{G}^j \mu = \eta_A \otimes \mathcal{V}$.

Proof: We want $\lim_{j \to \infty} \langle \mathcal{G}^j \mu \rangle = \langle \beta, \eta_A \otimes \mathcal{V} \rangle$, for every $\beta \in C(B^M; \mathbb{C})$. It suffices to assume $\beta = \alpha \otimes \phi$, where $\alpha \in \mathcal{C}(A^M; \mathbb{C})$ and $\phi \in \mathcal{C}(C^M; \mathbb{C})$. Since the characters of $A^M$ form a basis for the Banach space $C(A^M; \mathbb{C})$, it suffices to assume $\alpha \in \overline{A^M}$, and $\|\phi\|_\infty = 1$.

Then: $\langle \beta, \eta_A \otimes \mathcal{V} \rangle = \langle \alpha, \eta_A \rangle \cdot \langle \phi, \mathcal{V} \rangle = \left\{ \begin{array}{ll} \langle \phi, \mathcal{V} \rangle & \text{if } \alpha = \mathbb{1} \\ 0 & \text{if } \alpha \neq \mathbb{1} \end{array} \right.$

Now, for all $a \in B^M$, $\beta \circ \mathcal{G}^j(a \circ c) = (\alpha \otimes \phi)\mathcal{G}^j c \alpha \circ \mathcal{G}^j(c) = (\alpha \circ \mathcal{G}^j c)(\alpha \circ \mathcal{G}^j(c))$. Thus, $\langle \beta, \mathcal{G}^j \mu \rangle = \langle \beta \circ \mathcal{G}^j, \mu \rangle = \int_{\mathcal{C}M} (\phi \circ \mathcal{G}^j(c)) \cdot (\alpha \circ \mathcal{G}^j(c), \lambda) \, d\nu[c]$. If $\alpha = \mathbb{1}$, then this integral is just equal to $\int_{\mathcal{C}M} \phi \circ \mathcal{G}^j(c) \, d\nu[c]$, which converges to $\langle \phi, \mathcal{V} \rangle$ by hypothesis. Hence, assume $\alpha \neq \mathbb{1}$; we then want to show that $\lim_{j \to \infty} \langle \beta, \mathcal{G}^j \mu \rangle = 0$.

Fix $\epsilon > 0$. Since $\lambda \in \mathcal{H} \mathcal{M}[A^M]$, find $r > 0$ so that, if $\alpha \in \overline{A^M}$ and $\text{rank} [\alpha] > r$, then $|\langle \alpha, \lambda \rangle| < \frac{\epsilon}{2}$. Let $D_j = \{ c \in C^M : \text{rank} [\alpha \circ \mathcal{G}^j c] > r \}$, for every $\lambda \in \lambda$. By equation (13), find $J \in \mathbb{N}$ so that, $\forall j \in J$, $|\langle \alpha, \lambda \rangle| < \frac{\epsilon}{2}$.

Then $|\langle \beta, \mathcal{G}^j \mu \rangle| = \left| \int_{\mathcal{C}M} (\phi \circ \mathcal{G}^j(c) \cdot (\alpha \circ \mathcal{G}^j(c), \lambda) \, d\nu[c] \right| \leq \int_{D_j} (\phi \circ \mathcal{G}^j(c) \cdot (\alpha \circ \mathcal{G}^j(c), \lambda) \, d\nu[c] \right| + \int_{\mathcal{C}M \setminus D_j} (\phi \circ \mathcal{G}^j(c) \cdot (\alpha \circ \mathcal{G}^j(c), \lambda) \, d\nu[c] \right|$ \leq \int_{D_j} (\phi \circ \mathcal{G}^j(c) \cdot |(\alpha \circ \mathcal{G}^j(c), \lambda) | \, d\nu[c] \right| + \int_{\mathcal{C}M \setminus D_j} (\phi \circ \mathcal{G}^j(c) \cdot (\alpha \circ \mathcal{G}^j(c), \lambda) \, d\nu[c] \right| \leq \int_{D_j} \frac{\epsilon}{2} \, d\nu[c] \right| + \int_{\mathcal{C}M \setminus D_j} \frac{\epsilon}{2} \, d\nu[c] \right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. 

Suppose $B$ is nilpotent, with upper central series (13). If $k \in [1..K]$, then $Q_k = Z_k/Z_{k-1}$ is abelian, and the decomposition $B = Q_1 \ast (Q_2 \ast \ldots (Q_{K-1} \ast Q_K) \ldots)$ induces a natural identification $B^M \cong Q_1^M \times Q_2^M \times \ldots \times Q_K^M$. If $\lambda_k \in \mathcal{M}[Q_k^M]$ for all $k$, then $\lambda_1 \otimes \ldots \otimes \lambda_K \in \mathcal{M}[B^M]$. Let $\mathcal{H} \mathcal{M}[B^M]$ denote the convex, weak*-closure in $\mathcal{M}[B^M]$ of the set

$$\{ \lambda_1 \otimes \ldots \otimes \lambda_K : \lambda_k \in \mathcal{H} \mathcal{M}[Q_k^M] \text{ for all } k \}$$.\]
Theorem 21  Suppose \( B \) is nilpotent, and \( \mathfrak{G} \) has a local map of the form \( g(b) = b_{n_1}v_1b_{n_2}v_2 \cdots b_{n_J}v_J \). For every \( v \in V \), suppose \( \ell_v = \sum_{v_j = v} n_j \) is relatively prime to \( \text{card} [B] \).

If \( \mu \in \mathcal{H}M[B,M] \), then \( \lim_{J \to \infty} \mathfrak{G}^j(\mu) = \eta_B \) along a set \( J \subset \mathbb{N} \) of density one. Thus, equation (12) holds.

Proof:  We’ll prove this by induction on \( K \), the length of the series (10). If \( K = 1 \), then \( B \) is abelian; then \( \mathfrak{G} \) is diffusive in density by Example 18, and the result follows from Theorem 12 of [8].

If \( K > 1 \), then let \( A = Q_1 = Z(B) \), and \( C = B/A \). Thus \( \mathcal{H}M[B,M] \) is the convex weak* closure of \( S = \{ \lambda \otimes \nu; \lambda \in \mathcal{H}M[A^N], \nu \in \mathcal{H}M[C^N] \} \), so it suffices to prove the theorem for \( \mu = \lambda \otimes \nu \in S \). Let \( \mathfrak{F} = \mathfrak{F} \ast \mathfrak{H} \). Then \( \mathfrak{H} \) has local map \( h(c) = c_{n_1}v_1c_{n_2}v_2 \cdots c_{n_J}v_J \), and, by hypothesis, all \( \ell_v \) are all relatively prime to \( \text{card} [C] \). But \( C \) has an upper central series like (11) of length \( K - 1 \), so by induction hypothesis, there is a set \( K \subset \mathbb{N} \) of density one so that \( \lim_{K \to \infty} \mathfrak{H}^k(\nu) = \eta_C \).

Since \( A = Z(B) \), let \( \mathfrak{L} \) be as in Part 3 of Proposition 8. As in Example 18, \( l(a) = \sum_{v \in V} \ell_v \cdot a_v \), and, by hypothesis, \( \ell_v \) are all relatively prime to \( \text{card} [A] \), so, as in Example 18, \( \mathfrak{L} \) is \( \mathbb{I} \)-diffusive for some subset \( \mathbb{I} \subset \mathbb{N} \) of density one. Proposition 19 then implies that \( \mathfrak{F} \) is relatively \( \mathbb{I} \)-diffusive. Let \( \mathbb{J} = \mathbb{I} \cap \mathbb{K} \), also a set of density one. Then apply Proposition 20 to conclude that \( \lim_{J \to \infty} \mathfrak{G}^j(\mu) = \eta_A \otimes \eta_C = \eta_B \).

Example 22:  Recall \( \mathfrak{G} : Q_8^Z \longrightarrow Q_8^Z \) from Example (9e). In this case, \( A \cong \mathbb{Z}/2 \), and \( \mathfrak{L} \), having local map \( l(a_0, a_1, a_2, a_3) = a_0 + a_1 + a_2 + a_3 \), is diffusive in density, so \( \mathfrak{F} \) is relatively diffusive in density. Meanwhile, \( C = \mathbb{Z}/2 \oplus \mathbb{Z}/2 \) and \( \mathfrak{H} \), with local map \( h(c_0, c_1, c_2, c_3) = c_0 + c_1 + c_2 + c_3 \), is also diffusive in density. Hence, equation (12) holds for any \( \mu \in \mathcal{H}M[Q_8^Z] \).

6 Conclusion

Multiplicative cellular automata over a group \( B \) inherit a natural structural decomposition from \( B \). Using this decomposition, we can compute the measurable entropy of MCA, and show that a broad class of initial measures converge to the Haar measure in Cesàro average. However, many questions remain. For example, it is unclear how to show relative diffusion when \( A \) is not central in \( B \). Indeed, even non-relative diffusion is mysterious for noncyclic abelian groups [7]. Also, computation of relative entropy will be much more complicated in
the case of ‘variably permutative’ relative CA, such as Example (11c); perhaps this requires some ‘relative’ version of Lyapunov exponents [15, 16].

Permutative MCA are a considerable generalization of the linear cellular automata previously studied, but they are still only a very special class of permutative cellular automata. The asymptotics of measures for general permutative CA [1] is still poorly understood.

References

[1] Alejandro Maass and Servet Martínez. On Cesàro limit distribution of a class of permutative cellular automata. Journal of Statistical Physics, 90(1-2):435–452, 1998.

[2] Alejandro Maass and Servet Martínez. Time averages for some classes of expansive one-dimensional cellular automata. In Eric Goles and Servet Martínez, editors, Cellular Automata and Complex Systems, pages 37–54. Kluwer Academic Publishers, Dordrecht, 1999.

[3] H. Furstenberg. Ergodic Theory and Combinatorial Number Theory. Princeton University Press, Princeton, New Jersey, 1981.

[4] G. Hedlund. Endomorphisms and automorphisms of the shift dynamical systems. Mathematical System Theory, 3:320–375, 1969.

[5] L. M. Abramov and V. A. Rokhlin. The entropy of a skew product of measure-preserving transformations. American Mathematical Society Translations (Series 2), 48:225–265, 1965.

[6] Douglas Lind. Applications of ergodic theory and sofic systems to cellular automata. Physica D, 10:36–44, 1984.

[7] Marcus Pivato and Reem Yassawi. Limit measures for affine cellular automata II. Submitted to Ergodic Theory and Dynamical Systems; preprint available at: http://arXiv.org/abs/math.DS/0108083, April 2001.

[8] Marcus Pivato and Reem Yassawi. Limit measures for affine cellular automata. To appear in Ergodic Theory and Dynamical Systems; preprint available at: http://arXiv.org/abs/math.DS/0108082, 2002.

[9] Giovanni Manzini Michele D’amico and Luciano Margara. On computing the entropy of cellular automata. In G. Winskel K. G. Larsen, S. Skyum, editor, 25th International Colloquium on Automata, Languages, and Programming (ICALP ’98), number 1443 in Lecture Notes in Computer Science, pages 470–481, Heidelberg, 1998. Springer-Verlag.

[10] Cris Moore. Quasi-linear cellular automata. Physica D, 103:100–132, 1997.

[11] Cris Moore. Predicting non-linear cellular automata quickly by decomposing them into linear ones. Physica D, 111(1-4):27–41, Jan 1998.

[12] Servet Martínez Pablo Ferrari, Alejandro Maass and Peter Ney. Cesàro mean distribution of group automata starting from measures with summable decay. Ergodic Theory and Dynamical Systems, 20(6):1657–1670, 2000.

[13] Karl Petersen. Ergodic Theory. Cambridge University Press, New York, 1989.

[14] W. R. Scott. Group Theory. Dover, 1987.
[15] Mark A. Shereshevsky. Lyapunov exponents for one-dimensional cellular automata. *Journal of Nonlinear Science*, 2:1–8, 1992.

[16] P. Tisseur. Cellular automata and Lyapunov exponents. *Nonlinearity*, 13:1547–1569, 2000.