Order Number 9405515

The Kolmogorov metric and a classification of linear cellular automata

D'Alotto, Louis A., Ph.D.

City University of New York, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A

THE KOLMOGOROV METRIC AND A CLASSIFICATION OF LINEAR CELLULAR AUTOMATA

by

LOUIS A. D'ALOTTO

A dissertation submitted to the Graduate Faculty in Computer Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1993

This manuscript has been read and accepted for the Graduate Faculty in Computer Science in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Sept 28 '93
Date

Chair of Examining Committee

Sept 28 '93
Date

Executive Officer

Edward Beckenstein

Divyendu Sinha

Stanley Habib

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

THE KOLMOGOROV METRIC AND A CLASSIFICATION

OF

LINEAR CELLULAR AUTOMATA

by

LOUIS A. D'ALOTTO

Advisor: Professor Charles Giardina

The idea of classifying linear cellular automata by their dynamical behavior was initiated by Stephen Wolfram. Robert Gilman later brought this idea into a mathematical perspective when he introduced a probabilistic/topological classification of linear automata. This dissertation generalizes the classification results of Gilman by utilizing a more general, and more applicable, non-archimedean metric that was originally devised by Kolmogorov. The classification utilizes methods and results from point set topology, product measure theory and ergodic theory. Linear automata are then divided into three classes whereby the automata in each class seem to correspond to different dynamical behavior. These classification results are then applied to the convolution operation of signal processing to classify the functions of this operation for signals taking values in Z_2.

This Dissertation is Dedicated

to

The Two That Make It All Worth While

My Wife Zana

and

Daughter Angela

Acknowledgments

I would first like to express my most profound gratitude to my advisor Charles Giardina for his leadership, inspiration, mathematical expertise and all his tireless efforts in guiding me through this work. My sincere appreciation goes to Edward Beckenstein whose mathematical knowledge and long term friendship have always been invaluable to me from my beginning undergraduate days at St. John's through the completion of this dissertation. I wish to thank my other committee members, Professor Stanley Habib and especially Divyendu Sinha for his numerous and useful suggestions which helped in the writing of this dissertation. I also wish to thank Robert Gilman for his original classification work on cellular automata and his conversations in helping me to understand its principles.

Appreciation goes to my parents Lou and Angela D'Alotto and my brother Roc for all their moral support and their belief in my capabilities for which I am very grateful. I would also like to thank my parents for all their financial support over the past years. Lastly, I would like to express my most special appreciation to my wife Zana for all her love, support and her faith in me to complete this dissertation, as well as for her skillful proofreading and spelling abilities.

Preface

With the development of smaller and faster computers, cellular automata have recently came into the spotlight as tools to model natural systems such as biological, chemical and physical as well as being used for models of computation. This dissertation grew from an idea to find a new and more applicable metric for a probabilistic classification theory of linear cellular automata developed in the mid 1980's by Robert Gilman (see ref. [G1]). In doing so, it was found that this new metric was actually a generalization of the metric used in the classification results of Gilman. The only prerequisites the reader needs for this work are introductory point set topology and real analysis. It is hoped that, through many examples, thorough explanations and illustrations, the presentation given here will simplify the dynamical theory of cellular automata.

This dissertation is organized in the following manner. Chapter I is an introductory chapter that deals with the development of cellular automata and a brief discussion of its dynamics. Results from product measure theory, ergodic theory and functional analysis are then discussed in chapter II. The space and the new metric are defined in chapter III. Chapter IV shows that this new metric is actually a generalization of the metric previously used in the probabilistic classification of [G1]. Chapter V displays the connection between the periodicity (reoccurrence) of a sequence and equicontinuity under the new metric. Chapter VI contains the heart of the dissertation and shows that similar classification results, as those obtained in [G1], also hold for this new metric. Chapter VII gives an application of these results to digital signal processing. The bibliography lists many significant texts and journal articles on cellular automata and the analytical methods used in proving the results in this dissertation.

Definitions, lemmas, theorems and examples are encountered throughout. The definitions and examples are numbered independently and in succession with respect to the chapter number. The lemmas and theorems follow each other and are numbered in succession with respect to the number of the chapter. Corollaries are numbered after the supporting theorem. For example, the first corollary that is a result of theorem 2.1 will be numbered corollary 2.1.1, regardless of the chapter in which it occurs. The end of a proof for a lemma or theorem is clearly denoted by a ▲ or a █ respectively. Cellular automata functions are defined on a space of functions. Keeping this in mind, I have frequently used the letter f to represent a point (not an automaton) of the space. The new metric used here was originally devised by Kolmogorov. I have therefore used three letters of the Russian (Cyrillic) alphabet А, Б and Г to represent the three different classes of linear automata.

Contents

Chapter I
Introduction.. 1

Chapter II
Preliminaries From Ergodic Theory and Functional Analysis 24

Chapter III
Definition of Cellular Automata and The Metric........... 34

Chapter IV
Two Metrics.. 43

Chapter V
Periodicity and Equicontinuity........................... 53

Chapter VI
Classes.. 62

Chapter VII
Applications To Digital Signal Processing................ 96

Bibliography...123

Symbols and Notation

The following list of symbols and notations complement the usual set theoretic notations and are used throughout this dissertation:

iff	if and only if
$\mathbb{Z}$	the set of integers
$\mathbb{N}$	the set of positive integers
$\mathbb{N}_0$	the set of positive integers union 0
$\circ$	function composition
$\in$	an element of
$\notin$	not an element of
$\emptyset$	the empty set
$\lvert\ \rvert$	absolute value
$\ni$	such that
$\exists$	there exists
$\forall$	for all
Π	product
$\Rightarrow$	implies
$\Leftrightarrow$	equivalent
▲	signifies the end of proof for a lemma
█	signifies the end of proof for a theorem
card(X)	denotes the cardinality of a set X
$C_\varepsilon(f)$	$= \{y \mid d(f,y) \leq \varepsilon, \text{ for } \varepsilon > 0\}$

$g^i(f) \quad = \quad \underbrace{g \circ g \circ g \circ \ldots \circ g}_{i \text{ compositions}}(f)$

$B_\varepsilon(f) \quad = \quad \{y \mid d(g^i(f), g^i(y)) \leq \varepsilon,\ \forall i \in N_0 \text{ and } \varepsilon > 0\}$

$O_g^+(f) \quad =$ The forward orbit of f under g = $\{g^i(f) \mid i \in N_0\}$

$O_g^-(f) \quad =$ The backward orbit of f under g = $\{g^{-i}(f) \mid i \in N_0\}$

$O_g(f) \quad =$ The full orbit of f under g = $\{g^i(f) \mid i \in Z\}$

When no confusion arises the previous sets will be denoted by:

$O^+(f)$, $O^-(f)$ and $O(f)$, respectively.

$A^o \quad =$ The interior of A.

$\overline{A} \quad =$ The closure of A.

$\omega(f) \quad =$ The set of contact (adherence) points of $O_g^+(f)$.

$T_\sigma \quad = \quad \{y \in Y \mid \overline{O_\sigma^+(y)} = Y\} \cap \{y \in Y \mid \overline{O_\sigma^-(y)} = Y\}$

$\lambda_\wedge \quad = \quad \min_{j \in S}\{\lambda(j)\}$

$\lambda_\vee \quad = \quad \max_{j \in S}\{\lambda(j)\}$

$c \leq a,b \quad \leftrightarrow \quad c \leq a,\ c \leq b$

To simplify functional notation, a set of parenthesis will be omitted from $G((f(a_1), f(a_2), \ldots, f(a_n)))$ and simply written as $G(f(a_1), f(a_2), \ldots, f(a_n))$.

Chapter I

Introduction

Cellular automata are mathematical models for complex natural systems in which an infinite lattice of finite state machines updates itself (produces a next generation), in parallel, according to a local rule. Cellular automata were initially introduced and studied by John von Neumann. He was mainly interested in modeling self-reproducing systems such as biological systems. Ironically, today von Neumann is associated with the single-CPU computer architecture. He should actually be remembered for his pioneering work in parallel computing via cellular automata. Cellular automata can be defined for any dimension greater than or equal to one. This work will be concerned with the one dimensional case. One dimensional cellular automata is also known as **Linear Cellular Automata**. That is, linear cellular automata has as its base a lattice of dimension one.

Von Neumann sought to design his automata as a universal computing system or **Turing machine**. A **finite state machine** includes a finite set of states (with a designated starting state), an input alphabet, and a transition function that assigns a next state to every state. A Turing machine is composed of everything in a finite state machine together with a bi-infinite tape. Turing machines have read and write capabilities on the tape and can move forward and backward along this tape. They can model all the computations that can be performed on a computer. The differences between cellular automata and models such as Turing machines can be easily identified. Unlike Turing machines, cellular automata have no pre-determined starting place, the computation proceeds in parallel across the whole lattice. Also, cellular automata do not have a terminal or halting state.

Cellular automata operations occur in discrete time with each step in time being a generation or **forward iteration.** It is assumed that changes in each site ("cell") occur synchronously. The state of a cellular automaton is completely determined by the values of the elements at each site. The particular action that occurs depends on a **local rule** or local function. The local rule depends on the adjacent processing elements, in the sites, to which a particular processing site is connected. These adjacent sites form a **neighborhood** of that processing site. The radius of this neighborhood will be referred to as the **range** of the local rule. For example a local rule, in the one dimensional case, may be defined as follows in figure 1:

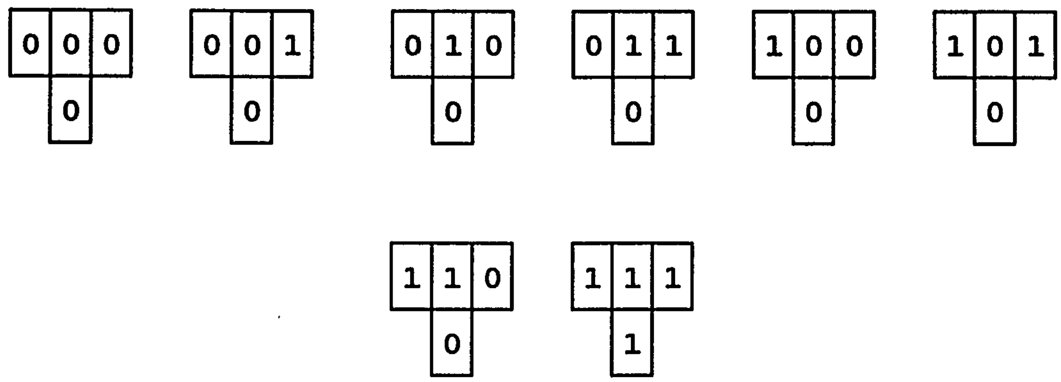


Figure 1

In this rule the range is 1 and three cells (along with their entries) form a neighborhood whereby a new cell with its entry are formed in the next generation. As seen, all configuration elements yield a zero element except the configuration of three ones. When given an initial configuration defined on an infinite lattice the rule operates synchronously on all 3-cell neighborhoods to form the next generation.

As mentioned above, cellular automata can be defined for dimensions other than one (that is other than the linear case). In the two-dimensional case, a neighborhood of a cell (call this cell "Y") is defined as a deleted neighborhood. That is, it is only the cells surrounding cell Y and not including Y. An example of a two-dimensional rule is John Conway's "Game of Life" (see [GA]). Conway was interested in finding a set of rules that would lead to bounded growth with eventual stability. He chose the following rules on a two-dimensional lattice:

1.) A 0-element (a blank) becomes a 1-element if its 2-dimensional neighborhood contains exactly three 1-elements.

2.) If a 1-element has in its neighborhood either two or three 1-elements, then it remains a 1-element.

3.) Otherwise the cell is a 0-element.

Using Conway's rules a number of stable, non-oscillatory, patterns were discovered. It was also found that many input patterns grew but finally stabilized into an overall pattern that consisted of collections of these stable, non-oscillatory initial patterns. It was also discovered that there exist a number of stable oscillatory patterns that oscillate with a period of two or four cycles. Lastly, there were stable oscillatory patterns which translated, in one direction, as they oscillated. The following figures use Conway's set of rules and demonstrate some oscillatory patterns.

The following initial pattern evolves to an oscillatory pattern, note that the pattern achieved on the fifth generation oscillates with a period of two:

			1			
		1	1	1		
			1			

Initial

		1	1	1		
		1		1		
		1	1	1		

1st-generation

			1			
		1		1		
	1				1	
		1		1		
			1			

2nd-generation

		1	1	1		
	1				1	
	1				1	
	1				1	
		1	1	1		

3rd-generation

			1			
		1	1	1		
	1		1		1	
1	1	1		1	1	1
	1		1		1	
		1	1	1		
			1			

4th-generation

			1	1	1			
	1						1	
	1						1	
	1						1	
			1	1	1			

5th-generation

				1				
				1				
				1				
1	1	1				1	1	1
				1				
				1				
				1				

6th-generation

			1	1	1			
	1						1	
	1						1	
	1						1	
			1	1	1			

7th-generation

				1				
				1				
				1				
1	1	1				1	1	1
				1				
				1				
				1				

8th-generation

Figure 2

The following example, known as a "glider", is a pattern which translates unidirectionally (up to the left) as it oscillates:

	1	1	1			
			1			
		1				

Initial

		1				
		1	1			
	1		1			

1st-generation

		1	1			
	1		1			
			1			

2nd-generation

		1	1			
			1	1		
		1				

3rd-generation

		1	1	1		
				1		
			1			

4th-generation

			1			
			1	1		
		1		1		

5th-generation

Figure 3

In the linear case, configurations of cellular automata are specified by bi-infinite sequences of site values. All the generations evolve unidirectional with each time step, that is in the "downward direction". Thus it is more accessible to study the forward iterations of these automata.

In the early 1980's Stephen Wolfram studied the iterations of linear cellular automata rules on a computer. In [W2] a "binary code scheme" is developed to refer to the different automata rules. There it was observed, after extensive computer simulation, that an automaton may behave differently with different initial configurations (different sites on and off). Wolfram further went on to observe that if the initial configuration is chosen at random, then the probability is high that the behavior of the automaton will lie in one of four classes (see ref. [W1] and [W2]). The patterns obtained, through iteration, with different initial configurations are seen to differ in their details; however they do display the same characteristic features. It is important to note that there may exist exceptional initial configurations that give rise to different behavior. However they exist with low or zero probability. Figures 4 and 4A display a pattern generated by totalistic code 36 (totalistic means that the new value in the next generation depends only on the total of all sites in the neighborhood). Figure 4 begins with a random initial configuration. Figure 4B begins with an initial configuration of all ones. A random initial configuration of zeros and ones will eventually, under iteration, be all zeros. However a sequence of all ones, as seen in figure 4B, will remain a sequence of all ones under repeated iteration of rule 36. These observations will later prove to be the foundation in developing a rigorous probabilistic classification of linear automata.

The first three classes correspond, heuristically, to the types of attractors of dynamical systems. Those being: (i) fixed point, (ii) periodic, and (iii) strange attractors. The fourth class corresponds to well-organized systems such as computers and are conjected to be capable of universal computation, that is their evolution can implement any finite algorithm (see [W1]). The four (Wolfram) classes of linear cellular automata are thus listed:

<u>Class</u>	<u>Description</u>
(1)	Evolution leads to a homogeneous state.
(2)	Evolution leads to a set of separated simple stable or periodic structures.
(3)	Evolution leads to a chaotic aperiodic pattern.
(4)	Evolution leads to complex localized structures and sometimes these structures are long-lived.

Class 1 cellular automata evolve after a finite number of generations from "almost all" initial configurations to a unique homogeneous state, where all sites have the same value. Their evolution completely obliterates any information about the initial configuration. That is to say after just one generation it may be impossible to reconstruct the exact initial configuration. In figure 4A note that for a random configuration of zeros and ones, all sites evolve to 0.

Class 2 cellular automata generate separated simple structures from initial site value configurations. Changes in the site values in the initial configuration almost always will affect final site values. The simple structures generated by class 2 cellular automata are either stable or are periodic of usually small periods. Some of these rules yield periodic persistent structures. For example, in code 24, sequences of 111 and 1111, surrounded by three zeros on each side, are seen to be persistent. See figure 5 for an example of the pattern generated by the iterations of code 24.

Class 3 cellular automata generate aperiodic (or "chaotic") patterns upon iteration. However automata in this class may exhibit some regularity. As shown in figures 6A and 6B on the following pages, large triangular "clearings" in which all sites have the same value (typically zero) appear. The regular patterns obtained with some class 3 rules appear to be self-similar fractals. That is their form is the same when viewed at different magnifications. Figure 6A displays a pattern generated by a class 3 cellular automaton, namely totalistic code 12, starting from a random initial configuration.

Evolution of class 4 automata yields the most complex structures. In most cases, starting at a random initial configuration, all sites seem to attain a zero value after a finite number of generations. However, in some cases stable or periodic structures which persist for an infinite number of iterations are formed. Also, in some cases, propagating structures are formed. Totalistic rules 20 and 52 are typical class 4 automata rules and are displayed in figures 7A, 7B and 8. In figure 7B note the persisting structures that can occur in rule 20.

In [W4] a formal language theory classification approach is discussed and provides a more complete characterization of the classes and their complexity. The four classes of linear automata generate distinctive patterns from finite initial configurations by their generations (forward iterations). The four classes are thus defined:

(1) Pattern disappears, under forward iterations, with time.

(2) Pattern evolves to a fixed finite size.

(3) Pattern grows indefinitely at a fixed rate.

(4) Pattern grows and contracts with its generations.

These classes are also characterized by the effects of small changes in the initial configurations:

(1) No change in final state.

(2) Changes only in a region of finite size.

(3) Changes over a region of unbounded increasing size.

(4) Irregular changes.

Thus in classes 1 and 2, information associated with an initial configuration propagates only a finite distance. However, in classes 3 and 4 it propagates an infinite distance.

The figures on the following pages illustrate Wolfram classes 1, 2, 3 and 4, respectively, with various typical rules.

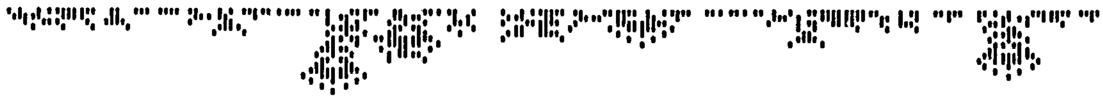


Figure 4A

Wolfram Class 1 Cellular Automata
Totalistic Code 36
Generated From a Random Initial Configuration

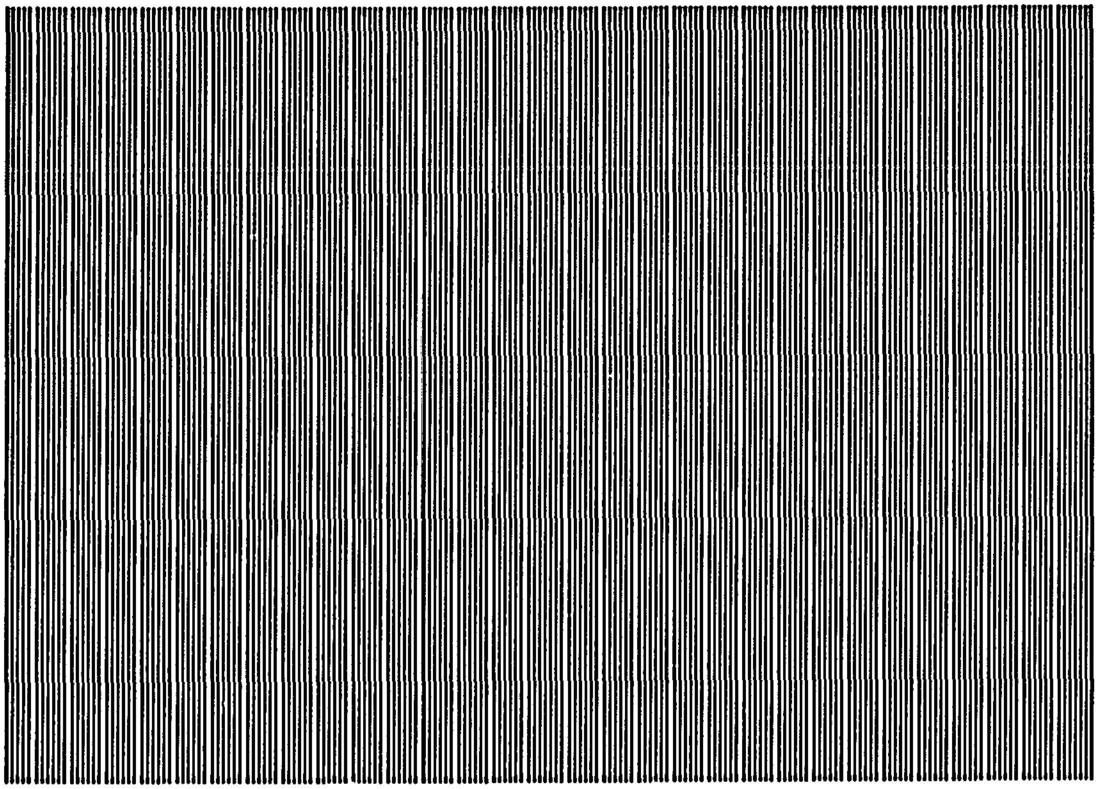


Figure 4B

Wolfram Class 1 Cellular Automata
Totalistic Code 36
Generated From an Initial Configuration of
All Sites With Value 1

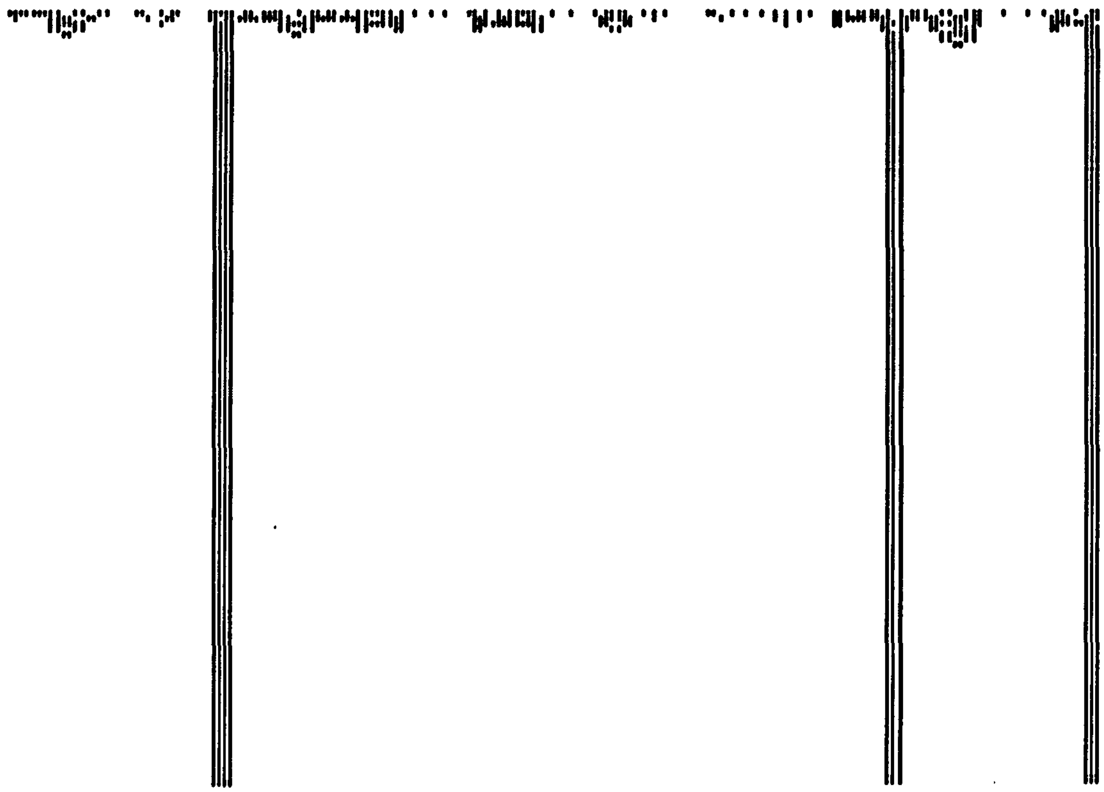


Figure 5

Wolfram Class 2 Cellular Automata
Totalistic Code 24
Generated From a Random Initial Configuration

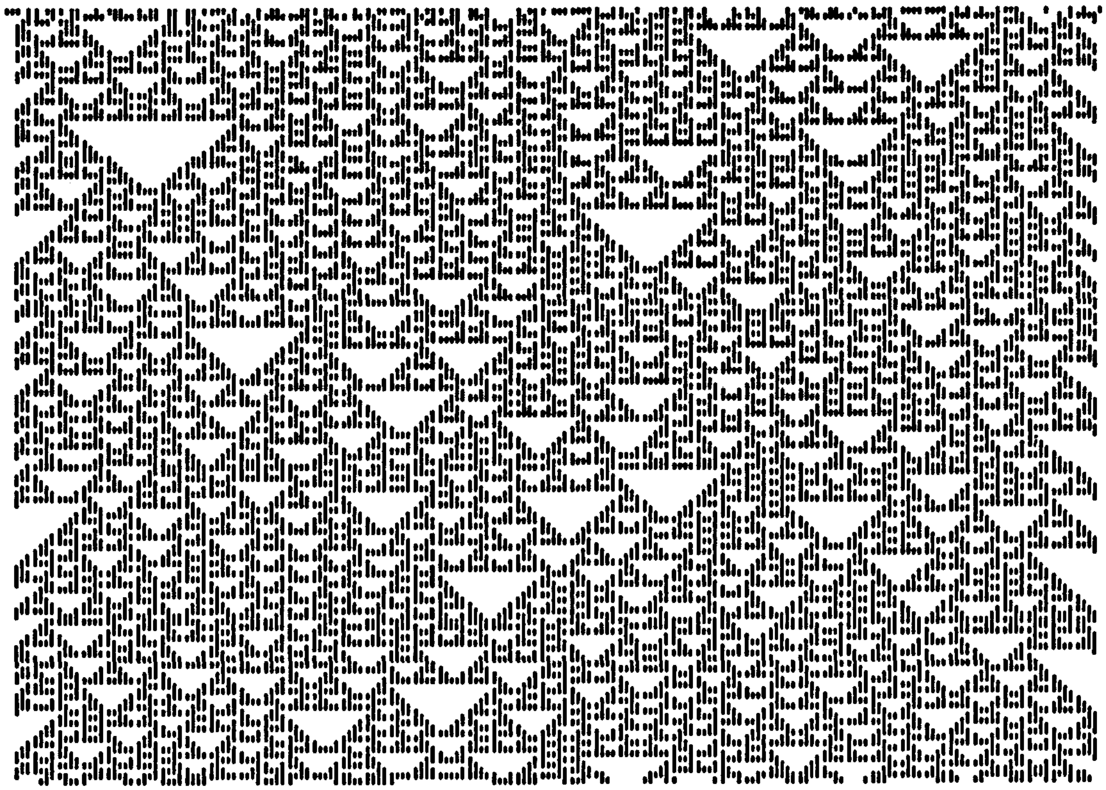


Figure 6A

Wolfram Class 3 Cellular Automata
Totalistic Code 12
Starting From a Random Initial Configuration

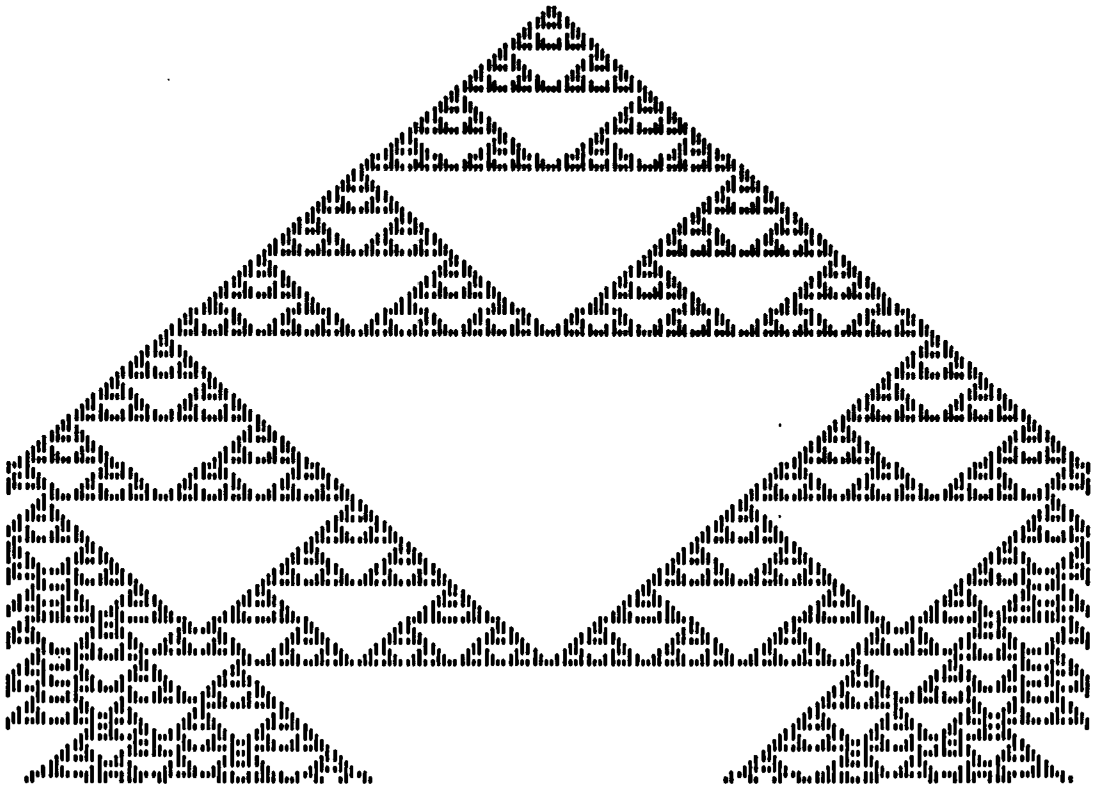


Figure 6B

Wolfram Class 3 Cellular Automata
Totalistic Code 12
Starting From a Two Point Initial Configuration

Figure 7A

Wolfram Class 4 Cellular Automata
Totalistic Code 20
Starting From a Random Initial Configuration

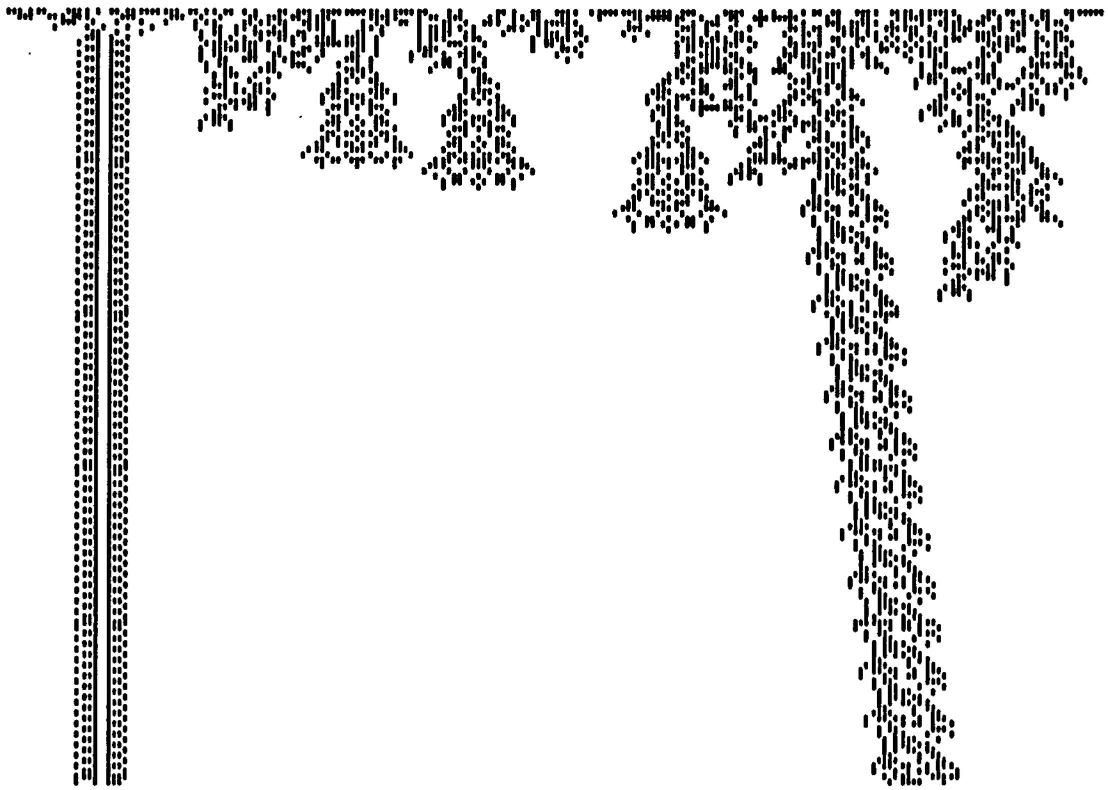


Figure 7B

Wolfram Class 4 Cellular Automata
Totalistic Code 20 With Persisting Structures
Starting From a Random Initial Configuration

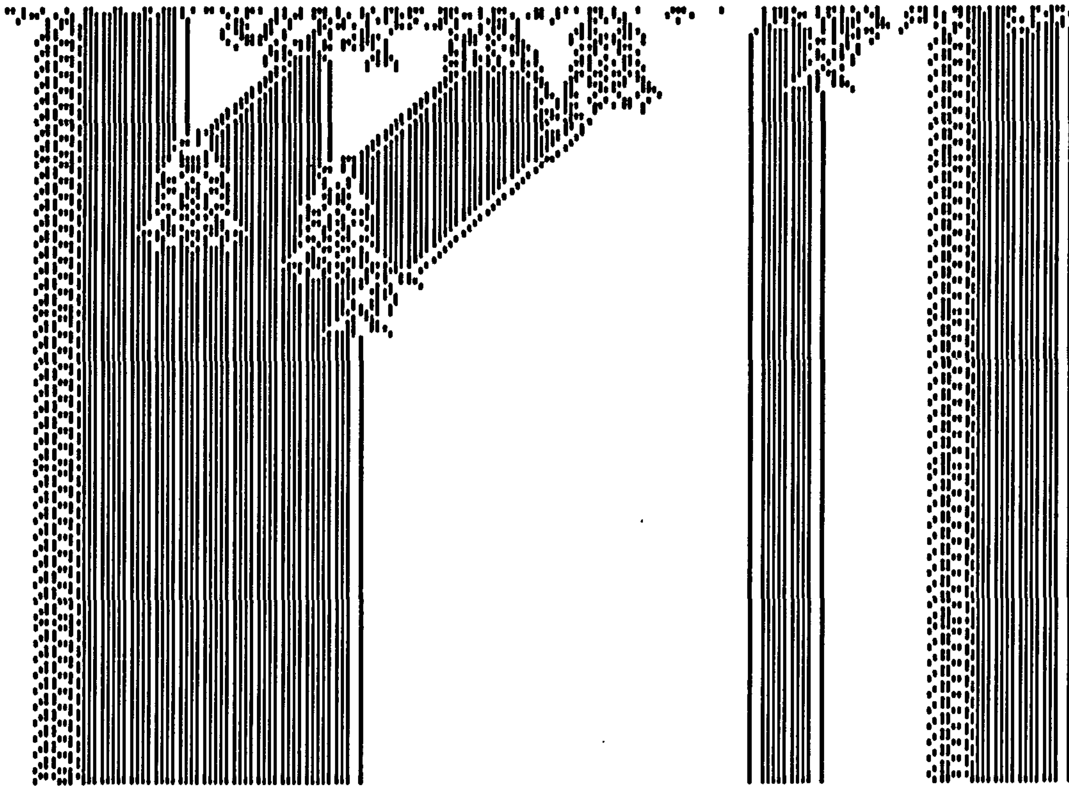


Figure 8

Wolfram Class 4 Cellular Automata
Totalistic Code 52
Generated From a Random Initial Configuration

The first attempt to classify linear cellular automata rigorously, with respect to their iterative ("dynamical") behavior, was done in 1987 by Gilman in **[G1]**. The classification used there is both topological and measure theoretic. A crucial factor in the construction of the classification is the use of the distance between sequences or "closeness". This calls for the use of a **metric**. The classification is determined by the probability of finding a sequence whose generations (forward iterations) will stay close to the generations of a given initial sequence. As a result, linear cellular automata are divided into three classes, *A*, *B* and *C*. Automata in class *A* are equicontinuous on a set of measure one. For a class *A* automaton, only a finite amount of information of that sequence is needed to determine the forward generations completely. Automata in class *B* satisfy a stochastic analogue of equicontinuity. These automata have the property that the probability is positive that a sequence can be found whose generations will remain close to the generations of an initial sequence. Automata in class *C* satisfy a stochastic analogue of expansiveness. These automata have the property that the probability is 0 that a sequence can be found whose generations will stay close to the generations of an initial sequence. Hence, for a class *C* automata, the probability is certain that a sequence can be found whose forward generations will diverge from the forward generations of a given sequence, even though the two sequences may initially be close together. Class *A* automata are independent of the choice of the probability measure. However classes *B* and *C* will depend upon the choice probability measure.

In **[G1]**, the metric used distinguishes distances between sequences by considering the first site where they differ, without respect to the type of value at each site. Hence if two sequences (configurations) differ in the $|n|^{th}$ site, the distance between them is the same as two completely different sequences that differ in the $|n|^{th}$ site.

In this dissertation a more general and applicable metric, originally devised by Kolmogorov, is used in proving an extension of the classification results given in **[G1]**. This new metric uses a real valued weight function on each of the distinct elements in a sequence. Therefore, by utilizing this weight function, the distance between two sequences is dependent upon the elements in the sequence. Thus the distance between two non-equal sequences which differ at the $|n|^{th}$ site will depend upon the common elements where they agree. The use of a weight function has many beneficial and interesting ramifications. For instance, each possible value at a site will have a probability of occurrence and that probability attached to each value will be in a one-to-one correspondence with the weight function. Also, with this new metric, applications will be seen in such areas as string matching and data compression. For instance, in sending, searching, or compressing a line of text, an *a, e, i, o* or *u* may want to be given a different weight than a consonant to stress its importance. Another important consequence of the new metric is that it contains the metric used in [G1] as a particular case. This adds to the versatility of the results obtained.

Chapter II

Preliminaries From Ergodic Theory and Functional Analysis

Notation: N is the set of positive integers, and $N_0 = N \cup \{0\}$. Intervals of integers will be represented by $[a,b] = \{x| \ a \le x \le b\}$, $[a,\infty)$ and $(-\infty,b]$. Intervals of integers, where necessary, will also be denoted with subscripts, for example $[-m_i,m_i]$, to distinguish different lengths.

Definition 2.1: Let $\mathbf{Y}$ be a set. A collection $\mathfrak{C}$ of subsets of $\mathbf{Y}$ is called a **semi-algebra** if the following three conditions hold: (i) $\emptyset \in \mathfrak{C}$; (ii) if A and B $\in \mathfrak{C}$ then $A \cap B \in \mathfrak{C}$; (iii) if $A \in \mathfrak{C}$, then $\mathbf{Y} - A = \cup^n_{i=1} D_i$ where each $D_i \in \mathfrak{C}$ and $D_i \cap D_j = \emptyset$ for $i \neq j$.

Definition 2.2: Let $\mathbf{Y}$ be a set. A **σ-algebra** of subsets of $\mathbf{Y}$ is a collection $\mathfrak{B}$ of subsets of $\mathbf{Y}$ such that the following three conditions hold: (i) $\mathbf{Y} \in \mathfrak{B}$; (ii) if $A \in \mathfrak{B}$ then $\mathbf{Y} - A \in \mathfrak{B}$; (iii) if $A_n \in \mathfrak{B}$ for $n \ge 1$ then $\cup^\infty_{n=1} A_n \in \mathfrak{B}$.

Definition 2.3: Let $\mathbf{X}$ be a metric space, the σ-algebra generated by the open sets is called the σ-algebra of **Borel subsets**.

Definition 2.4: Let **Y** be a set. An **algebra** $\mathscr{E}$ is a collection of subsets of **Y** such that the following three conditions hold: (i) $\emptyset \in \mathscr{E}$; (ii) if $A,B \in \mathscr{E}$, then $A \cap B \in \mathscr{E}$; (iii) if $A \in \mathscr{E}$, then $\mathbf{Y} - A \in \mathscr{E}$.

Note: Every algebra is a semi-algebra and every σ-algebra is an algebra.

Definition 2.5: A finite measure space is a triple $(\mathbf{X},\mathfrak{B},\mu)$, where μ is a finite measure on the space $(\mathbf{X},\mathfrak{B})$, if $\mu(\mathbf{X}) = 1$ then $(\mathbf{X},\mathfrak{B},\mu)$ is called a **probability space** and μ is called a **probability measure** on $(\mathbf{X},\mathfrak{B})$.

Let **S** be a finite alphabet and let $\mathbf{X} = \prod_{i=-\infty}^{\infty}\mathbf{S}$ (the direct product of **S** with itself). A point of **X** has the following form:

$$\ldots, x_{-i}, \ldots, x_{-1}, x_0, x_1, \ldots, x_i, \ldots \quad \text{with each } x_i \in \mathbf{S}$$

Let **S** be a finite alphabet and ν a probability measure on 2^S, the collection of all subsets of **S**, then $(\mathbf{S},2^S,\nu)$ denotes the probability space.

Example 2.1: Let $\mathbf{S} = \{0,1\}$, then $2^S = \{\emptyset,\{0\},\{1\},\{0,1\}\}$.

The measure ν is given by the probability vector $(p_0, p_1, \ldots, p_{n-1})$, which means for each i the point s_i has measure p_i (i.e. $\nu(\{s_i\}) = p_i$) and

$$\text{For each } i,\ p_i > 0 \text{ with } \sum_{i=0}^{n-1} p_i = 1$$

Let $A_j \in 2^S$ for $h \le j \le l$. The set

$$\prod_{i=-\infty}^{h-1} S \times \prod_{j=h}^{l} A_j \times \prod_{i=l+1}^{\infty} S = \{(x_i)_{-\infty}^{\infty} \in X \mid x_j \in A_j \text{ for } h \le j \le l\}$$

is called a **measurable rectangle.** The collection of all measurable rectangle subsets of X forms a semi-algebra $\mathfrak{C}$ that generates a σ-algebra $\mathfrak{B}$.

Let $(X, \mathfrak{B}, \mu) = \prod_{-\infty}^{\infty}(S, 2^S, \nu)$ the countable bi-infinite product of the space $(S, 2^S, \nu)$ with itself. Take elementary rectangles where each A_j is taken to be a point of the finite alphabet S. In this case a rectangle has the following form:

$$\{(x_i)_{-\infty}^{\infty} \mid x_j = s_j \text{ for } h \le j \le l\}$$

This set can be more easily denoted by:

$$({}_h[s_h, s_{h+1}, \ldots, s_{l-1}, s_l]_l)$$

and is called a **block** with end points h and l. Note that h and l are not necessarly both positive or negative, nor necessarly symmetric about the 0^{th} place. Blocks have measure:

$$\mu({}_h[s_h, s_{h+1}, \ldots, s_{l-1}, s_l]_l) = \prod_{i=h}^{l} p_i$$

This measure μ is the product measure defined on the product space $(\mathbf{X}, \mathfrak{B}, \mu)$. The collection of measurable rectangles forms an open set basis for the topology on **X**. Hence a non-empty open set will have positive measure. The space **X** is the space of all bi-infinite sequences with elements from **S**. The notation $\mathbf{S}^{\mathbf{Z}}$ will also be used to denote the space of all bi-infinite sequences with elements from the finite alphabet **S**.

<u>Definition 2.6</u>: Given a finite alphabet **S** of size s > 1, let $\mathbf{X} = \mathbf{S}^{\mathbf{Z}}$. The map σ: **X** --> **X**, defined by $\sigma(\{x_i\}) = \{x_{i+1}\}$, is called the **left two-sided shift** ("two-sided" refers to the fact that the sequences (elements of **X**) are bi-infinite i.e. the infinite product ranges over all the integers **Z**, hence the motivation for the $\mathbf{S}^{\mathbf{Z}}$ notation). Similarly, $\sigma(\{x_i\}) = \{x_{i-1}\}$ is called the **right two-sided shift.** Notice that the two-sided (right or left) shift is a homeomorphism from the space **X**, defined above, to itself.

For the following definitions and theorems $(Y_1, \mathfrak{B}_1, \mu_1)$ and $(Y_2, \mathfrak{B}_2, \mu_2)$ are probability spaces where $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are σ-algebras of subsets of the spaces Y_1 and Y_2 with measures μ_1 and μ_2, respectively.

<u>Definition 2.7</u>: A transformation $T: Y_1 \dashrightarrow Y_2$ is said to be **measure preserving** if T is measurable and $\mu_1(T^{-1}B_2) = \mu_2(B_2)$ $\forall B_2 \in \mathfrak{B}_2$.

<u>Definition 2.8</u>: A transformation $T: Y_1 \dashrightarrow Y_2$ is said to be **invertible measure preserving** if T is measure preserving, bijective, and T^{-1} is also measure preserving.

It is often difficult to check whether a transformation is measure preserving since explicit knowledge of the elements of the σ-algebra $\mathfrak{B}$ is usually needed. This is not the case in practice. For example if the space $\mathbf{Y}$ equals the interval [0,1] on the real line, $\mathfrak{C}$ may be the semi-algebra of all subintervals of $\mathbf{Y}$. For the case mentioned above, when $\mathbf{Y}$ is a direct product space, the semi-algebra $\mathfrak{C}$ may be the collection of all measurable rectangles. The following theorem gives this type of insight into when a transformation is measure preserving:

<u>Theorem 2.1</u>: Suppose $(Y_1, \mathfrak{B}_1, \mu_1)$ and $(Y_2, \mathfrak{B}_2, \mu_2)$ are probability spaces and T: $Y_1 \dashrightarrow Y_2$ is a transformation. Let $\mathfrak{C}_1$ and $\mathfrak{C}_2$ be semi-algebras which generate $\mathfrak{B}_1$ and $\mathfrak{B}_2$, respectively. If for each $A_2 \in \mathfrak{C}_2$ there is $T^{-1}(A_2) \in \mathfrak{C}_1$ and $\mu_1(T^{-1}(A_2)) = \mu_2(A_2)$ then T is measure preserving.

<u>Proof</u>: [see ref. W page 20]

<u>Lemma 2.2</u>: The full two-sided shift map, defined by $\sigma(\{x_i\}) = \{x_{i+1}\}$, is invertible measure preserving.

<u>Proof</u>: Let $n \geq 2$ be a fixed integer and choose $(p_0, p_1, \ldots, p_{n-1})$ to be a probabilty vector with non-zero entries, that is $p_i > 0$ for each i and $\Sigma_{i=0}^{n-1} p_i = 1$). Let $(\mathcal{S}, 2^S, \nu)$ denote the measure space where $\mathcal{S} = \{s_0, s_1, \ldots, s_{n-1}\}$ and the point s_i has measure equal to p_i. If $\mathfrak{C}$ is the semi-algebra of all measurable rectangles then $\mu(\sigma^{-1}A) = \mu(A)$ and $\mu(\sigma A) = \mu(A)$ $\forall$ $A \in \mathfrak{C}$. The invariance condition of the measure is written in the form:

$$\mu(\{x: x_{i_1} \in A_1, \ldots, x_{i_r} \in A_r\}) = \mu(\{x: x_{i_1+k} \in A_1, \ldots, x_{i_r+k} \in A_r\}), \quad -\infty < k < \infty$$

and the A_i's are each taken to be a single point in $\mathcal{S} = \{s_0, s_1, \ldots, s_{n-1}\}$.

Hence by theorem 2.1, σ is measure preserving.

Definition 2.9: Let $(X,\mathcal{B},\mu)$ be a probability space. A measure preserving transformation T of $(X,\mathcal{B},\mu)$ to itself is said to be ergodic iff the only members B of $\mathcal{B}$ with $T^{-1}B = B$ also satisfy the following $\mu(B) = 0$ or $\mu(B) = 1$.

If $T^{-1}B = B$ for $B \in \mathcal{B}$ then $T^{-1}(X-B) = X-B$, and the transformation T can be more easily understood by studying the two simplier transformations $T|_B$ and $T|_{X-B}$ (where $T|_B$ is the restriction of T to B). If $0 < \mu(B) < 1$ then T will be simplified. However, if $\mu(B) = 0$ or $\mu(X-B) = 0$ then T has not been simplified to any significant extent (since dismissing a set of measure 0 is not significant in measure theory). Thus ergodic transformations are those measure preserving transformations, defined on a probability space, which cannot be decomposed into two simplier transformations $T|_B$ and $T|_{X-B}$ with $0 < \mu(B) < 1$.

Theorem 2.3: The full two-sided shift σ (left or right) is ergodic.

Proof: [see ref. W page 32]

Definition 2.10: $g^i(f)$ is used to represent $\underbrace{g\circ g\circ g\circ\cdots\circ g}_{i \text{ compositions}}(f)$

Note, with the previous definition, $g^0(f) = f$

Refer to $g^i(f)$ as the i^{th} iterate of f under the map g.

Similarly, $g^{-i}(f) = \underbrace{g^{-1} \circ g^{-1} \circ g^{-1} \circ \ldots \circ g^{-1}}_{i \text{ compositions}}(f)$

<u>Definition 2.11</u>: If g: **Y** --> **Y**, f is a **periodic point** of g if $g^n(f) = f$ for some $n \geq 1$. The smallest such n is called the **period** of f. A point of period one is called a **fixed point.**

<u>Definition 2.12</u>: The **forward orbit** of f, under g, is the set of points f, g(f), $g^2(f)$,..., $g^i(f)$,.... and will be denoted by $O_g^+(f)$.

<u>Definition 2.13</u>: If g is a homeomorphism of a topological space **X**, define the **backward orbit** of f, under g, as the set of points f, $g^{-1}(f)$, $g^{-2}(f)$,.. ..., $g^{-i}(f)$,.... and denote it by $O_g^-(f)$.

<u>Definition 2.14</u>: For a homeomorphism g of a topological space **X**, the **full orbit** of f, under g, is the set $O_g(f) = \{g^n(f) \mid n \in \mathbf{Z}\}$.

When no confusion arises, the notational function subscript in $O_g(f)$ will be omitted and the orbits of f will just be represented by $O^+(f)$, $O^-(f)$ and $O(f)$.

<u>Theorem 2.4</u>: Let **X** be a compact metric space, $\mathcal{B}(\mathbf{X})$ the σ-algebra of Borel subsets of **X** and let μ be a probability measure on $(\mathbf{X},\mathcal{B}(\mathbf{X}))$ such that $\mu(U) > 0$ for every non-empty open set U. Suppose T: **X** --> **X** is a continuous transformation which preserves μ and is ergodic. Then almost all (in the sense of the measure μ) points of **X** have a dense orbit under T. That is $\mu(\{x \in \mathbf{X} \mid \overline{O_T^+(x)} = \mathbf{X}\}) = 1$.

<u>Proof</u>: See [**W**, Theorem 1.7].

Two final definitions from ergodic theory on topological transitivity and regular measures will now be stated:

<u>Definition 2.15</u>: A measure μ, on a compact topological space **X**, is **regular** if for every $\varepsilon > 0$ and every measurable set E there exists a compact set V and a open set U with $V \subset E \subset U$ such that $\mu(U - V) < \varepsilon$.

<u>Definition 2.16</u>: A homeomorphism g: **X** --> **X** where **X** is a compact topological space is called **topologically transitive** if there exists some $x \in \mathbf{X}$ with $O_g(x) = \{g^n(x) \mid n \in \mathbf{Z}\}$ dense in **X**.

In the following chapters non-archimedean metric spaces will provide the topological setting for numerous results.

<u>Definition 2.17</u>: A metric d is said to be non-archimedean if it satisfies the ultrametric inequality

$$d(x,y) \leq \max\{d(x,z),d(z,y)\}.$$

In these spaces it follows that:

<u>Theorem 2.5</u>: Every point in the open spherical neighborhood of x of radius ε, $S_\varepsilon(x) = \{y \in X \mid d(x,y) \leq \varepsilon\}$, is a center, i.e., if $y \in S_\varepsilon(x)$, then $S_\varepsilon(x) = S_\varepsilon(y)$.

<u>Proof</u>: [see **NBB**, page 6]

<u>Corollary 2.5.1</u>: If $S_\varepsilon(x) \cap S_\alpha(y) \neq \emptyset$ and $\varepsilon \leq \alpha$, then

$$S_\varepsilon(x) \subset S_\alpha(x) = S_\alpha(y)$$

The previous theorem and corollary state that if a metric satisfies the non-archimedean property then given two open balls either one is contained in the other or they intersect trivially.

Chapter III

Definition of Cellular Automata and The Metric

Let **S** be an alphabet of size s such that $2 \le s < \infty$ and let $\mathbf{X} = (\mathbf{S} \cup \{*\})^{\mathbf{Z}}$ be the set of all maps from the integers to $\mathbf{S} \cup \{*\}$. That is, for $f \in \mathbf{X}$, $f: \mathbf{Z} \to \mathbf{S} \cup \{*\}$. The set $\mathbf{S} \cup \{*\}$ is compact with the discrete topology. Therefore, being the infinite product of compact sets, the space **X** is compact.

Definition 3.1: The restriction of a map $f \in \mathbf{X}$ to a non-empty interval $[i,j]$ of **Z**, where $i \le j$, is called a **word.** Words are written as $f[i,j]$ for $-\infty < i \le j < \infty$ or $f[k,\infty)$, $f(-\infty,k]$ and $f(-\infty,\infty)$ for right infinite, left infinite, and bi-infinite intervals of **Z**, respectively. $f(-\infty,\infty)$ is just denoted by f.

The length of a word w is $|w|$. Hence, if $w = f[i,j]$ then $|w| = j - i + 1$. Note that $i = j$ is possible, and it represents the word of length 1.

Example 3.1:

Let $\mathbf{S} = \{a,b\}$ and

$$f = (\ldots * \; * \; * \; a \; a \; b \; a \; \underset{\uparrow 0^{th}}{a} \; b \; b \; b \; a \; a \; a \; a \ldots)$$

letting $w = f[0,0] = a$, it follows that $|w| = 1$.
If $w' = f[-2,2] = b\ a\ a\ b\ b$, then $|w| = 2 - (-2) + 1 = 5$.

Definition 3.2: For any $f, y \in X$ define the binary infimum operator $\wedge$ on X to be:

$$(f \wedge y)(n) = \begin{cases} f(n) \quad \forall n \in Z & \text{if } f = y, \text{ with } f(n) \neq * \\ * & \text{if } f(0) \neq y(0) \\ f(n) \quad \forall n \in [-m,m] & \text{for } m = \text{maximum value} \ni f(n) = y(n) \text{ and } f(n) \neq * \text{ in } [-m,m] \\ * & \text{elsewhere} \end{cases}$$

Thus $f \wedge y$ consists of the largest center stretch of values from S where f and y agree (with no *) and is * valued outside this center.

Example 3.2: using the alphabet $S = \{1,2,3\}$

$$f = (\ldots.1\ 1\ 1\ 1\ 1\ \underset{0^{th}}{2}\ 3\ 3\ 3\ 3\ 3\ldots.)$$

$$y = (\ldots.3\ 3\ 3\ 3\ 1\ \underset{0^{th}}{2}\ 3\ 2\ 2\ 2\ 2\ldots.)$$

then $$f \wedge y = (\ldots.*\ *\ *\ 1\ \underset{0^{th}}{2}\ 3\ *\ *\ *\ldots.)$$

Note: It will be convenient to employ bound vector notation of digital signal processing to represent elements of **X**.

In general, write: $f = (a_0\ a_1\ a_2\ \ldots.\ a_n)^q_p$ whenever

$$f(i) = \begin{cases} q & \text{if } i < p \\ a_0 & \text{if } i = p \\ a_1 & \text{if } i = p+1 \\ a_2 & \text{if } i = p+2 \\ \vdots & \\ a_n & \text{if } i = p+n \\ q & \text{if } i \geq p+n+1 \end{cases}$$

Hence,

$$f \wedge y = (\ldots.*\ *\ *\ 1\ \underset{\uparrow\, 0^{th}}{2}\ 3\ *\ *\ *\ldots.)$$

can be written $f \wedge y = (1\ 2\ 3)_{-1}^{*}$.

It can be easily verified that:

i.) $f \wedge (y \wedge z) = (f \wedge y) \wedge z$

ii.) $f \wedge y = y \wedge f$

iii.) $f \wedge f = f$

iv.) $* \wedge f = *$

and that the binary relation $\leq$, on **X**, defined by $f \leq y$ if and only if $f = f \wedge y$, forms a partial order on **X**. As **X** is closed under the infimum operation $\wedge$ and satisfies (i.), the pair $<\mathbf{X},\wedge>$ is a semigroup. Since (ii.) - (iv.) are also satisfied $<\mathbf{X},\wedge>$ forms a lower (infimum) semilattice with the lower unit *.

Let **Y** = S^Z be a subspace of **X** defined as the set of all maps from the integers to the alphabet **S**, i.e. f: **Z** --> **S**. As the space **X** is compact so is **Y**. The space **Y** can also be considered as the space of all bi-infinite sequences with elements taken from **S**. Define a metric on the space **Y** as follows:

$$d(f,y) = \begin{cases} 1 & \text{if } f(0) \neq y(0) \\ 0 & \text{if } f = y \\ \prod_{i=-m}^{m} \lambda_i & \text{if } (f \wedge y) = \\ & \quad (\ldots * * f(-m)\ldots f(0)\ldots f(m) * * \ldots) \end{cases}$$

where λ is any real-valued function defined on **S** and taking values in the open interval (0,1).

i.e. λ : **S** ----> (0,1) where λ_i is defined by $\lambda_i = \lambda(f(i))$

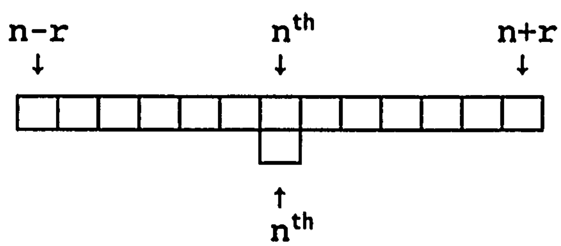
So $0 < \lambda_i < 1$

(**Y**,d) forms a metric space where the metric d is **non-archimedean**, i.e. $d(f,y) \leq \max\{d(f,z),d(z,y)\}$.

The metric just defined will be referred to as The Kolmogorov Metric.

Linear, or one-dimensional, cellular automata are induced by arbitrary (local) maps G: $S^{2r+1} \rightarrow S$ (these local maps are sometimes called **local rules** or **block maps** in the literature) where $r \in N_0$. The value r is called the **range** of the cellular automaton. The automaton map g induced by G is defined by g(f) = z with z(n) = G(f(n-r),...,f(n),...,f(n+r)).

The following diagram illustrates the cells that influence the next generation n^{th} cell in a local map of range r:



Linear cellular automata maps are characterized as being symmetrical about the n^{th} cell in the sense that the next generation n^{th} cell is dependent on r cells to its left, r cells to its right as well as the center cell directly above.

<u>Example 3.3</u>:

i.) The idenity map δ is the automaton of range 0 induced by G(a) = a.

ii.) The right 1-shift (or just shift) map σ is the automaton of range 1 induced by $G(\alpha,\beta,\gamma) = \alpha$.

<u>Definition 3.3</u>: $C_\varepsilon(f) = \{z \mid d(f,z) \leq \varepsilon, \text{ for } \varepsilon > 0\}$, $C_\varepsilon(f)$ is the open ball of radius ε around f.

Note: i.) Since the metric is non-archimedean, given any two disks $C_\varepsilon(f)$, $C_\alpha(y)$, either $C_\varepsilon(f) \cap C_\alpha(y) = \emptyset$ or one contains the other.

ii.) In this topology the C_ε sets are also closed.

For fixed $\varepsilon > 0$ the relation $f \sim y$ if $d(f,y) \leq \varepsilon$ is an equivalence relation with equivalence classes $\{C_\varepsilon(f)\}$. The fact that $\sim$ forms an equivalence relation with equivalence classes $\{C_\varepsilon(f)\}$ follows from the non-archimedean property of the metric d.

Example 3.4: Consider the alphabet $S = \{1,2,3\}$ and let

$$\lambda(1) = .5$$
$$\lambda(2) = .2$$
$$\lambda(3) = .1$$

If $f = (2\ 2\ 3)^1_{-1}$ and $z = (1\ 2\ 2\ 3\ 1)^3_{-2}$ then since $f \wedge z = (1\ 2\ 2\ 3\ 1)^*_{-2}$ it follows that

$$d(f,z) = (.5)(.2)(.2)(.1)(.5) = 0.001$$

Moreover, $z \in C_\varepsilon(f)$ for any $\varepsilon \geq 0.001$

Notice that $y = (1)^1_0$ is not in $C_\varepsilon(f)$ for any $\varepsilon < 1$. Additionally, observe that if $x = (1\ 1\ 2\ 2\ 3\ 1\ 1)^2_{-3}$ then x is also in $C_\varepsilon(f)$ since

$d(x,f) = (.5)(.5)(.2)(.2)(.1)(.5)(.5) = 0.00025$

Another equivalence class is obtained in the same manner using the following definition:

<u>Definition 3.4</u>:

$$B_\varepsilon(f) = \{y |\ d(g^i(f), g^i(y)) \leq \varepsilon,\ \forall i \epsilon N_0 \ \text{ and } \varepsilon > 0\}$$

Notice that $B_\varepsilon(f) \subset C_\varepsilon(f)$. In chapters V and VI it will be shown that under given conditions the opposite inclusion holds.

The $\{B_\varepsilon(f)\}$ are equivalence classes with the relation $f \approx y$ iff $d(g^i(f), g^i(y)) \leq \varepsilon$, $\forall i \in N_0$. Again, the fact that $\approx$ forms an equivalence relation follows from the non-archimedean property of the metric d.

The behavior of f, that is the iterations of f under a linear automaton map g, can be visualized as an array $(a_{i,j})$, with rows called i and columns j, here each entry $a_{i,j} = (g^i(f))(j)$ $\forall i \epsilon N_0$ and $j \epsilon \mathbf{Z}$. Then $B_\varepsilon(f)$ = the set of all y whose behavior agrees with the behavior of f on the infinite **vertical jagged edge strip** defined by the intervals $[-m_i, m_i]$ such that $d(g^i(f), g^i(y)) \leq \varepsilon$ $\forall i \in N_0$. That is, $B_\varepsilon(f)$ is the set of y where $(g^i(f))[-m_i, m_i] = (g^i(y))[-m_i, m_i]$, for each $i \epsilon N_0$ where, for given $\lambda: \mathbf{S} \dashrightarrow (0,1)$, the minimum length of the $[-m_i, m_i]$ intervals is determined by ε and the intervals $[-m_i, m_i]$ themselves are determined by the places where f and y agree.

In general, the terminology that y has behavior $B_\varepsilon(f)$ on the intervals $[-m_i+k, m_i+k]$ if $\sigma^{-k}(y) \in B_\varepsilon(f)$ (σ is the right-shift map) for all $i \in N_0$ and $k \in N_0$ is used. Note that, for a fixed automaton g and fixed ε the shape of the jagged edge strips will depend on the elements y in $B_\varepsilon(f)$ similiar to the "length" of agreement for elements in $C_\varepsilon(f)$ (see example 3.5); however the minimum length of the intervals will remain fixed for each ε. The sets $B_\varepsilon(f)$ are useful to study since many computer experiments (and applications) of linear automata display the central window of the array $(a_{i,j})$. The classification scheme used here will be based on the analysis of the sets $B_\varepsilon(f)$.

<u>Example 3.5</u>: The right 6 - shift σ_6 induced by:

$G: S^{13} \dashrightarrow S, \quad G(a,b,c,d,e,f,g,h,i,j,k,l,m) = a$

defined on $Y = S^Z$, where $S = \{0,1\}$

and suppose: $\lambda(0) = 1/10$ and $\lambda(1) = 9/10$

and choose $\varepsilon = 1/500$

If $y = (0)^0{}_0$ and

$f = (1\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0$
$0\ 0\ 1\ 1\ 0\ 0\ 0)^0{}_{-34}$

The forward orbit of y, under σ_6, always yields $(0)^0{}_0$ and consider the forward orbit of f under σ_6:

```
                                           0th
                                            ↓
f =       ...0 0 0 0 0 0 0 0 1 1|0 0 0 0 0|0 0 0 0 0 0 0..
σ6(f) =   ...0 0 0 0 0 0 0 0 0 0 0|0 0 0|1 1 0 0 0 0 0 0..
σ6^2(f)=  ...0 0 0 1 0|0 0 0 0 0 0 0 0 0 0 0 0 0 0 0|1 1..
σ6^3(f)=  ...0 0 0 1 1 0 0 0 0 1|0 0 0 0 0|0 0 0 0 0 0 0..
σ6^4(f)=  ...0 0 1 0 0 0 0 0 0 1 1|0 0 0|0 1 0 0 1 1 0 0..
σ6^5(f)=        ...0 0 0 0 0 1 0|0 0 0 0 0|1 1 0 0 0 1 0..
σ6^6(f)=        ...0 0 0 0 0 0 0 0|0 0 0|1 0 0 0 0 0 0 0..
σ6^7(f)=        ...0 0|0 0 0 0 0 0 0 0 0 0 0 0 0 0 0|1 0..
```

Since the function $\sigma^7(f)$ has all 0's to the left, the widths of the intervals $[-m_i, m_i]$ will increase, by 12, for each iteration $i \geq 7$.

Then y agrees with f on the infinite vertical jagged edge strip such that $d(\sigma_6^{\,i}(f), \sigma_6^{\,i}(y)) \leq \varepsilon \quad \forall i \in N_0$.

Hence $y \in B_\varepsilon(f)$.

<u>Note</u>: By changing the value of ε to $\varepsilon = 1/2000$ then y does not belong to the class $B_\varepsilon(f)$.

Chapter IV

Two Metrics

In **[G1]** the following metric on the set **Y** is utilized. $d^\wedge(x,y) = 2^{-n}$, where $n = \inf\{|i|: x(i) \neq y(i)\}$ and $d^\wedge(x,y)=0$ if $x = y$. The cap ^ symbol will be used when referring to a structure that uses this metric Thus if $d^\wedge(x,y) = 2^{-n}$ it follows that $x[-n+1,n-1] = y[-n+1,n-1]$ and $x[-n,n] \neq y[-n,n]$. In this reference, the open disks $C_n{}^\wedge(x)$ around the point x are defined by the cylinder set $C(i,j,w) = \{x \in \mathbf{Y}: x[i,j] = w\}$, i.e. $C_n{}^\wedge(x) = C(-n,n,x[-n,n])$ is the open disk of radius 2^{-n} around x. Thus $y \in C_n{}^\wedge(x)$ means that $d^\wedge(x,y) < 2^{-n}$ and so surely $x[-n-k,n+k] = y[-n-k,n+k]$ for some $k \in N_0$. This metric is also non-archimedean and for fixed $n \in N_0$ the relation $x \sim y$ iff $d^\wedge(x,y) < 2^{-n}$ is an equivalence relation with equivalence classes $\{C_n{}^\wedge(x)\}$. Again, continuing in this manner $x \approx y$ is an equivalence relation iff $d^\wedge(g^i(x),g^i(y)) < 2^{-n}$ $\forall i \in N_0$, and equivalence classes $\{B_n{}^\wedge(x)\}$. Here the set $B_n{}^\wedge(x)$ is the set of all $y \ni (g^i(y))[-n,n] = (g^i(x))[-n,n]$ $\forall i \in N_0$. Visualize the array $(a_{i,j})$ with entries $a_{i,j}=(g^i(x))(j)$, then $B_n{}^\wedge(x)$ is the set of all y that agree with x on the infinite straight edge vertical strip under the interval [-n,n]. As an example: suppose one considers the right shift map σ of range 1 induced by the local map $G(a,b,c) = a$.

Consider the class $B_1\hat{}(x)$ for $y = (1)^1{}_0$, if

$$x = (\ldots 1\ 1\ 1\ \underset{\uparrow\,0^{th}}{1}\ 1\ 0\ 0\ 0\ 0\ 0 \ldots)$$

$$\begin{array}{rcl|ccc|l}
x & = & \ldots 1\ 1\ 1\ 1 & 1 & \overset{0^{th}}{\downarrow}1 & 1 & 0\ 0\ 0\ 0\ 0 \ldots \\
\sigma(x) & = & \ldots 1\ 1\ 1\ 1 & 1 & 1 & 1 & 1\ 0\ 0\ 0\ 0 \ldots \\
\sigma^2(x) & = & \ldots 1\ 1\ 1\ 1 & 1 & 1 & 1 & 1\ 1\ 0\ 0\ 0 \ldots \\
\sigma^3(x) & = & \ldots 1\ 1\ 1\ 1 & 1 & 1 & 1 & 1\ 1\ 1\ 0\ 0 \ldots \\
\vdots & & & & \vdots & &
\end{array}$$

Hence $y \in B_1\hat{}(x)$.

From the the following lemmas and observation, it will be shown that the Kolmogorov metric is more general than and for given conditions actually contains the metric used in **[G1]** as a special case. Two important quantities will be used here for the first time. These are $\max_{j\in S}\{\lambda(j)\}$ and $\min_{j\in S}\{\lambda(j)\}$.

<u>Lemma 4.1a</u>: $\forall$ f $\in$ **Y**, given $\varepsilon \geq (\max_{j\in S}\{\lambda(j)\})^{2n+1}$ for $n \in N_0$ then $C_n\hat{}(f) \subset C_\varepsilon(f)$.

<u>Proof</u>: $x \in C_n\hat{}(f) \Rightarrow x[-n-k,n+k] = f[-n-k,n+k]$ for some $k \in N_0$. this implies that

$$d(x,f) = \prod_{i=-n-k}^{n+k} \lambda_i$$

or that $x = f$, in which case $x \in C_\varepsilon(f)$.

Assume

$$d(x,f) = \prod_{i=-n-k}^{n+k} \lambda_i \quad \text{for some } k \geq 0$$

to show that $x \in C_\varepsilon(f)$. Then it must be true that

$$\prod_{i=-n-k}^{n+k} \lambda_i \leq \varepsilon$$

This must hold true for all λ_i. In particular it must hold true for the maximum value, thus it must be shown that

$$\varepsilon \geq \prod_{i=-n-k}^{n+k} (\max_{j \in S} \lambda(j)) = (\max_{j \in S} \lambda(j))^{2(n+k)+1}$$

But from the hypothesis

$$\varepsilon \geq (\max_{j \in S} \{\lambda(j)\})^{2n+1} \geq (\max_{j \in S} \{\lambda(j)\})^{2(n+k)+1}$$

and $C_n\hat{}(f) \subset C_\varepsilon(f)$. ▲

Example of Lemma 4.1a:

Lemma 4.1a says that if $\varepsilon \geq (\max_{j \in S}\{\lambda(j)\})^{2n+1}$ for $n \in N_0$, then $C_n\hat{}(f) \subset C_\varepsilon(f)$. Thus, with the above condition, there may be elements in $C_\varepsilon(f)$ that are not in $C_n\hat{}(f)$, this is demonstrated in the following example:

Using the alphabet $S = \{0,1\}$, for given $\lambda(0) = 1/10$

$\lambda(1) = 1/2$ and

for $n = 1$

choose $\varepsilon \geq (1/2)^{2n+1} = 1/8$ i.e. choose $\varepsilon(\alpha) = 1/8 + \alpha$ $\forall \alpha \geq 0$

$$\text{if } f = (0)_0^0 \text{ then } C_{\varepsilon(\alpha)}(f) = \ldots . * * * \underset{\uparrow\ 0^{th}}{0} * * * \ldots .$$

$$\text{and } C_1\hat{}(f) = \ldots . * * 0 \underset{\uparrow\ 0^{th}}{0} 0 * * \ldots .$$

Hence if $x = (101)_{-1}^{1} = \ldots.111101111\ldots.$

Then $x \in C_{\varepsilon(\alpha)}(f)$ but $x \notin C_1\hat{}(f)$.

As illustrated, for $\varepsilon \geq (\max_{j \in S}\{\lambda(j)\})^{2n+1}$ $\exists$ elements in $C_\varepsilon(f)$ which are not in $C_n\hat{}(f)$.

<u>Lemma 4.1b</u>: For any $f \in Y$, given $\varepsilon \geq (\max_{j \in S}\{\lambda(j)\})^{2n+1}$ for $n \in N_0$, then $B_n\hat{}(f) \subset B_\varepsilon(f)$.

<u>Proof</u>: $x \in B_n\hat{}(f) \rightarrow (g^i(x))[-n-k,n+k] = (g^i(f))[-n-k,n+k]$ for some $k \in N_0$ and $\forall\ i \in N_0$ (note that k is dependent on i). This implies that

$$d(g^p(x), g^p(f)) = \prod_{i=-n-k}^{n+k} \lambda_i \quad \text{for } p = 0,1,2,3,\ldots.$$

or that $g^p(x) = g^p(f)$ for all or some $p \in N_0$, in either case $x \in B_\varepsilon(f)$. Assume

$$d(g^p(x), g^p(f)) = \prod_{i=-n-k}^{n+k} \lambda_i \quad \text{for } p \in N_0$$

to show that $x \in B_\varepsilon(f)$. Then it must be true that

$$\prod_{i=-n-k}^{n+k} \lambda_i \leq \varepsilon$$

This must hold true for all λ_i. In particular it must hold true for the maximum value, thus it must be shown that

$$\varepsilon \geq \prod_{i=-n-k}^{n+k} (\max_{j\in S}\lambda(j)) = (\max_{j\in S}\lambda(j))^{2(n+k)+1}$$

But from the hypothesis $\varepsilon \geq (\max_{j\in No} \lambda(j))^{2n+1} \geq (\max_{j\in No} \lambda(j))^{2(n+k)+1}$ and $B_n\hat{}(f) \subset B_\varepsilon(f)$. ▲

Example of Lemma 4.1b: Suppose the condition on ε is changed. For the alphabet $S = \{0,1\}$ with $\lambda(0) = 1/2$ and $\lambda(1) = 1/10$

use the following rule: G(1,1,1) = 1

otherwise = 0

For $\varepsilon < (\max_{j\in No}\{\lambda(j)\})^{2n+1}$ with $n = 1$, $\varepsilon < (1/2)^3 = 1/8$.

So choose $\varepsilon(\alpha) = 1/8 - \alpha$ for any $\alpha\in\mathbf{R} \ni 0 < \alpha < 1/8$ and let

0^{th} ↓

f	= ...0 0 0 0 0 0	0 0 1	1 1 1 1 1 0 0 0 0...
g(f)	= ...0 0 0 0 0 0	0 0 0	1 1 1 1 0 0 0 0 0...
g(g(f))	= ...0 0 0 0 0 0	0 0 0	0 1 1 0 0 0 0 0 0...

$g^i(f) = (0)_0^0$ for all natural numbers $i > 2$.

0^{th} ↓

If x = ...0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0...

$g^i(x) = (0)_0^0$ for all $i \in N$ then

since $\lambda(g^i(f(j))) = 1/2$ for $i = 1$ and all $j \in [-n,n]$

$d(g(f),g(x)) = 1/8 \nless \varepsilon(\alpha)$

Hence $x \in B_1\hat{}(f)$ but $x \notin B_{\varepsilon(\alpha)}(f)$.

Lemma 4.2a: If $d\hat{}(x,f) = 2^{-(n+1)}$ and $x \in C_\varepsilon(f)$ then $\varepsilon \geq (\min_{j\in S}\{\lambda(j)\})^{2n+1}$ for $n \in N_0$.

Proof: $d\hat{}(x,f) = 2^{-(n+1)}$ implies that $x[-n,n] = f[-n,n]$ and furthermore, that $x[-n-1,n+1] \neq f[-n-1,n+1]$. Therefore

$$d(x,f) = \prod_{i=-n}^{n} \lambda_i$$

but since $x \in C_\varepsilon(f)$ it follows that

$$\varepsilon \geq \prod_{i=-n}^{n} \lambda_i$$

which must hold for $\lambda_i = \lambda(f(i))$, in particular it must hold for the smallest of these values. Thus $\varepsilon \geq (\min_{j\in S} \lambda(j))^{2n+1}$. ▲

Lemma 4.2b: If $d\hat{}(g^p(x),g^p(f)) = 2^{-(n+1)}$ for all $p \in N_0$ and $x \in B_\varepsilon(f)$ then $\varepsilon \geq (\min_{j\in S}\{\lambda(j)\})^{2n+1}$ for $n \in N_0$.

Proof: $d\hat{}(g^p(x),g^p(f)) = 2^{-(n+1)}$ implies that

$(g^p(x))[-n,n] = (g^p(x))f[-n,n]$ for all $p \in N_0$ and furthermore that $(g^p(x))[-n-1,n+1] \neq (g^p(x))f[-n-1,n+1]$ for all $p \in N_0$. Therefore

$$d(g^p(x),g^p(f)) = \prod_{i=-n}^{n} \lambda_i \quad \text{for all } p\in N_0$$

but since $x \in B_\varepsilon(f)$ it follows that

$$\varepsilon \geq \prod_{i=-n}^{n} \lambda_i$$

which must also hold for $\lambda_i = \lambda(f(i))$, in particular the minimum of these values of λ. Thus $\varepsilon \geq (\min_{j \in S} \lambda(j))^{2n+1}$. ▲

<u>Example of lemma 4.2b</u>: Let $S = \{0,1,2\}$ and let g be the automaton of range 2 induced by the following rule:

G(*,*,0,*,*) = 0	G(*,*,1,0,0) = 1	G(*,*,2,0,0) = 1
	G(*,*,1,0,1) = 1	G(*,*,2,0,1) = 1
	G(*,*,1,0,2) = 2	G(*,*,2,0,2) = 2
	G(*,*,1,1,0) = 1	G(*,*,2,1,0) = 1
	G(*,*,1,1,1) = 1	G(*,*,2,1,1) = 1
	G(*,*,1,1,2) = 1	G(*,*,2,1,2) = 1
	G(*,*,1,2,0) = 2	G(*,*,2,2,0) = 2
	G(*,*,1,2,1) = 2	G(*,*,2,2,1) = 2
	G(*,*,1,2,2) = 2	G(*,*,2,2,2) = 2

(Here * represesents any element in S)

```
                   0th
                    ↓
Let f = ....* * |0 0 0|1 0 2 2 1 1 0 1 2 * * *....
                |0 0 0|2 0 2 1 1 1 0
                |0 0 0|2 0 1 1 1
                |0 0 0|1 0 1
                |0 0 0|1
                |     |
                |     |
```

$$\text{and } x = \;\dots\ast\ \ast \left|\begin{array}{ccc} 0 & \overset{0^{th}}{\overset{\downarrow}{0}} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ & & \\ & & \end{array}\right|\begin{array}{llllllllllllll} 0 & 2 & 1 & 1 & 2 & 2 & 0 & 1 & 1 & 2 & \ast & \ast & \ast\dots \\ 0 & 1 & 1 & 2 & 2 & 1 & 0 & 1 & & & & & \\ 0 & 1 & 2 & 2 & 1 & 1 & & & & & & & \\ 0 & 2 & 2 & 1 & & & & & & & & & \\ 0 & 2 & & & & & & & & & & & \\ & & & & & & & & & & & & \\ & & & & & & & & & & & & \end{array}$$

The condition $d\hat{}(g^p(x), g^p(f)) = 2^{-(n+1)}$ $\forall\, p \in N_0$ and for $n \in N_0$ is satisfied.

If the following values of λ are given:

$$\lambda(0) = 0.1$$

$$\lambda(1) = 0.2$$

$$\lambda(2) = 0.5$$

and $\varepsilon = .01$ then $x \in B_\varepsilon(f)$. Hence, by lemma 4.2b, $\varepsilon \geq (\min_{j\in S}\{\lambda(j)\})^{2n+1} = (0.1)^3$ for the case presented here. Which is readily seen to be true since $d(g^p(x), g^p(f)) = (0.01)^3$ $\forall\, p \in N_0$.

<u>Lemma 4.3a</u>: For any $f \in Y$, if $\varepsilon \leq (\min_{j\in S}\{\lambda(j)\})^{2n+1}$ for $n \in N_0$, then $C_\varepsilon(f) \subset C_n\hat{}(f)$.

<u>Proof</u>: If $x \in C_\varepsilon(f)$ and

$$d(x, f) = \prod_{i=-m}^{m} \lambda_i$$

then $m \geq n$.

If not (i.e. m < n) then

$$\varepsilon \geq \prod_{i=-m}^{m} \lambda_i \geq \prod_{i=-n}^{n} \lambda_i \geq (\min_{i\in S}\{\lambda(i)\})^{2n+1}$$

and is a contradiction. Thus x and f agree at least in [-n,n] which implies $d^\wedge(x,f) < 2^{-n}$. Hence it follows that $x \in C_n^\wedge(f)$. ▲

<u>Lemma 4.3b</u>: For any $f \in Y$, if $\varepsilon \leq (\min_{j\in S}\{\lambda(j)\})^{2n+1}$ for $n \in \mathbb{N}_0$, then $B_\varepsilon(f) \subset B_n^\wedge(f)$.

<u>Proof</u>: If $x \in B_\varepsilon(f)$ and

$$d(g^p(x), g^p(f)) = \prod_{i=-m-k}^{m+k} \lambda_i \quad \textit{for each } p \in \mathbb{N}_0 \textit{ and some } k \geq 0$$

then $m + k \geq n$.

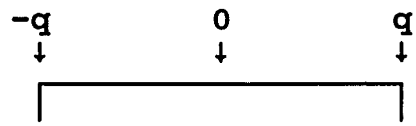
Suppose not (i.e. $m + k < n$), then

$$\varepsilon \geq \prod_{i=-m-k}^{m+k} \lambda_i \geq \prod_{i=-n}^{n} \lambda_i \geq (\min_{i\in S}\{\lambda(i)\})^{2n+1}$$

and is a contradiction. Thus $g^p(x)$ and $g^p(f)$ agree at least in [-n,n] $\forall\, p \in \mathbb{N}_0$, which implies $d(g^p(x), g^p(f)) < 2^{-n}$. Hence $x \in B_n^\wedge(f)$. ▲

<u>Observation</u>: Using the Kolmogorov metric, a minimum bound criteria can be set, for elements of $C_\varepsilon(f)$ and $B_\varepsilon(f)$, on the sizes of the intervals where $(g^i(f))[-m_i,m_i] = (g^i(y))[-m_i,m_i]$, for each $i \in N_0$. Indeed, if $\varepsilon \le (\min_{j \in S}\{\lambda(j)\})^{2q+1}$ is chosen then $\inf_{i \in No}(m_i) \ge q$ for $q \in N_0$.

As illustrated:



Hence, by choosing ε properly, a lower bound can be placed on the size of the intervals $[-m_i,m_i]$.

Now it will be shown that the metric used in **[G1]** is a special case of the Kolmogorov metric.

<u>Lemma 4.4</u>: If $\lambda(j) = 1/2 \quad \forall\, j \in S$ and $\varepsilon = 1/2^{2n+1}$ then the open disks around any point f and the $B_\varepsilon(f)$ and $B_n\hat{}(f)$ classes for both metrics coincide.

<u>Proof</u>: For $x \in Y \ni x \in C_n\hat{}(f) \Rightarrow x[-n,n] = f[-n,n]$ and for $x \in C_\varepsilon(f) \Rightarrow x[-n,n] = f[-n,n]$. The second implication follows from the previous observation, that is for

$\varepsilon \le (\min_{j \in S}\{\lambda(j)\})^{2n+1}$ then $\inf_{i \in No}(m_i) \ge n$.

For $x \in B_n\hat{}(f) \Rightarrow (g^i(x))[-n,n] = (g^i(f))[-n,n] \quad \forall i \in N_0$ and $\forall\, f \in Y$ and $x \in B_\varepsilon(f) \Rightarrow (g^i(x))[-n,n] = (g^i(f))[-n,n] \; \forall i \in N_0$ and $\forall f \in Y$. Again, the latter implication follows from the previous observation. ▲

Chapter V

Periodicity and Equicontinuity

Definition 5.1: Let (Y,d) be the metric space previously defined. Let $\mathcal{F}$ be a subset of the collection of elements of the space Y^Y which are automata (recall that Y^Y consists of all functions $g: Y \to Y$). If $y_0 \in Y$, the family of functions $\mathcal{F}$ is **equicontinuous** at y_0 if given $\varepsilon > 0$, there exists a δ-neighborhood C_δ of y_0 such that for all $y \in C_\delta$ and all $g \in \mathcal{F}$, $d(g(y),g(y_0)) \leq \varepsilon$.

The subset $\mathcal{F}$, of functions from Y^Y, that is of study here is the set of all forward iterates of an automaton g, the set $\{g^i \mid i \in N_0\}$. Therefore, instead of saying the set of iterates g^i $\forall i \in N_0$ is equicontinuous at f, it is simply said that g is equicontinuous at f.

Definition 5.2: A non-empty word $y[-q,q]$ is **ultimately periodic** if there exists $k \geq 0$ and $q \geq 0$ such that $(g^{n+i}(y))[-q,q] = (g^i(y))[-q,q]$ for some $n \geq 1$ and all $i \geq k$. The smallest n such that $(g^{n+i}(y))[-q,q] = (g^i(y))[-q,q]$ is called the **period** of the word. i.e. $(g^i(y))[-q,q]$ has period n for $i \geq k$.

<u>Definition 5.3</u>: For $\varepsilon \leq (\min_{j \in S}\{\lambda(j)\})^{2q+1}$, $B_\varepsilon(f)$ is **ultimately periodic** if the word f[-q,q] is ultimately periodic.

<u>Notation</u>: $\overline{O^+(f)}$ is the closure of $O^+(f)$ and for any $A \subseteq X$, A^o will be used to denote the interior of A. $\omega(f)$ will be used to denote the set of contact (adherence) points of $O^+(f)$.

<u>Lemma 5.1</u>: Let g: Y --> Y be a continuous function of the compact metric space Y and $f \in Y$. Then $g(\omega(f)) \subset \omega(f)$.

<u>Proof</u>: By definition of the set of contact points $\omega(f)$ of $O^+(f)$

$$\omega(f) = \{z \in Y | \exists n_i \to \infty \ with \ g^{n_i}(f) \to z\}$$

If $z \in g(\omega(f))$ then $\exists$ $y \in \omega(f)$ such that $z = g(y)$. It must be shown that z also belongs in $\omega(f)$. Since y is a contact point of $O^+(f)$ there exists a subsequence that converges to y. That is:

$$g^{n_i}(f) \to y \ \ as \ \ n_i \to \infty$$

and by the continuity of g,

$$g(g^{n_i}(f)) \to g(y) = z \ \ as \ \ n_i \to \infty$$

Hence z is a contact point of $O^+(f)$ (i.e. $z \in \omega(f)$). ▲

The previous lemma states that if y is a contact point of $O^+(f)$ then so is $g(y)$ and will be used in the proof of the following lemma.

<u>Lemma 5.2</u>: Consider any $f \in \Upsilon$ and $\varepsilon > 0$, then the following statements hold:

i.) $B_\varepsilon(f)$ is closed.

ii.) g is equicontinuous at f iff $f \in B_\varepsilon(f)^\circ$ for all $\varepsilon > 0$.

iii.) the restriction of g to $\overline{O^+(f)}$ is equicontinuous iff $B_\varepsilon(f)$ is ultimately periodic for all $\varepsilon > 0$.

<u>Proof</u>: To show $B_\varepsilon(f)$ is closed it suffices to show that $B_\varepsilon(f)$ contains all its limit points. Recall that

$$B_\varepsilon(f)=\{y \mid d(g^i(y),g^i(f)) \le \varepsilon,\ \forall i \in N_0\}$$

and suppose x is a limit point of $B_\varepsilon(f)$. Then there exists a sequence $\{z_n \in B_\varepsilon(f) \mid n \in N_0\}$ with $z_n \to x$ as $n \to \infty$. Choose N large enough so that for all $n \ge N$, $d(z_n,x) \le \varepsilon$. By the continuity of g, $z_n \to x$ implies $g(z_n) \to g(x)$ and $g^i(z_n) \to g^i(x)$ for $i \in N_0$. That is, for large enough N and $n \ge N$,

$$d(g^i(x),g^i(z_n)) \le \varepsilon.$$

Since $z_n \in B_\varepsilon(f)$ for each n, $d(g^i(z_n), g^i(f)) \le \varepsilon$. Hence, $d(g^i(x), g^i(f)) \le \max\{d(g^i(x), g^i(z_n)); d(g^i(z_n), g^i(f))\} \le \varepsilon$ and $x \in B_\varepsilon(f)$.

To show part (ii.), from the definition of equicontinuity of g, g is equicontinuous at f iff for every $\varepsilon > 0$ $\exists$ a neighborhood C_δ of f such that $\forall y \in C_\delta$ and all g^i ($i \in N_0$),

$$d(g^i(f), g^i(y)) \leq \varepsilon.$$

Hence $f \in B_\varepsilon(f)^\circ$. By the definition of the interior of a set, $f \in B_\varepsilon(f)^\circ$ iff $\exists$ an open neighborhood C_δ around f contained in $B_\varepsilon(f)$. This implies that $\forall\ y \in C_\delta \subset B_\varepsilon(f)$, therefore $d(g^i(y), g^i(f)) \leq \varepsilon$, $\forall i \in N_0$. It remains to prove (iii.). If g restricted to $O^+(f)$ is equicontinuous, choose $\delta > 0$ such that for $y,z \in \overline{O^+(f)}$,

$$d(y,z) \leq \delta \rightarrow d(g^i(y), g^i(z)) \leq \varepsilon \leq (\min_{j \in S}\{\lambda(j)\})^{2q+1}.$$

$\overline{O^+(f)}$, being a closed subset of a compact metric space, is compact and hence sequentially compact. Therefore, there exists $j,k \in N_0$, $j < k$ with $d(g^j(f), g^k(f)) \leq \delta$. Taking $y = g^j(f)$ and $z = g^k(f)$ implies $d(g^{j+i}(f), g^{k+i}(f)) \leq \varepsilon$ by equicontinuity. Hence $B_\varepsilon(f)$ is ultimately periodic.

To prove the converse, suppose that for all $\varepsilon > 0$, $B_\varepsilon(f)$ is ultimately periodic. The sequence of words $f[-q,q]$, $(g(f))[-q,q], \ldots, (g^i(f))[-q,q], \ldots$ has period $p(\varepsilon)$ when $i \geq k(\varepsilon)$.

It suffices to show that every $y \in \overline{O^+(f)}$ staisfies the same condition.

In fact it need only be shown that

$$d(g^i(y), g^{i+p}(y)) \le \varepsilon \le (\min_{j\in S}\{\lambda(j)\})^{2q+1} \text{ if}$$

$i \ge k(\varepsilon) > 0$ (i.e. $d(g^i(y), g^i(g^p(y))) \le \varepsilon \le (\min_{j\in S}\{\lambda(j)\})^{2q+1}$).
The proof is complete if $y \in O^+(f)$ by the ultimately periodic condition, so assume $y \in \omega(f)$. Let $z = g^i(y)$. By lemma 5.1, if y is a contact point of $O^+(f)$ then so is $g^p(y)$. Therefore choose $t \ge k(\varepsilon)$ with

$$d(g^t(f), z) \le \alpha = (\min_{j\in S}\{\lambda(j)\})^{2q+pr+1}$$

(recall that r is the range of the automaton).
From the way in which g is induced by G it follows that

$$d(g^{t+p}(f), g^p(z)) \le \varepsilon \le (\min_{j\in S}\{\lambda(j)\})^{2q+1}$$

and, by choice of t, $d(g^t(f), g^{t+p}(f)) \le \varepsilon$.
Now, by the non-archimedean property of the metric,

$$d(g^p(z), g^t(f)) \le \max\{d(g^{t+p}(f), g^p(z)),\ d(g^{t+p}(f), g^t(f))\} \le \varepsilon$$

and

$$d(g^p(z), z) \le \max\{d(g^t(f), z), d(g^p(z), g^t(f))\}$$
$$\le \varepsilon \le (\min_{j\in S}\{\lambda(j)\})^{2q+1}$$

∎

As a result g is **equicontinuous** at f iff for all $\varepsilon > 0$ $\exists$ $\delta > 0$ such that $C_\delta(f)$, the open ball of radius δ around f, is contained in $B_\varepsilon(f)$.

The following example demonstrates the implication of lemma 5.2 part (iii.).

<u>Example 5.1</u>:

Take the automaton g induced by the following rule on $S = \{0,1\}$:

$$G: S^3 \dashrightarrow S$$

$$G(1,1,1) = 1$$

$$G(a,b,c) = 0 \quad \text{otherwise}$$

Let $f = (0\ 0\ 0\ 0\ 0\ 0\ 0)^1_{-3}$, that is

$$\begin{array}{lcl} & & \qquad\qquad\qquad\qquad\quad 0^{th} \\ & & \qquad\qquad\qquad\qquad\quad \downarrow \\ f & = & \ldots.1\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ldots. \\ g(f) & = & \ldots.1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ldots. \\ g^2(f) & = & \ldots.1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ldots. \\ g^3(f) & = & \ldots.1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ldots. \\ \vdots & & \qquad\qquad\qquad\qquad\quad \vdots \end{array}$$

$$\begin{array}{lcl} & & \qquad\qquad\qquad\qquad\quad 0^{th} \\ & & \qquad\qquad\qquad\qquad\quad \downarrow \\ z & = & \ldots.1\ 1\ \underbrace{0\ 0\ldots..0\ 0\ 0\ 0\ 0\ 0\ 0\ldots..0\ 0}_{n \text{ zeros}}\ 1\ 1\ldots. \end{array}$$

$$y = (0)^0_0, \quad \text{that is } y = \ldots.0\ 0\ 0\ 0\ \overset{0^{th}}{\overset{\downarrow}{0}}\ 0\ 0\ 0\ 0\ldots. \ .$$

Note that both $y, z \in \overline{O^+(f)}$, z being a point in $O^+(f)$ and y a point of $\omega(f)$.

If $\lambda(0) = 0.9$, $\lambda(1) = 0.2$ and $\varepsilon = 0.0001$ there exists $\delta > 0$ such that $d(y,z) \le \delta = (0.9)^n$ (in this case) implies $d(g^i(y),g^i(z)) \le 0.0001 \quad \forall\ i \in N_0$.

Hence g restricted to $\overline{O^+(f)}$ is equicontinuous. As seen, $B_{.0001}(f)$ is ultimately periodic of period one or **ultimately fixed.** For this value of ε, the minimum ultimalely periodic interval is [-44,44] and has a minimum length of 89.

Another example of the previous lemma demonstrates the fact that given $B_\varepsilon(f)$ is ultimately periodic then g, restricted to $\overline{O^+(f)}$, is equicontinuous.

Use the following automaton rule defined on the alphabet S = {0,1} induced by the local rule (Wolfram totalistic rule 24):

$$G: S^5 \dashrightarrow S \quad \text{where}$$

$$G(a,b,c,d,e) = \begin{cases} 1 & \text{if } a+b+c+d+e = 3 \text{ or } 4 \\ 0 & \text{otherwise} \end{cases}$$

Choose $f = (1\ 1\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 1)^{0}_{-6}$

Hence,

$$\begin{array}{rcl}
 & & \qquad\qquad\qquad\qquad\qquad\quad 0^{th} \\
 & & \qquad\qquad\qquad\qquad\qquad\quad \downarrow \\
f & = & \ldots.0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ldots. \\
g\,(f) & = & \ldots.0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ldots. \\
g^2(f) & = & \ldots.0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ldots. \\
\vdots & & \qquad\qquad\qquad\qquad\qquad\quad \vdots \\
g^n(f) & = & \ldots.0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ldots. \\
\vdots & & \qquad\qquad\qquad\qquad\qquad\quad \vdots
\end{array}$$

The sequence is fixed for all iterates $n \geq 2$, and therefore ultimately fixed.

Without loss of generality, λ can be chosen equal to 1/2 for both 0 and 1. Thus for $\varepsilon \leq (1/2)^{2q+1}$, where in this case $q \leq 11$, and $\alpha = (1/2)^{23}$ the open ball (here * = 0 or 1)

$$C_\alpha((0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0)^*_{-11}) \subset B_\varepsilon(f)$$

Hence, g is equicontinuous at f. For any $\varepsilon \leq (1/2)^{2q+1}$ an open ball $C_\alpha(f)$ can be found that is contained in $B_\varepsilon(f)$.
In figure 9 the first 150 iterates are displayed. It is seen that any point x belonging to the open ball

$$C_\alpha((0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0)^*_{-12})$$

(note: * = 0 or 1) also belongs to $B_\varepsilon(f)$. That is, any site less than -12 and greater than 12 is chosen arbitrarily. Notice the two persistent structures in the center of the diagram. these represent the forward iterations of sequences of three ones at the -6, -5, -4, 7, 8, and 9th places.

It is important to note that the $B_\varepsilon(f)$ sets are not necessarily open, as the following example illustrates:
If σ is the shift map induced by the local rule G: S^3 --> S where G(a,b,c) = c then $\forall$ f $\in$ Y, $B_\varepsilon(f)$ does not contain any open ball. For instance for $S = \{0,1\}$ and $f = (0)^0_0$ with $\lambda(0) = \lambda(1) = 1/2$, $g^i(f) = f$ for all i. For $\alpha = 1/32$

$$C_\alpha(f) = \{(0\ 0\ 0\ 0\ 0)^*_{-2}\}$$

and $C_\alpha(f) \not\subset B_\varepsilon(f)$ since $z = (0\ 0\ 0\ 0\ 0\ 1)^*_{-2} \in C_\alpha(f)$ and

$$z \notin B_\varepsilon(f).$$

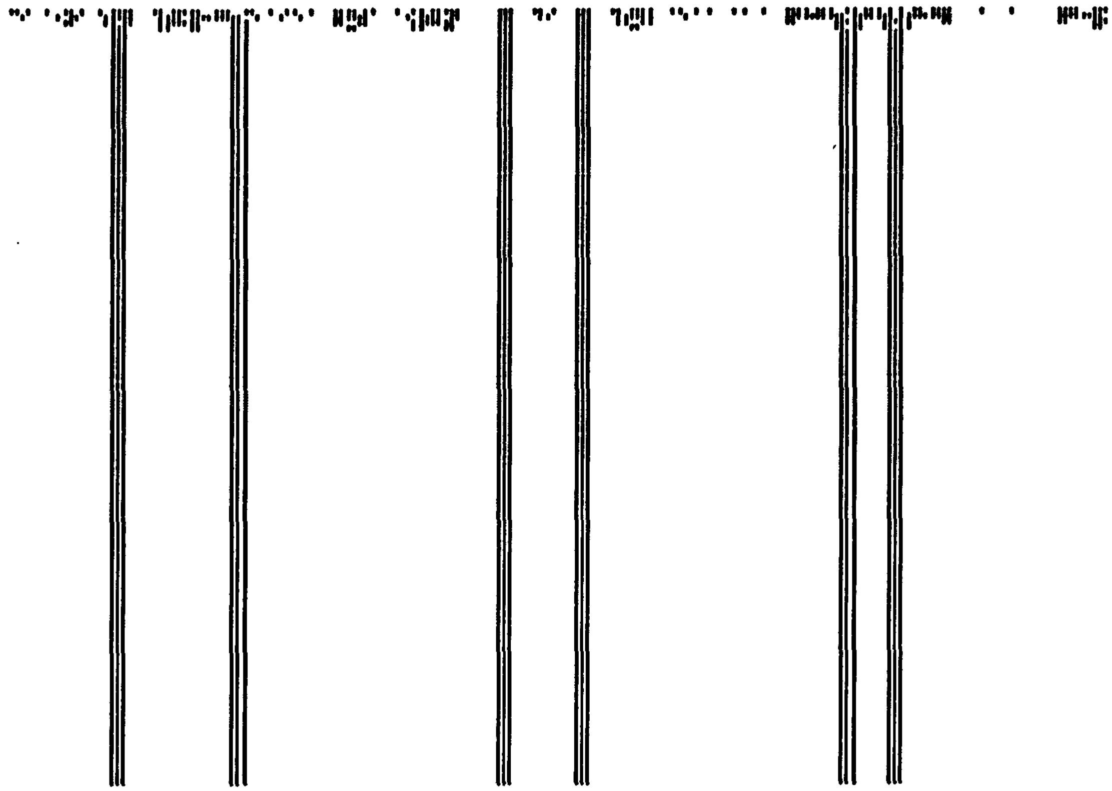


Figure 9

Chapter VI

Classes

The set $B_\varepsilon(f)$ consists of points (functions) from $Y = S^Z$ hence the shift map σ can be defined on $B_\varepsilon(f)$. Referring to the terminology in chapter 3, y has behavior $B_\varepsilon(f)$ on $[-m_i+k, m_i+k]$ if, for some $k \in Z$, $\sigma^{-k}(y) \in B_\varepsilon(f)$ for all $i \in N_0$. Since σ is a homeomorphism, y has behavior $B_\varepsilon(f)$ on $[-m_i+k, m_i+k]$ can be restated as $y \in \sigma^k(B_\varepsilon(f))$ for all $i \in N_0$. That is the behavior of y is known on the vertical jagged edge strip determined by the intervals $[-m_i+k, m_i+k]$ for all $i \in N_0$. For example let $S = \{0,1\}$, g is the additive mod 2 rule defined by the local rule $G: S^3 \dashrightarrow S$ where $G(a,b,c) = (b + c) \bmod 2$. Let $\lambda(0) = 1/2$ and $\lambda(1) = 1/2$ and choose $\varepsilon = (1/2)^3$. Then $\inf_{i \in N_0}(m_i) = 1$ and if

				0^{th} ↓								
$f =$	0	0	0	0	0	1	1	1	1	1	1	1....
$g(f) =$	0	0	0	0	1	0	0	0	0	0	0	0....
$g^2(f) =$	0	0	0	1	1	0	0	0	0	0	0	0....
$g^3(f) =$	0	0	1	0	1	0	0	0	0	0	0	0....
$g^4(f) =$	0	1	1	1	1	0	0	0	0	0	0	0....

and

$$
\begin{array}{rcl|ccc|cccccccc}
 & & & & 0^{th} \downarrow & & & & & & & & & \\
y & = & \ldots.0\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\ldots. \\
g\,(y) & = & \ldots.0\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\ldots. \\
g^2(y) & = & \ldots.1\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\ldots. \\
g^3(y) & = & \ldots.0\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\ldots. \\
g^4(y) & = & \ldots.1\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\ldots.
\end{array}
$$

then $y \in \sigma^5(B_\varepsilon(f))$,

that is y has behavior $B_\varepsilon(f)$ on $[-m_i+5, m_i+5]$.

Use the array visualization for the behavior of f (the iterations of f under a linear automaton map g), $(a_{i,j})$, with i rows and j columns, where each entry $a_{i,j} = (g^i(f))(j)$ $\forall i \in N_0$ and $j \in \mathbb{Z}$. Then if g has range r and if $(a_{i,j})$ is known on two vertical jagged edge strips, each of width at least r, and the top row of $(a_{i,j})$ between the two strips is known, then that part of $(a_{i,j})$ between the two jagged vertical strips is uniquely determined. This yields, as a consequence:

<u>Lemma 6.1</u>: Given a finite alphabet S let $\sigma: Y \to Y$ be the right shift map and let $f \in Y$. Let $\lambda_{\wedge} = \min_{j \in S}\{\lambda(j)\}$ and $\lambda_{\vee} = \max_{j \in S}\{\lambda(j)\}$. If g is an automaton map of range $r \leq 2n + 1$ for some $n \in N_0$ and:

i.) $f \in \sigma^{-k}(B_{\varepsilon}(y)) \cap \sigma^{u}(B_{\varepsilon}(z))$

for $0 < \varepsilon \leq (\lambda_{\wedge})^{2n+1}$ and $k,u \in N_0$

ii.) let $c,q \in N_0$ where $0 \leq c \leq k,u \leq q$

and

$0 < \alpha \leq (\lambda_{\wedge})^{2q+1}$

$0 < (\lambda_{\vee})^{2c+1} \leq \delta$

Then:

$\sigma^{-k}(B_{\varepsilon}(y)) \cap \sigma^{u}(B_{\varepsilon}(z)) \cap C_{\alpha}(f) \subset B_{\delta}(f)$

<u>Remarks</u>: a.) Condition (i) merely states that f has behavior $B_{\varepsilon}(y)$ and behavior $B_{\varepsilon}(z)$ on the intervals $[-m_i-k, m_i-k]$ and $[-t_i+u, t_i+u]$, respectively. The upper bound of $(\lambda_{\wedge})^{2n+1}$ for ε places a minimum on the size of the intervals $[-m_i-k, m_i-k]$ and $[-t_i+u, t_i+u]$.

b.) The upper bound of $(\lambda_{\wedge})^{2q+1}$, for α, in condition (ii) places a minimun on the size of the intervals for elements of $C_{\alpha}(f)$, that is if $x \in C_{\alpha}(f)$ then at least $x[-q,q] = f[-q,q]$. This is to insure that the portion of the top row between the two jagged strips is known. The lower bound of $(\lambda_{\vee})^{2c+1}$, for δ, insures that the widths of the intervals $[-c_i, c_i]$ for $B_\delta(f)$ do not necessarily extent pass (outside) the two vertical jagged edge strips. That is if $x \in B_\delta(f)$ then $x[-c,c] = f[-c,c]$.

<u>Proof</u>: Suppose $x \in \sigma^{-k}(B_\varepsilon(y)) \cap \sigma^{u}(B_\varepsilon(z)) \cap C_\alpha(f)$ it must be shown that $x \in B_\delta(f)$. That is it must be shown that $d(g^i(x), g^i(f)) \le \delta \quad \forall\ i \in N_0$.

The lemma is true for $i = 0$ since $d(x,f) \le \alpha \le \delta$.

Hence at least $x[-q,q] = f[-q,q]$, which implies

$$x[-k,u] = f[-k,u]$$

$x \in \sigma^{-k}(B_\varepsilon(y)) \cap \sigma^{u}(B_\varepsilon(z))$ and $f \in \sigma^{-k}(B_\varepsilon(y)) \cap \sigma^{u}(B_\varepsilon(z))$ implies:

$$(g^i(f))[-k,n-k] = (g^i(x))[-k,n-k] \quad \text{and}$$
$$(g^i(f))[u-n,u] = (g^i(x))[u-n,u]\ \forall\ i \in N_0.$$

Since $x[-q,q] = f[-q,q]$ was already established, the previous two equations will be necessary for $i \ge 2$.

Since $r \le 2n+1$, $(g(f))[n-k+1] = (g(x))[n-k+1]$ and

$$(g(f))[n-k+2] = (g(x))[n-k+2]$$
$$\vdots \qquad\qquad \vdots$$
$$(g(f))[u-n-2] = (g(x))[u-n-2]$$

and finally,

$$(g(f))[u-n-1] = (g(x))[u-n-1].$$

Therefore

$$(g(f))[-k,u] = (g(x))[-k,u].$$

So assume $(g^i(f))[-k,u] = (g^i(x))[-k,u]$, then it follows that

$$(g^{i+1}(f))[-k,u] = (g^{i+1}(x))[-k,u].$$

Thus, by induction,

$$(g^i(f))[-k,u] = (g^i(x))[-k,u] \quad \forall\ i \in N_0.$$

To show that $x \in B_\delta(f)$ it must be shown that

$$d(g^i(f),g^i(x)) \le \delta.$$

Let $w = [-k,u]$, now $|w| = u - (-k) + 1 = u + k + 1$

and $2c + 1 \le u + k + 1$ since $c \le k,u$. Hence

$$d(g^i(f),g^i(x)) \le (\lambda_V)^{2c+1} \le \delta. \quad \blacktriangle$$

The previous lemma will be crucial in proving the classification theorems; the illustration is given below in figure 10:

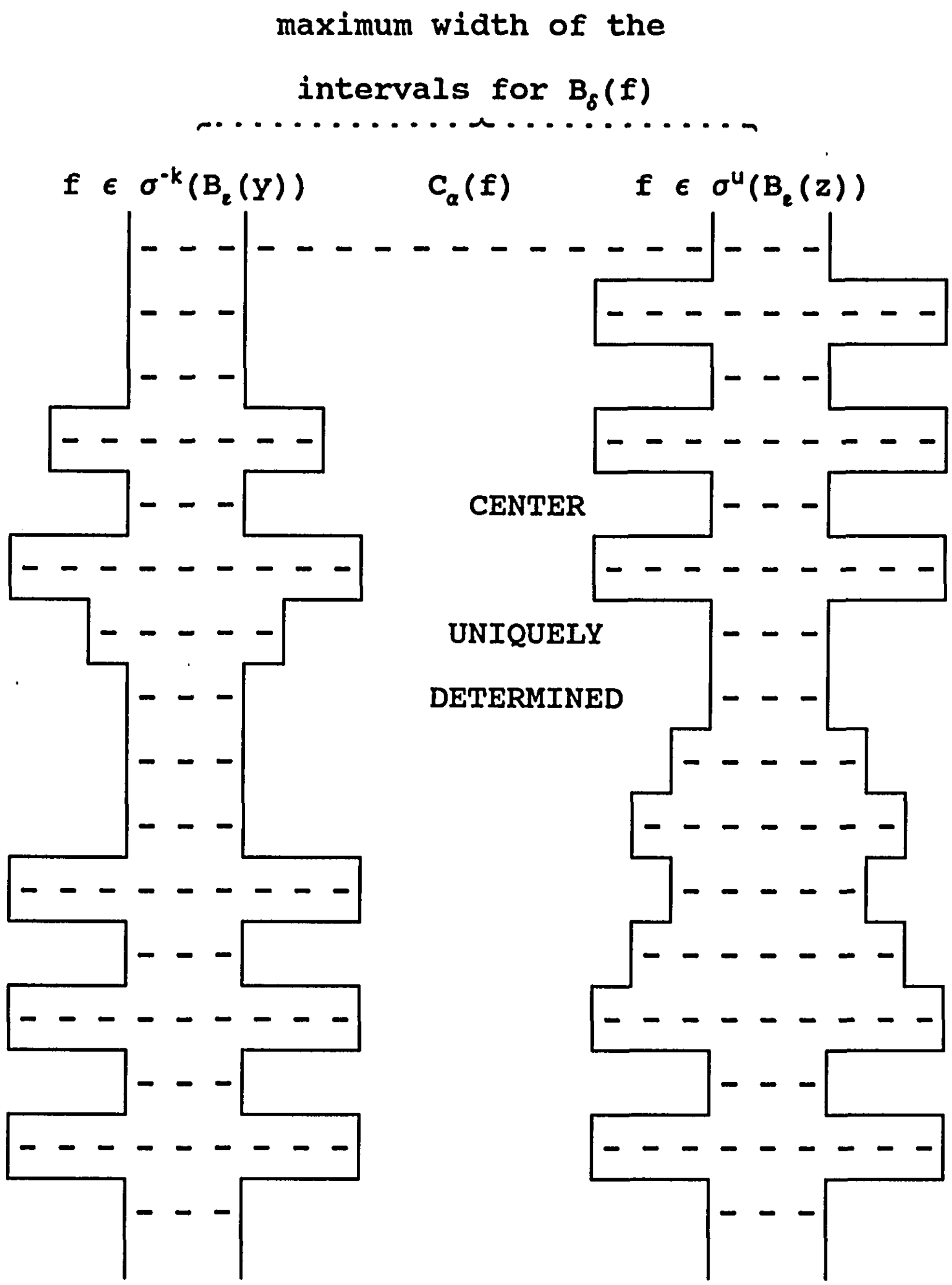
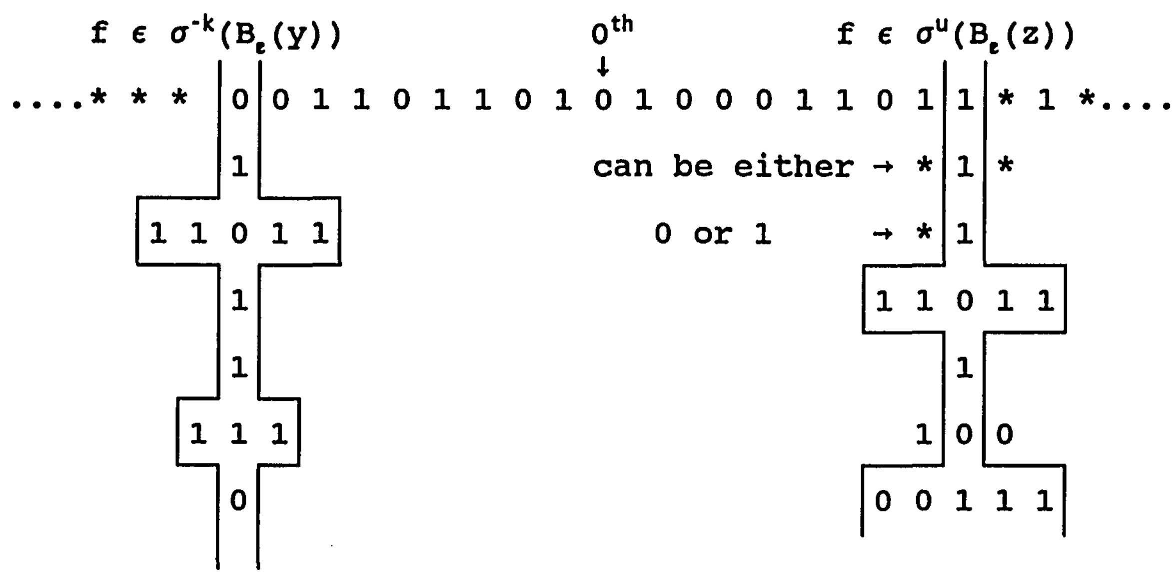


Figure 10

<u>Example 6.1</u>: Suppose that the hypothesis that $r \leq 2n + 1$ is deleted, then $r \not\leq 2n + 1$. It will be shown that if both the right and left infinite vertical jagged edge strips and top row between the strips are known then the region between these two strips is NOT uniquely determined. Let $\mathcal{S} = \{0,1\}$ and suppose $\lambda(0) = \lambda(1) = 1/2$, now $\varepsilon \leq (1/2)^{2n+1}$. Let $c \leq u = k = q = 9$ so that the condition $0 \leq c \leq k,u \leq q$ holds and $\alpha = (\frac{1}{2})^{19}$ for $C_\alpha(f)$. If $g = \sigma_2$ is chosen to be the left 2 - shift induced by the local rule G_σ: $\mathcal{S}^5$ --> $\mathcal{S}$ where $G_\sigma(a,b,c,d,e) = e$ then the map G_σ has range $r = 2$. Let $n = 0$, thus $r = 2 \not\leq 2n + 1 = 1$. Hence $\varepsilon \leq 1/2$ and for the B_ε's $\inf_{i \in No}(m_i) = 1$.



Here * represents an element of $\mathcal{S}$ (i.e. either 0 or 1). Note that the center region is NOT uniquely determined since the * can be either 0 or 1. Hence, for $\delta \geq (\max_{j \in S}\{\lambda(j)\})^{2c+1}$,

$$\sigma^{-k}(B_\varepsilon(y)) \cap \sigma^{u}(B_\varepsilon(z)) \cap C_\alpha(f) \not\subset B_\delta(f).$$

The automaton pulls the value of * accross the infinite jagged vertical strip into the center region. This cannot occur if the most narrow section of the strip has width at least r.

Choose a probability distribution on the alphabet $\mathbf{S}$ such that each element $a \in \mathbf{S}$ has positive probability. i.e. let $n \geq 2$ be a fixed integer and let $(p_0, p_1, p_2, \ldots., p_{n-1})$ be a probability vector whose entries $p_i > 0$ for each i and $\sum_{n=0}^{n-1} p_i = 1$.

<u>Definition 6.1</u>: μ is the corresponding product measure on $\mathbf{Y}$.

<u>Examples 6.2</u>:

i.) For $\mathbf{S} = \{0,1,2,\ldots.,n-1\}$, $\mu(\{i\}) = p_i$.

ii.) For $\mathbf{S} = \mathbf{Z}_p$, the equiprobable measure on $\mathbf{S}$ is 1/p for each element.

<u>Definition 6.2</u>: $\mathbf{T}_\sigma$ is the set of $f \in \mathbf{Y}$ with dense forward and backward orbits under the shift σ. Using more formal notation:

$$\mathbf{T}_\sigma = \{y \in \mathbf{Y} \mid \overline{O_\sigma^{+}(y)} = \mathbf{Y}\} \cap \{y \in \mathbf{Y} \mid \overline{O_\sigma^{-}(y)} = \mathbf{Y}\}.$$

<u>Note</u>: $f \in \mathbf{T}_\sigma$ iff every finite word can be written as f[i,j] for $i,j \in \mathbf{Z}$ with $0 \leq i \leq j$ and also every finite word can be written as f[i,j] for $i,j \in \mathbf{Z}$ with $i \leq j \leq 0$.

Lemma 6.2: $\mu(T_\sigma) = 1$, in particular σ is topologically transitive (see also [W, theorems 5.15 and 5.16]).

Proof: By theorem 2.4 $\mu(\{x \in X \mid \overline{O_\sigma^+(x)} = X\}) = 1$.
However

$$T_\sigma = \{y \in Y \mid \overline{O_\sigma^+(y)} = Y\} \cap \{y \in Y \mid \overline{O_\sigma^-(y)} = Y\}$$

But

$$\mu(\{y \in Y \mid \overline{O_\sigma^+(y)} = Y\}) + \mu(\{y \in Y \mid \overline{O_\sigma^-(y)} = Y\}) = \mu(\{y \in Y \mid \overline{O_\sigma^+(y)} = Y\} \cap \{y \in Y \mid \overline{O_\sigma^-(y)} = Y\}) + \mu(\{y \in Y \mid \overline{O_\sigma^+(y)} = Y\} \cup \{y \in Y \mid \overline{O_\sigma^-(y)} = Y\})$$

and since μ is a probability measure, the result follows. ▲

Lemma 6.3: $Y - T_\sigma$ is first category.

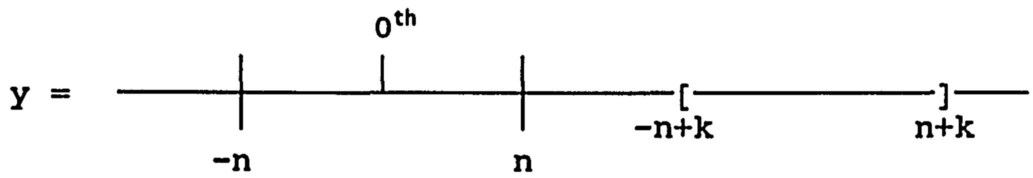
Proof: First note that for each ε $\exists$ a minimum interval $[-n,n]$, $n \in N_0$, such that, for any point $z \in C_\varepsilon(f)$, $z[-n,n] = f[-n,n]$.

$$W_{C_\varepsilon} = \bigcap_{i=-\infty}^{\infty} \sigma^i(Y - C_\varepsilon(f)) = \bigcap_{i=-\infty}^{\infty} \{\sigma^i(x) \mid (\sigma^i(x))[-n,n] \neq f[-n,n]\}$$

The set W_{C_ε} is closed being the infinite intersection of sets that are both open and closed.

claim: W_{C_ε} has empty interior.

proof of claim: Suppose Not, then $\exists$ some $y \in W_{C_\varepsilon}$ with an open ball $C_{\varepsilon'}(y)$ around it. Hence $\sigma^{-k}(y[-n+k,n+k]) = f[-n,n]$ for some $k \in N$, as illustrated:



and there is a contradiction. Hence W_{C_ε} has empty interior.

$$Y - T_\sigma = \bigcup W_{C_\varepsilon}$$

where the union is countable since there are a countable number of equivalence classes of C_ε balls.

Therefore $Y - T_\sigma$ is first category. ▲

Since the whole space Y is second category, T_σ must be second category. Hence it would seem that there would be more points with dense forward and backward orbits under the shift than points without this property. However it is surprisingly easy to find points of the less common type as the next two examples show.

<u>Examples 6.3</u>:

i.) For $S = \{0,1\}$ and if $f = (\ldots.1110111\ldots.) = (101)^{1}{}_{-1}$ then $f \notin T_{\sigma}$, Hence $f \in Y - T_{\sigma}$.

ii.) The whole set of all periodic points, under the shift, belongs to $Y - T_{\sigma}$. This is easily seen since for a point f to be periodic under the shift $\sigma^{p}(f) = f$ is needed. Therefore $f(p+i) = f(i)$ for each $i \in Z$. A periodic point of σ of period p is a fixed point of σ^{p}, hence the points fixed by σ^{p} have the following form:

$$\ldots\ldots,f(p-1),f(0),f(p+1),\ldots.f(p-1),\underset{\uparrow\, 0^{th}}{f(0)},f(p+1),\ldots$$
$$\ldots.,f(p-1),f(0),f(p+1)\ldots.)$$

iii.) For $S = \{0,1\}$ define

f = (....001 100 000 11 01 10 00 1 0 1 00 10 01 11 000 100

3-blocks 2-blocks 2-blocks 3-blocks

100 001 010 101 110 011 111 0000 1000 0001....)

3-blocks 4-blocks

i.e. f consists of all permutation n-blocks of 0's and 1's then all permutation n+1-blocks, etc. listed consecutively starting at the 0^{th} entry and continuing both on the left and right sides. Obviously, some forward and backward iterate of σ will produce any finite word, of any length, desired. Hence $f \in T_{\sigma}$.

The following definition develops a measure-theoretic analogue for the interior of a set.

<u>Definition 6.3</u>: For any measurable set E,

$$\rho_E(f) = \lim_{\alpha\to 0}\frac{\mu(C_\alpha(f)\cap E)}{\mu(C_\alpha(f))}$$

is the **density** of E at the point f.

<u>Lemma 6.4</u>: (Lebesgue Density) For almost all $f \in E$, $\rho_E(f) = 1$.

<u>Proof</u>: Let E_m be the set of f in E such that:

$$\liminf_{\alpha\to 0}\frac{\mu(C_\alpha(f)\cap E)}{\mu(C_\alpha(f))} \geq 1-\frac{1}{m}$$

If $\mu(E_m) = \mu(E)$ for all m, the proof is complete. Therefore, assume $\mu(E_m) < \mu(E)$ for some m. Replacing E by $E - E_m$ implies $\mu(E) > 0$ and that for each $f \in E$

$$\mu(C_\alpha(f)\cap E) < (1-\frac{1}{m})\mu(C_\alpha(f))$$

for infinitely many α and infinitely many $C_\alpha(f)$, where $0 < \alpha < 1$.

The measure μ is regular [see **W**, theorem 6.1] and hence the set E may be replaced by a closed subset and assume that E is compact.

For any $\varepsilon > 0$, E is covered by a finite union of balls,

$$C_i = C_{\alpha_i}(y_i)$$

with

$$\mu(E) \geq \sum \mu(C_i) - \varepsilon\mu(E)$$

Factoring gives

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

Choose ε such that $1/(1+\varepsilon) > 1-1/m$. Given two C_i's whose intersection is non-empty one contains the other, hence it can be assumed that $C_i \cap C_j = \emptyset$ for $i \neq j$. For each i and $C_i \cap E$ choose $\alpha \geq \alpha_i$ such that

$$\mu(C_\alpha(f) \cap E) < \left(1-\frac{1}{m}\right)\mu(C_\alpha(f))$$

is true. A finite number of the $C_\alpha(f)$'s are pairwise disjoint and cover E. Since each $C_\alpha(f)$ lies entirely in a C_i,

$$\mu(C_\alpha(f) \cap E) < \left(1-\frac{1}{m}\right)\mu(C_\alpha(f))$$

implies

$$\mu(E) < \left(1-\frac{1}{m}\right)\sum \mu(C_i)$$

which contradicts

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

hence

$$\mu(E_m) \nless \mu(E)$$

for all m. ▲

The following lemma is necessary since the shift map σ does not preserve the metric.

<u>Lemma 6.5</u>: [see **G1**, lemma 2.5] For any measurable set E $\subset$ **X** let

$$E' = \{f \in E\colon \rho_{\sigma^i(E)}(\sigma^i(f)) = 1, i \in Z\}$$

where ρ is defined as above.

For any measurable subsets E, F $\subset$ **X**:

i.) $\mu(E') = \mu(E)$

ii.) $\sigma^i(E') = [\sigma^i(E)]'$, $i \in \mathbf{Z}$

iii.) $E \subset F$ implies $E' \subset F'$

iv.) $E' \cap F' \subset (E \cap F)'$

The definition of equicontinuous linear cellular automata was discussed in chapter V. Other possible types of linear automata will now be defined. The following definition uses the product measure to define a stochastic analogue of equicontinuity. Figure 11 displays the implication of the definition of almost equicontinuous.

<u>Definition 6.4</u>: g is **almost equicontinuous** at f if for all $\varepsilon > 0$

$$\lim_{\alpha \to 0} \frac{\mu(C_\alpha(f) \cap B_\varepsilon(f))}{\mu(C_\alpha(f))} = 1$$

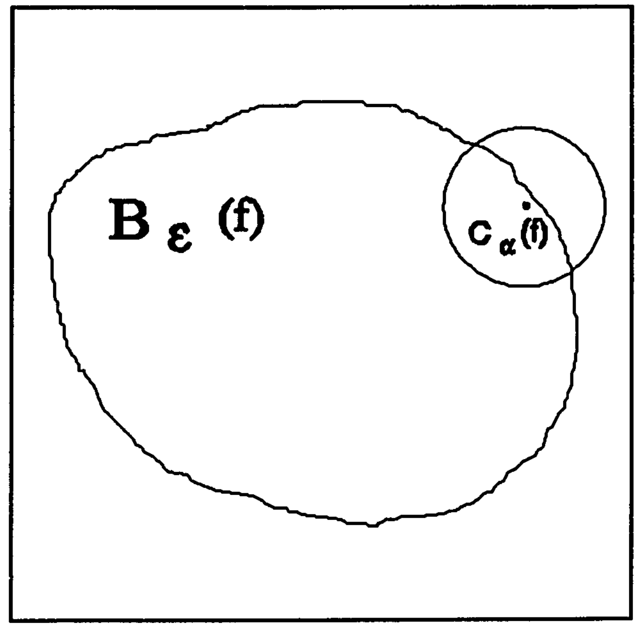


Figure 11

<u>Definition 6.5</u>: g is **almost expansive** if there is $\varepsilon > 0 \ni$ for all $f \in Y$ $\mu(B_\varepsilon(f)) = 0$.

<u>Definition 6.6</u>: g is **expansive** if $\exists\ \varepsilon > 0 \ni$ for all f $B_\varepsilon(f) = \{f\}$.

<u>Definition 6.7</u>: The three classes of linear automata are defined:

i.) $g \in A$ if g is equicontinuous at some $f \in Y$.

ii.) $g \in Б$ if g is almost equicontinuous at some $f \in Y$ but $g \notin A$.

iii.) $g \in \Gamma$ if g is almost expansive.

Clearly, all expansive linear automata belong to class Γ.

The following theorems give a description about the classes Α, Б and Γ. In particular, they imply that these classes form a partition of linear automata.

Theorem 6.6: The following are equivalent:

i.) $g \in$ Α

ii.) g is equicontinuous at some $f \in \mathbf{Y}$

iii.) g is equicontinuous on a set of measure 1

iv.) g is equicontinuous on $\mathbf{T}_\sigma$

v.) for some $n \geq (r - 1)/2$ and $0 < \varepsilon \leq (\min_{j \in S}\{\lambda(j)\})^{2n+1}$ there exists a class $B_\varepsilon(f)$ with $B_\varepsilon(f)^\circ \neq \emptyset$

vi.) $\forall\ \varepsilon > 0$ there exists a class $B_\varepsilon(f)$ with $B_\varepsilon(f)^\circ \neq \emptyset$

Proof: It will be shown that

$$(i.) \Leftrightarrow (ii.)$$

$$(iv.) \Rightarrow (iii.) \Rightarrow (ii.) \Rightarrow (vi.) \Rightarrow (v.)$$

and finally (v.) $\Rightarrow$ (iv.).

By the definition of Α (i.) is equivalent to (ii.). Now, since g is equicontinuous on $\mathbf{T}_\sigma$ and $\mu(\mathbf{T}_\sigma) = 1$, g is equicontinuous on a set of measure 1. Hence (iv.) $\Rightarrow$ (iii.). g is equicontinuous on a set of measure 1, and therefore this set is non-empty. Hence (iii.) $\Rightarrow$ (ii.). By lemma 5.2 part (ii), (ii.) $\Rightarrow$ (vi.). Given (vi.) $\forall\ \varepsilon > 0$ there exists a class $B_\varepsilon(f)$ with $B_\varepsilon(f)^\circ \neq \emptyset$ and this is particularly true when

$n \geq (r-1)/2$ and $\varepsilon \leq (\min_{j\in S}\{\lambda(j)\})^{2n+1}$, therefore (vi.) → (v.). It remains to show (v.) → (iv.). To do this choose $B_\varepsilon(f)$ as in (v.). If $y \in T_\sigma$, then $\forall\ \delta > 0\ \exists\ c \in N_0$ and $\exists\ k,u \geq c\ \ni$ $0 < (\max_{j\in S}\{\lambda(j)\})^{2c+1} \leq \delta$ with $\sigma^u(y) \in B_\varepsilon(f)^\circ$ and $\sigma^{-k}(y) \in B_\varepsilon(f)^\circ$ (by the denseness of T_σ). Since σ is a homeomorphism, $y \in \sigma^{-u}(B_\varepsilon(f))^\circ \cap \sigma^k(B_\varepsilon(f))^\circ$. By lemma 6.1 $y \in B_\delta(y)^\circ$. Since δ was arbitrarily choosen, lemma 5.2 part (ii) implies g is equicontinuous at y. ∎

The following examples have important implications:

<u>Examples 6.4</u>: i.) For the alphabet $S = \{0,1\}$ Let g be the automaton induced by G: S^3 --> S where $G(1,1,1) = 1$, and $G(x,y,z) = 0$ if x, y and z are not all equal to 1 (see figure 12). Using bound vector notation, for $f = (0)^0{}_0$ let $\lambda(j) = 1/2$ $\forall\ j \in S$. Choose $\alpha \leq 1/8$ and $\varepsilon = 1/2$, now $C_\alpha(f) \subset B_\varepsilon(f)$ hence $g \in A$ by Theorem 6.6 part (v.). Note that $(1)^1{}_0 \notin T_\sigma$ and that $\lim_{n\to\infty} g^n(y) = (0)^0{}_0$ for all $y \in Y$ except for $y = (1)^1{}_0$, in that case $\lim_{n\to\infty} g^n(y) = (1)^1{}_0$. Therefore class A contains members which are not equicontinuous at all $f \in Y$. Referring to figure 12, and using the terminology of [W2], a random configuration of sites is chosen. The sites containing a 1 are on and a pixel is placed at that site. The sites containing a zero are off and are left blank. The automaton, as seen, evolves to a homogeneous state where all points are eventually fixed to $(0)^0{}_0$.

Another, but slightly more complex example is:

ii.) For the alphabet $\mathbf{S} = \{1,2,3\}$ let g be the automaton induced by G: $\mathbf{S}^3 \dashrightarrow \mathbf{S}$ where G(2,2,2) = 2, G(3,3,3) = 3 and G(a,b,c) = 1 otherwise. Again, using bound vector notation, for f = $(0)^0_0$ let $\lambda(j) = 1/2$ $\forall j \in \mathbf{S}$. Choose $\alpha \le 1/8$ and $\varepsilon = 1/2$, now $C_\alpha(f) \subset B_\varepsilon(f)$ hence $g \in A$ by Theorem 6.6 part (v.). Note that $(2)^2_0 \notin \mathbf{T}_\sigma$ and $(3)^3_0 \notin \mathbf{T}_\sigma$. It is easily seen that $\lim_{n\to\infty} g^n(y)=(1)^1_0$ except for $y = (2)^2_0$ or $y = (3)^3_0$. Hence, it is again demonstrated that class A can contain members which are not equicontinuous at all $f \in \mathbf{Y}$.

It is seen that if $g \in A$ and $f \in \mathbf{T}_\sigma$, then $f \in B_\varepsilon(f)^\circ$ for all $\varepsilon > 0$. Thus $C_\alpha(f) \subset B_\varepsilon(f)$ for some $\alpha > 0$. This demonstrates an important characteristic of class A cellular automata. The approximate behavior of f, on the infinite vertical jagged edge strip defined by the intervals $[-m_i,m_i]$ for $i \in N_0$ (whose width is determined by ε), is decided by a finite amount of information about f, namely f(n) for $|n| \le q$, where q is the size of the interval determined by α. It will be shown, later, that this is not the case for the other classes of linear automata.

Figure 12

Class A Cellular Automata
RULE: G(1,1,1) = 1 Otherwise G(a,b,c) = 0

<u>Theorem 6.7</u>: The following are equivalent:

i.) $g \in A \cup \bar{B}$

ii.) g is almost equicontinuous at some $f \in Y$

iii.) g is almost equicontinuous on a set of measure 1

iv.) for some $n \geq (r-1)/2$ and $0 < \varepsilon \leq (\min_{j \in S}(\lambda(j)))^{2n+1}$ there exists a class $B_\varepsilon(f)$ with $\mu(B_\varepsilon(f)) > 0$

v.) for all $\varepsilon > 0$ there exists a class $B_\varepsilon(f)$ with $\mu(B_\varepsilon(f)) > 0$

<u>Proof</u>: (i.) and (ii.) are equivalent from the definitions. Next the following implications will be shown:

$$(iii.) \Rightarrow (ii.)$$
$$(ii.) \Rightarrow (v.)$$
$$(v.) \Rightarrow (iv.)$$
$$(iv.) \Rightarrow (iii.)$$

Since a set of positive measure is non-empty (iii.) $\Rightarrow$ (ii.). To show (ii.) $\Rightarrow$ (v.) suppose (ii.) holds and $\mu(B_\varepsilon(f)) = 0$ for some $\varepsilon > 0$. Now $\mu(B_\varepsilon(f)) = \mu(B_\varepsilon(f) \cap C_\alpha(f)) + \mu(B_\varepsilon(f) - C_\alpha(f))$. Hence $\mu(B_\varepsilon(f) \cap C_\alpha(f)) = 0$ for $\alpha > 0$ and g is not almost equicontinuous at f. To show (v.) $\Rightarrow$ (iv.) choose a class $B_\varepsilon(f)$ with $\varepsilon \leq (\min_{j \in S}(\lambda(j)))^{2n+1}$ for $n \geq (r-1)/2$. It remains to show that (iv.) implies (iii.). Choose $B_\varepsilon(f)$ as in (iv.). Define $\mathbf{U}$ as $\mathbf{U} = \mathbf{U}_+ \cup \mathbf{U}_-$ where

$$U_+ = \bigcap_{N=0}^{\infty} \bigcup_{i=N}^{\infty} \sigma^{-i}(B_\varepsilon(f)')$$

and

$$U_- = \bigcap_{N=0}^{\infty} \bigcup_{i=N}^{\infty} \sigma^{i}(B_\varepsilon(f)')$$

U is the set of y ϵ Y whose full orbit, under the shift σ, intersects $B_\varepsilon(f)'$ infinetly often. By lemma 6.5,

$$\mu(B_\varepsilon(f)') = \mu(B_\varepsilon(f)) > 0.$$

Now,

$$\sigma^{-1}(\bigcup_{i=N}^{\infty} \sigma^{-i}(B_\varepsilon(f)')) = \bigcup_{i=N+1}^{\infty} \sigma^{-i}(B_\varepsilon(f)')$$

hence

$$\mu(\bigcup_{i=N}^{\infty} \sigma^{-i}(B_\varepsilon(f)')) = \mu(\bigcup_{i=N+1}^{\infty} \sigma^{-i}(B_\varepsilon(f)'))$$

and

$$\mu(\bigcup_{i=0}^{\infty} \sigma^{-i}(B_\varepsilon(f)')) = \mu(\bigcup_{i=N}^{\infty} \sigma^{-i}(B_\varepsilon(f)')) \quad \textit{for all } N \geq 0.$$

Since

$$\bigcup_{i=0}^{\infty} \sigma^{-i}(B_\varepsilon(f)') \supset \bigcup_{i=1}^{\infty} \sigma^{-i}(B_\varepsilon(f)') \supset \bigcup_{i=2}^{\infty} \sigma^{-i}(B_\varepsilon(f)') \supset \ldots.$$

it follows that

$$\mu(U_+) = \mu(\bigcap_{N=0}^{\infty} \bigcup_{i=N}^{\infty} \sigma^{-i}(B_\varepsilon(f)')) = \mu(\bigcup_{i=0}^{\infty} \sigma^{-i}(B_\varepsilon(f)'))$$

A similiar argument shows

$$\mu(U_-) = \mu(\bigcup_{i=0}^{\infty} \sigma^{i}(B_\varepsilon(f)')).$$

Since $\mu(B_\varepsilon(f)') > 0$ and σ is measure preserving it follows that $\mu(U) > 0$. U is σ-invariant since

$$\sigma^{-1}(U_+) = \bigcap_{N=0}^{\infty} \bigcup_{i=N}^{\infty} \sigma^{-(i+1)}(B_\varepsilon(f)') = \bigcap_{N=0}^{\infty} \bigcup_{i=N+1}^{\infty} \sigma^{-i}(B_\varepsilon(f)') = U_+$$

and similiary, $\sigma^{-1}(U_-) = U_-$. Hence, by ergodicity $\mu(U) = 1$. To prove (iv.) implies (iii.) it suffices to show that for all $y \in U$ and $\delta > 0$, $y \in B_\delta(y)'$, since then y will be almost equicontinuous on a set of measure 1. If $y \in U$, then $\forall$ $\delta > 0$ $\exists$ $c \in N_0$ and $\exists$ positive integers $k,u \geq c$ such that $0 < (\max_{j\in S}\{\lambda(j)\})^{2c+1} \leq \delta$ with $\sigma^u(y) \in B_\varepsilon(f)'$ and $\sigma^{-k}(y) \in B_\varepsilon(f)'$ (follows since U is the set of y whose orbit under σ enters $B_\varepsilon(f)'$ infinitely often). Since σ is a homeomorphism

$$y \in \sigma^k(B_\varepsilon(f)') \text{ and } y \in \sigma^{-u}(B_\varepsilon(f)'),$$

thus $y \in \sigma^k(B_\varepsilon(f)') \cap \sigma^{-u}(B_\varepsilon(f)')$ and by lemma 6.5

$$y \in [\sigma^k(B_\varepsilon(f))]' \cap [\sigma^{-u}(B_\varepsilon(f))]'.$$

Hence by lemma 6.1 $y \in B_\delta(y)'$ and since δ was chosen arbitrarily, g is almost equicontinuous at y. ∎

Theorem 6.8: The following are equivalent:

i.) $g \in \Gamma$

ii.) g is almost expansive

iii.) there exists $\varepsilon > 0$ such that $\mu(B_\varepsilon(f)) = 0$ for all $f \in Y$

iv.) for all $0 < \varepsilon \le (\min_{j \in S}\{\lambda(j)\})^{2n+1}$ where $n \in N_0$, $n \ge (r-1)/2$ and all $f \in Y$, $\mu(B_\varepsilon(f)) = 0$.

Proof: By the definition of almost expansive (i.) ⇔ (ii.) ⇔ (iii.). In (iv.) choose $\varepsilon = (\min_{j \in S}\{\lambda(j)\})^{2n+1}$ and by the definition of almost expansive (iv.) ⇒ (ii.).
Lastly, (iii.) ⇒ (iv.) by the equivalency of parts (iv.) and (v.) of the previous theorem. ∎

Example 6.5:

i.) The shift map σ, induced by the local rule $G(a,b,c) = c$, is not expansive since for $\varepsilon > 0$ $B_\varepsilon(f) \neq f$ for some f, but belongs to class Γ. To see this consider the easily generalizable case for all $f \in Y$ where $S = \{0,1\}$ and let $\lambda(j) = 1/4$ $\forall j \in S$. For all $f \in Y$ it follows that $\mu(B_\varepsilon(f)) = 0$ for $\varepsilon = (\min_{j \in S}\{\lambda(j)\})^{2n+1} = (1/4)^{2n+1}$. Choose $n = 0$ and $\varepsilon = 1/4$. Choose a probability distribution p on $\mathcal{S}$ in the following manner:

$p(1) = p$ where $0 < p < 1$ and $p(0) = 1 - p = q$ and assume $p \geq q$.

$$\text{If } f = \ldots.1\ 1\ 1\ \overset{\overset{0^{th}}{\downarrow}}{1}\ 1\ 0\ 0\ 0\ 0\ 0\ldots.$$

then $B_{1/4}(f)$ equals the set of all sequences of the following form (here $* = 0$ or 1):

$$(1\ 1\ 0\ 0\ 0\ 0\ 0)^*_0, \quad \text{call this set } D$$

Define the following sequence of sets:

$$D_1 = \ldots.*\ *\ *\ \overset{\overset{0^{th}}{\downarrow}}{1}\ *\ *\ *\ *\ *\ *\ldots. \qquad \mu(D_1) = p$$

$$D_2 = \ldots.*\ *\ *\ 1\ 1\ *\ *\ *\ *\ *\ldots. \qquad \mu(D_2) = p^2$$

$$D_3 = \ldots.*\ *\ *\ 1\ 1\ 0\ *\ *\ *\ *\ldots. \qquad \mu(D_3) = p^2q \leq p^3$$

$$D_4 = \ldots.*\ *\ *\ 1\ 1\ 0\ 0\ *\ *\ *\ldots. \qquad \mu(D_4) = p^2q^2 \leq p^4$$

$$D_5 = \ldots.*\ *\ *\ 1\ 1\ 0\ 0\ 0\ *\ *\ldots. \qquad \mu(D_5) = p^2q^3 \leq p^5$$

$$\vdots$$

$$D_n = \ldots.*\ *\ *\ \underbrace{1\ 1\ 0\ 0\ 0\ 0\ldots0\ 0\ 0}_{\longleftarrow\ n\ \longrightarrow}\ *\ *\ *\ldots.$$

Therefore $\mu(D_n) \leq p^n$ for each n.

$\{D_n\}$ is a decreasing sequence of sets, that is

$$D_1 \supset D_2 \supset \ldots. \supset D_n \supset \ldots.$$

$$\text{and} \qquad D = \bigcap_{n=1}^{\infty} D_n$$

$$\text{Hence} \quad \mu(D) = \mu(\lim_{n\to\infty} D_n) = \lim_{n\to\infty} \mu(D_n) \le \lim_{n\to\infty} p^n = 0$$

Since it only matters what f is on the right and the way the shift map is induced, all sequences belonging to $B_{1/k}(f)$ must agree with f on the right. A similiar argument holds true for any $f \in \mathbf{Y}$.

ii.) The automaton induced by the following addition mod 2 rule for $S = \{0,1\}$:

$G(a,b,c) = [b + c] \bmod 2$ is almost expansive, that is $\exists$ $\delta > 0 \ni \forall f \in \mathbf{Y}$ $\mu(B_\delta(f)) = 0$, hence belongs to class Γ.

As illustrated:

1 1 1	0 0 0	0 1 1	0 0 1	0 1 0	1 0 1	1 0 0	1 1 0
0	0	0	1	1	1	0	1

see figures 13 and 13A.

For example, choose a random sequence:

```
                                      0th
                                       ↓
  f    = ...0 1 1 1 0 0 1|0 1 1 0 0 0 1|0 0 1 1 1 1 0....
g (f)  =   ...0 0 1 0 1 1 1|0 1 0 0 1|1 0 1 0 0 0....
g²(f)  =     ...1 1 1 0 0 1 1|1 0 1|0 1 1 1 0 0....
g³(f)  =       ...0 1 0 1 0 0 1|1|1 1 0 0 1 0....
g⁴(f)  =         ...1 1 1 0 1 0 0 0 1 0 1 1....
  .                              .
  .                              .
  .                              .
```

If another sequence agrees with f on all but one entry, such as:

```
                                  0th
                                   ↓       ↓
y      = ...0 1 1 1 0 0 1|0 1 1 0 0 0 1|1 0 1 1 1 1 0....
g (y)  =    ...0 0 1 0 1 1 1|0 1 0 0 1|0 1 1 0 0 0....
g²(y)  =       ...1 1 1 0 0 1 1|1 0 1|1 1 0 1 0 0....
g³(y)  =          ...0 1 0 1 0 0 1|1|0 0 1 1 1 0....
g⁴(y)  =             ...1 1 1 0 1 0 1 0 1 0 0 1....
                                   ↑
                           change propagates
```

then this change will propagate towards the center and eventually affect the center portion of the array.

Suppose that $\lambda(0) = \lambda(1) = 1/2$ and choose $\varepsilon \leq (1/2)^{2n+1}$, say $n = 1$ so $\varepsilon = 1/8$.

Hence $\inf_{i \in No}(m_i) = 1$. However, even a single change on the right will propagate into the interval $[-1,1]$.

A similiar argument as in the previous example can be used to show $g \in \Gamma$.

iii.) The following is a generalization of an example which can be found in [G1]:

Let $S = \{0,1,2\}$. Define $G: S^3 \to S$ in the following manner:

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

(Here * = 0, 1, or 2)

If the array visualization $(a_{i,j})$ is used, where $a_{i,j}$ = $(g^i(f))(j)$, 0 is a background element and in passing from one row of $(a_{i,j})$ to the next row it is seen that the 1's move left, the 2's move straight down and a 1 and 2 which collide annihilate each other (that is they produce a 0). The behavior of g can be connected with a random walk on the integers (see [G1] or [PG] for a similiar analysis). Let (p_0,p_1,p_2) be the probability vector that assigns p_0, p_1, p_2 to 0, 1, 2, respectively in defining the measure μ. That is $p(1) = p_1$, $p(2) = p_2$ and $p(0) = p_0$. Consider the random walk with p(1), p(2), and p(0) being the probability of moving left, moving right, and staying stationary, respectively. Suppose f(0) = 2, then the probability that this 2 is never annihilated equals the probability of no return to 0 starting at 1 in the random walk. In a random walk, the probability of returning to 0 starting at 1 is:

$$Q_1 = \begin{cases} 1 & \text{if } p(2) \le p(1) \\ p(1)/p(2) & \text{if } p(2) > p(1) \end{cases}$$

Hence the probability of no return to 0, starting at 1, in the random walk is:

$$P_1 = 1 - Q_1 = \begin{cases} 0 & \text{if } p(2) \le p(1) \\ 1 - p(1)/p(2) & \text{if } p(2) > p(1) \end{cases}$$

As seen, the probability of no return to 0, starting at 1, in the random walk is positive if and only if if p(2) > p(1).

Let $f = (2)^2_0$ and let $\lambda(j) = 1/2 \ \forall \ j \in S$. Thus $p(2) > p(1)$ implies that $\mu(B_{½}(f)) > 0$. Hence, by theorem 6.7 $g \in A \cup Б$. To show that g actually belongs in Б take any $n \in N_0$, if f is chosen such that $f(k) = 0$ for all $k > n$ then $(g^i(f))[0]$ is eventually 0 or 2. If $f(k) = 1$ for all $k > n$ then $(g^i(f))[0]$ is eventually 1. Therefore no finite amount of information about f determines $B_{½}(f)$ hence $g \notin A$. See figure 14. A random configuration, of elements from {0,1,2}, is chosen; the off sites represent 0's and are left blank. The sites containing a 1 or 2 show a pixel.

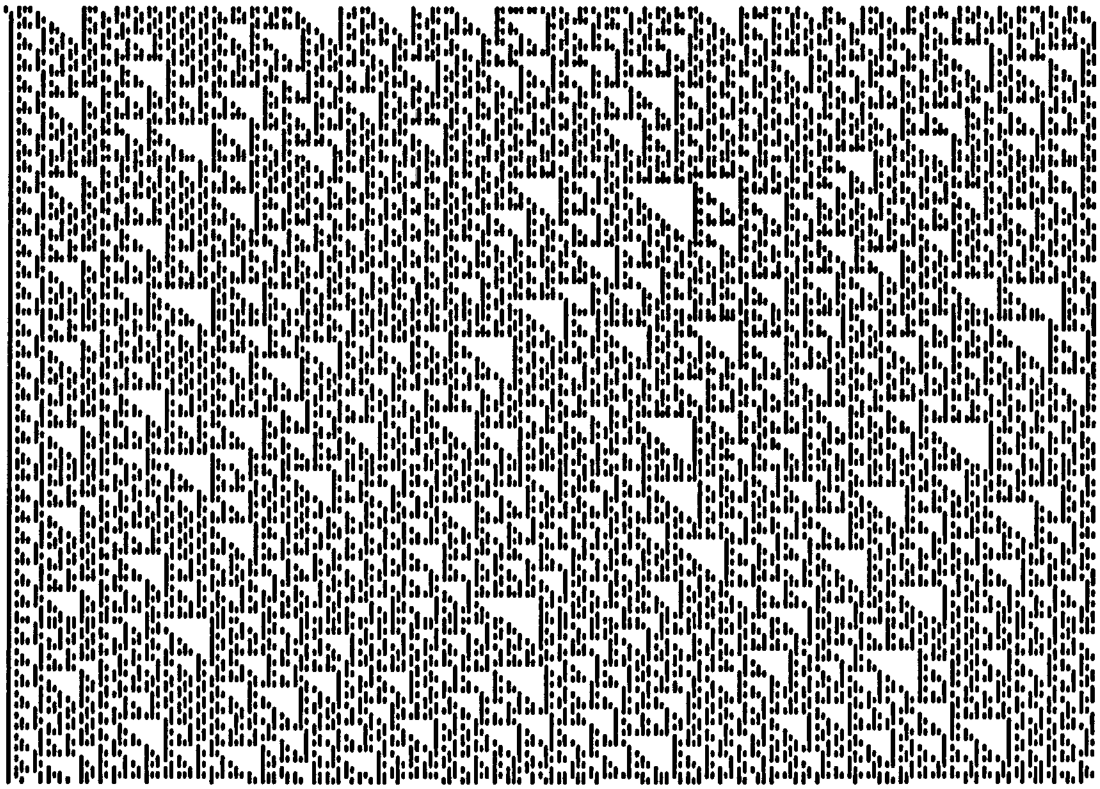


Figure 13A

Class Γ Cellular Automata
RULE: G(a,b,c) = [b + c] mod 2 From a Random Start

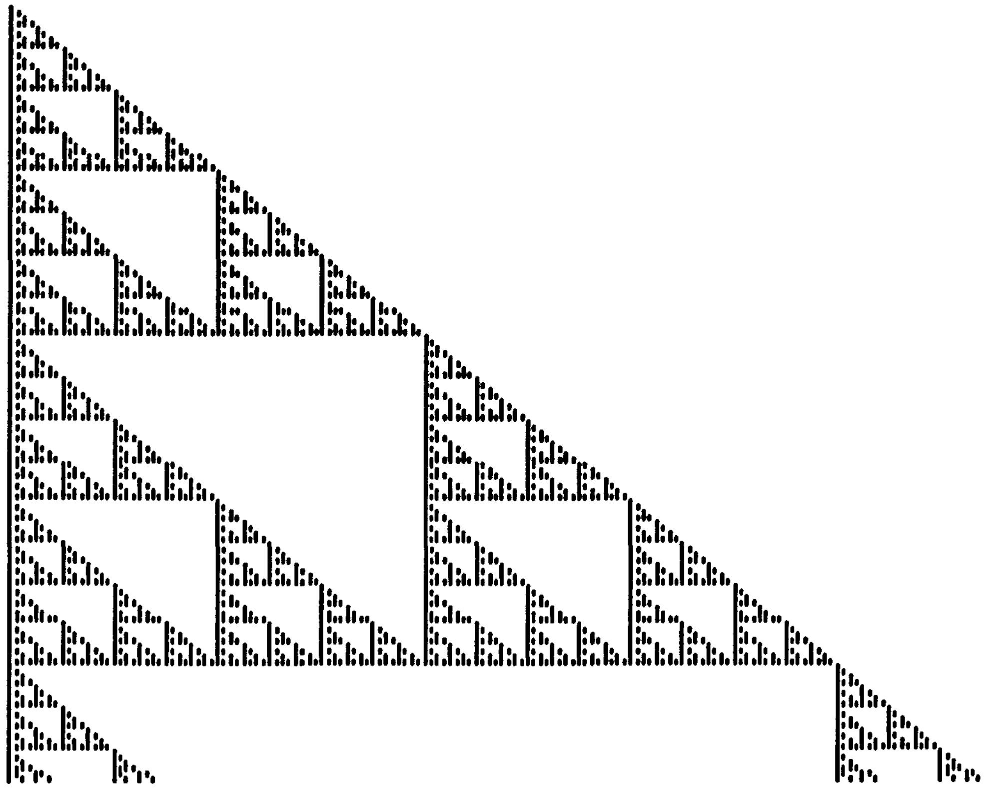


Figure 13B

Class Γ Cellular Automata
RULE: G(a,b,c) = [b + c] mod 2
Starting From a Single Initial Point
All other points are 0

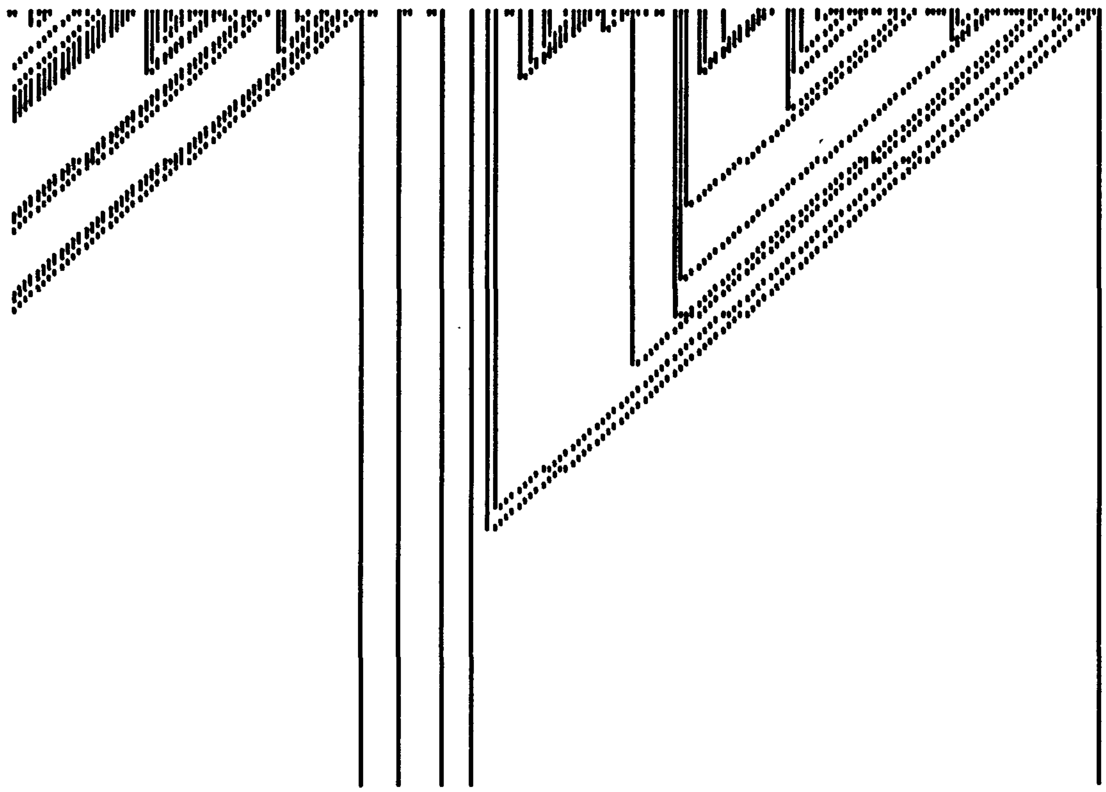


Figure 14

Class Б Cellular Automata
Starting From a Random Initial Configuration of 0, 1 and 2

Linear automata in class $A \cup Б$ behave quite differently. Automata in class A have the property that if $g \in A$ and $f \in T_\sigma$, then the forward orbit of $g(f)$ (i.e. $g^i(f)$ for $i \in N_0$) will stay arbitrarily close to the forward orbit of $g(y)$, assuming y is close enough to f. However, automata belonging to class $Б$ have the property that the probability that $g^i(y)$ will diverge from $g^i(f)$ is positive. As the following theorem demonstrates:

<u>Theorem 6.9</u>: If $g \in Б$, then for all $f \in Y$, all $\alpha > 0$, $r \leq 2n+1$ and $\varepsilon \leq (\min_{j \in S}(\lambda(j)))^{2n+1}$, $\mu(C_\alpha(f) \cap B_\varepsilon(f)) < \mu(C_\alpha(f))$.

<u>Proof</u>: If the strict inequality above fails, then $\mu(C_\alpha(f) \cap B_\varepsilon(f)) = \mu(C_\alpha(f))$ for some $\alpha > 0$ and $f \in Y$. Since $C_\alpha(f)$ is open and $B_\varepsilon(f)$ closed, $C_\alpha(f) - B_\varepsilon(f)$ is open. Now

$$\mu(C_\alpha(f)) = \mu(C_\alpha(f) \cap B_\varepsilon(f)) + \mu(C_\alpha(f) - B_\varepsilon(f))$$

and

$$\mu(C_\alpha(f)) - \mu(C_\alpha(f) \cap B_\varepsilon(f)) = 0.$$

Hence $C_\alpha(f) - B_\varepsilon(f)$ has measure 0 and is therefore empty. By theorem 6.6(v.), $g \in A$. ∎

The previous theorem asserts that if $g \in \bar{B}$ and for $f \in \mathbf{Y}$ then $B_\varepsilon(f)$ cannot be determined by any finite amount of information about f. Theorem 6.6 shows that if $g \in A$ and $f \in \mathbf{T}_\sigma$, then $f \in B_\varepsilon(f)^\circ$ and hence the open ball $C_\alpha(f)$ is contained in $B_\varepsilon(f)$. This says that the behaviour of f, on the infinite vertical jagged edge strip defined by the intervals $[-m_i, m_i]$ (whose width is determined by ε), is decided by a finite amount of information about f.

The following example shows that $B_\varepsilon(f)^\circ \neq \emptyset$ is possible for $g \notin A$ if $\varepsilon \leq (\min_{j \in S}\{\lambda(j)\})^{2n+1}$ when $r \nleq 2n+1$.

<u>Example 6.6</u>: Let $S = \{0,1,2\}$ and let g be the automaton of range 2 induced by the following rule:

$$G(*,*,0,1,*) = G(*,*,0,2,*) = 0$$

$$G(*,*,*,0,2) = G(*,*,2,2,*) = G(*,*,1,2,*) = 2$$

$$\text{and} \quad G(*,*,a,b,c) = 1 \text{ otherwise}$$

(Here * represents any element in S)

Suppose $\lambda(j) = 1/2 \ \forall j \in S$

Since $r = 2$, let $r > 2n+1$ by choosing $n = 0$

Then $\varepsilon \leq (\min_{j \in S}\{\lambda(j)\})^{2n+1} = 1/2$. Take $\varepsilon = 1/2$ and let

```
                0th
                 ↓
         ┌──────────────┐
f = ....*│ *  * │0│1  0│2  0  1  1  0  1  *  *  *....
                │0│2  0  1  0  1  1
                │0│1  0  1  0
                │0│1
                │0│
                │0│
                │ │
```

(where * represents any element in S)

The column of zeros in the center with a non-zero element on the right act like a wall and allow nothing to change them. The non-zero elements, on the right of the column of zeros can change between the elements 1 and 2.

Now $B_{1/2}(f)^{\circ} \neq \emptyset$

Indeed, $C_{1/32}(f)$ is contained in $B_{1/2}(f)$ and thus $B_{1/2}(f)$ contains an open ball and here $B_{1/2}(f)^{\circ} \neq \emptyset$.

However if $\varepsilon \leq (\min_{j\in S}(\lambda(j)))^{2n+1}$ and $r \leq 2n + 1$ then, since the elements on the right of the column of zeros can change between 1 and 2, no finite amount of information determines $B_{\varepsilon}(f)$ and hence g does not belong to class A.

Chapter VII

Applications To Digital Signal Processing

Definition 7.1: A **binary digital signal** is a function

$$f: \mathbf{Z} \text{ --> } \mathbf{Z}_2$$

Hence a binary digital signal is an element belonging to $\mathbf{S}^{\mathbf{Z}}$, where $\mathbf{S} = \mathbf{Z}_2$.

Definition 7.2: The **support** region of a function f, denoted by supp(f), is a subset of $\mathbf{Z}$ outside of which f has value zero.

Example 7.1:

i.) For the signal $f = (1011)^0_{-1}$, supp(f) = {-1,1,2}

ii.) For the signal $f = (1\ 1\ 1\ 1\ 1....)^0_p$, supp(f) = {p, p+1, p+2,....}

Note that as the previous example illustrates, the support of a signal is not necessarily a finite set.

Definition 7.3: A **time limited signal** is a signal with finite support. That is card(supp(f)) < ∞.

Definition 7.4: The set of non-zero values of elements in S is denoted by $S - \{0\}$. The co-zero set of a time limited signal G, denoted coz(G), is defined as $G^{-1}(S-\{0\})$. Therefore, coz(G) is that subset of Z where G does not have a zero value.

Example 7.2: $G = (10011)_0^0$ $\quad$ $coz(G) = \{0,3,4\}$

Definition 7.5: card(coz(f)) is defined as the cardinality of the co-zero set for the signal f.

Example 7.3: $f = (11000101)_{-2}^0$ $\quad$ $coz(f) = \{-2,-1,3,5\}$

$card(coz(f)) = 4$

Note: $card(coz(f)) = 0 \Leftrightarrow coz(f) = \emptyset$

The alphabet S, in use here, is a ring isomorphic to the factor ring $Z_2 \cong Z/2Z$ where the two operations are multiplication and addition modulo 2 i.e. $(Z_2, \cdot, +)$.

Example: in $Z_2 = \{[0],[1]\}$

$$0 \cdot 1 = 0 \quad \text{and} \quad 0 + 1 = 1$$

Note: The set S^Z forms a R-module over the ring $(Z_2, \cdot, +)$.

<u>Definition 7.6</u>: For $f \in Y$, and G a time limited signal taking values in Z_2, the next signal is defined pointwise by the convolution:

$$(f*G)(n) = \sum_{\substack{n-k \in coz(G) \\ k \in coz(f)}} f(k) \cdot G(n-k)$$

Where arithmetic is performed using the ring operations in Z_2.

Two important quantities of the co-zero set of the convolution cellular automaton map will be needed. These are min(coz(G)) and max(coz(G)), where min and max are the minimum and maximum values, respectively, of the co-zero set.

<u>Definition 7.7</u>: $a = \max(coz(G))$ and $b = \min(coz(G))$ and $r = \max(|a|,|b|)$

The quantity r is the **range** of the convolution.

Hence $G: S^{2r+1} \text{----}> S$

and

$g: S^Z \text{----}> S^Z$

The signal G is the local rule or local map for each point n. The local map G induces the global **convolution cellular automata** (hereby denoted CCA) map g(f) defined by g(f) = y with each $y(n) = (f*G)(n) = G(f(n-r),\ldots,f(n),\ldots f(n+r))$.

<u>Example 7.4</u>: Consider the binary signal

$$f = (\ldots 101010101 \ldots) \qquad G = (011)_0^0$$

(↑ $0^{\underline{th}}$)

$$coz(G) = \{1,2\}$$

in Z_2

$$(f*G)(-1) = \sum_{-1-k \in coz(G)} f(k) \cdot G(-1-k)$$

$$= f(-3)\cdot G(2) + f(-2)\cdot G(1)$$

$$= 0\cdot 1 + 1\cdot 1$$

$$= 0 + 1 = 1$$

$$(f*G)(0) = \sum_{-k \in coz(G)} f(k) \cdot G(-k)$$

$$= f(-2)\cdot G(2) + f(-1)\cdot G(1)$$

$$= 1\cdot 1 + 0\cdot 1 = 1$$

$$\sum_{1-k \in coz(G)} f(k) \cdot G(1-k) = f(-1)\cdot G(2) + f(0)\cdot G(1)$$

$$= 0 + 1 = 1$$

Continuing in this way, on both the negative and positive sides of the signal f, the following signal is obtained:

$$g(f) = f*G = (\ldots 11111111 \ldots) = (1)_0^1$$

Upon another iteration of G on f the following signal is obtained:

$$g^2(f) = (f*G)*G = (\ldots 00000000 \ldots) = (0)_0^0$$

and successive iterations $g^i(f) = (0)_0^0$ for all $i \geq 2$.

The following definitions will define the fundamental operations on digital signals:

I.) <u>Definition 7.8</u>: The RANGE of a digital signal is the integer value of the map f at each point in coz(f).

Illustrated by the Block Diagram:

II.) <u>Definition 7.9</u>: The operation ADD is a binary operation (i.e. ADD: $S^Z \times S^Z \longrightarrow S^Z$) defined pointwise by the rule:

$$ADD(f,y)(k) = f(k) + y(k)$$

Illustrated by the Block Diagram:

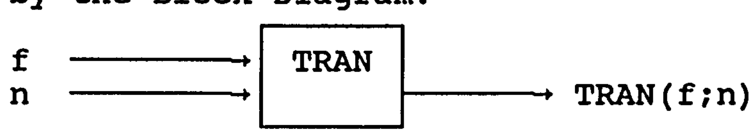
III.) <u>Definition 7.10</u>: The operation TRAN is a binary operation defined as follows

$$TRAN: S^Z \times Z \longrightarrow S^Z$$

where

$$(TRAN(f;n))(k) = f(n-k) \quad \text{and } n \in Z$$

Illustrated by the Block Diagram:



<u>Example 7.5</u>:

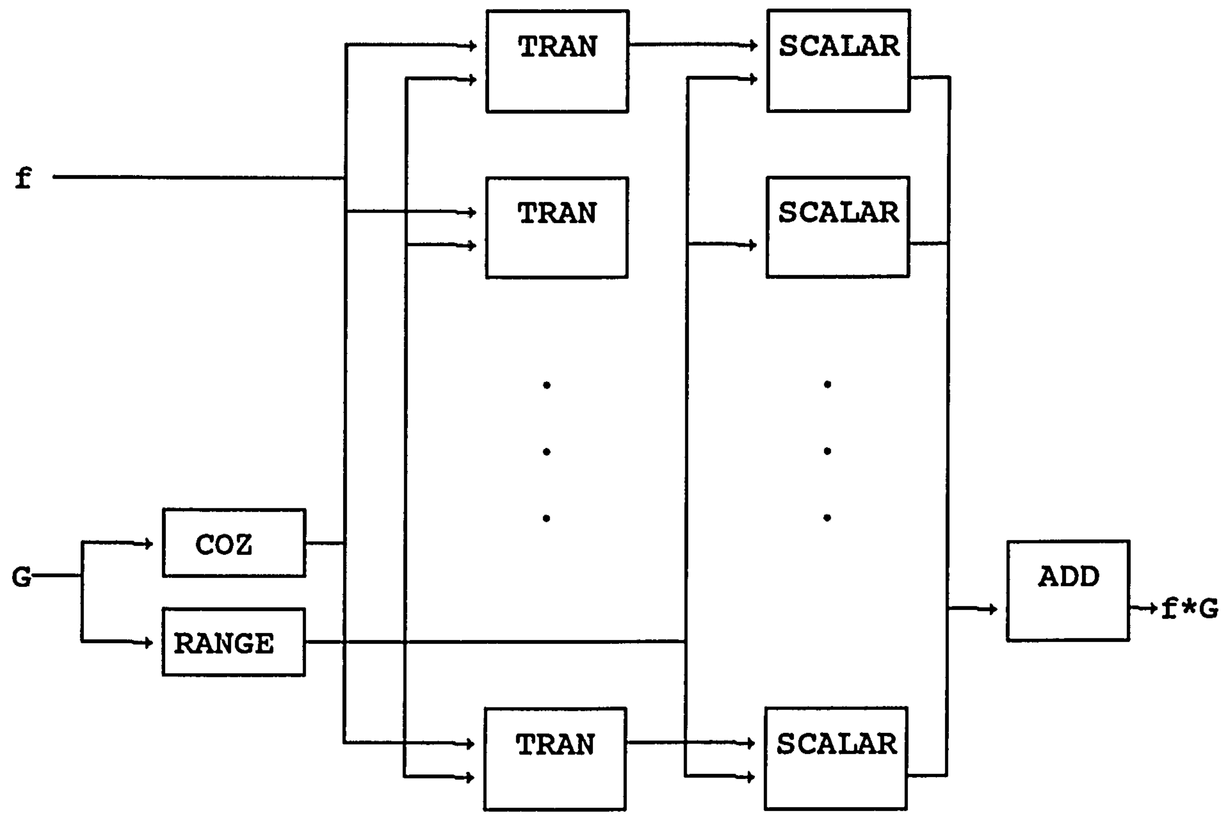
IV.) <u>Definition 7.11</u>: The SCALAR operation is the same as scalar multiplication in a vector space and is defined by:

$$\text{SCALAR}(f;a) = a \cdot f$$

Illustrated by the Block Diagram:

Parallel Convolution Algorithm:

The convolution of the signal f (not necessary time limited) with the local map G can be computed using the Parallel Convolution Algorithm.



Parallel Convolution Algorithm

The following pages demonstrate examples of iterations of the Cellular Convolution Map. The signals are binary signals and the arithmetic is performed using the binary operations of the factor ring Z_2:

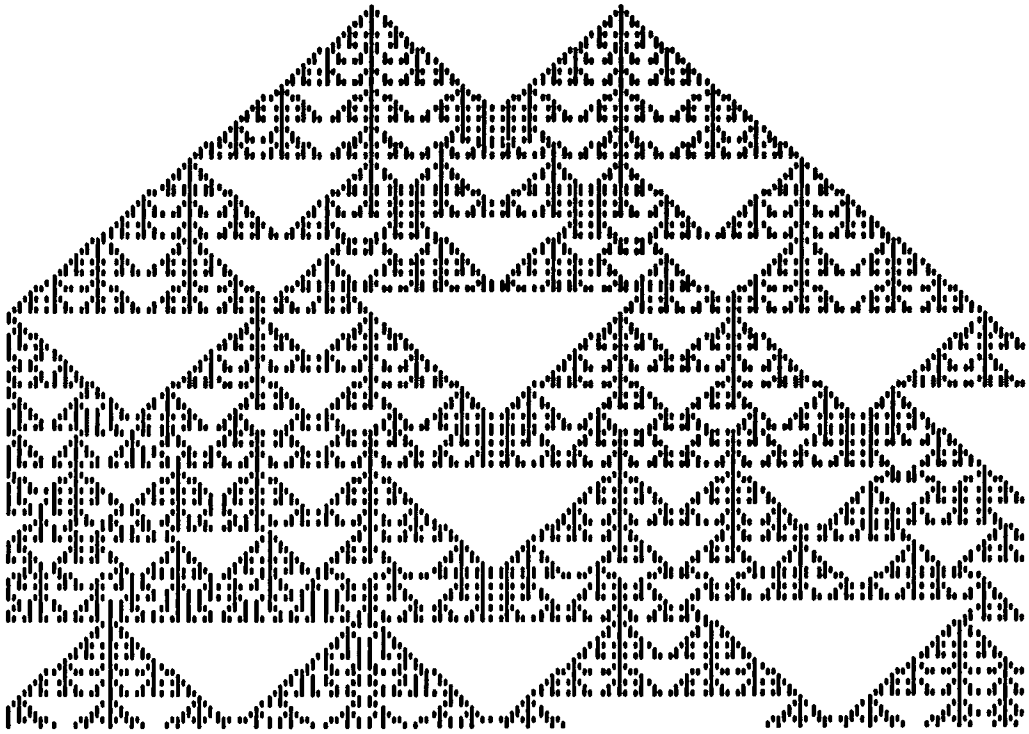


Figure 15

CCA Map coz(G) = {-1,0,1}
card(coz(f)) = 2

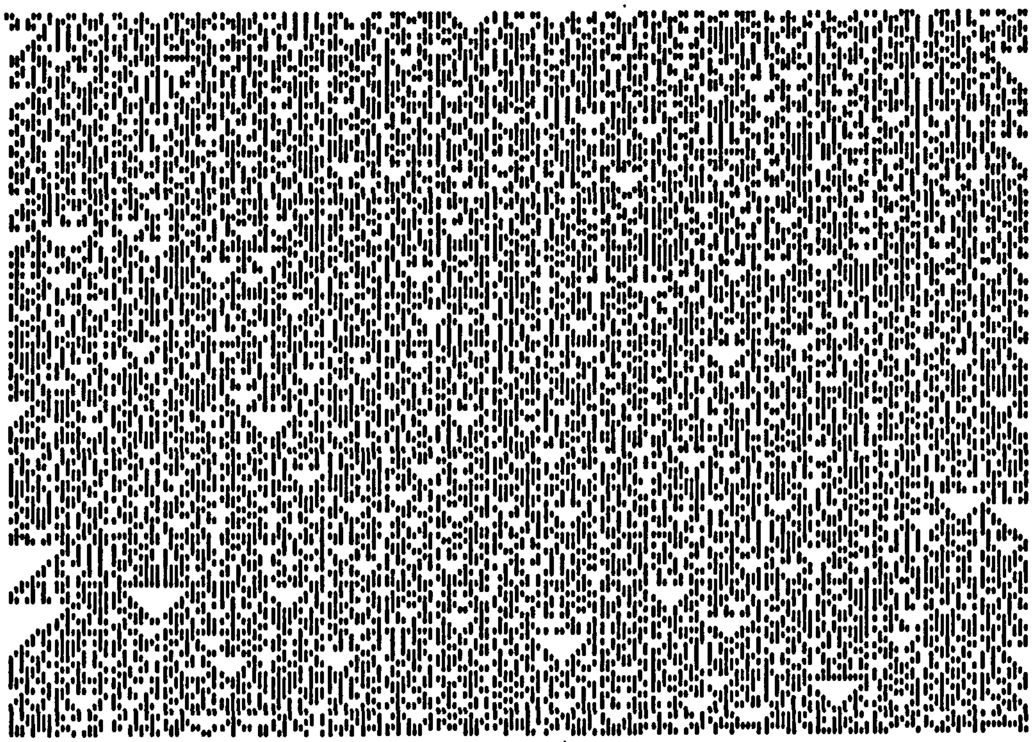


Figure 16

CCA Map coz(G) = {-1,0,1}
coz(f) is Chosen Randomly

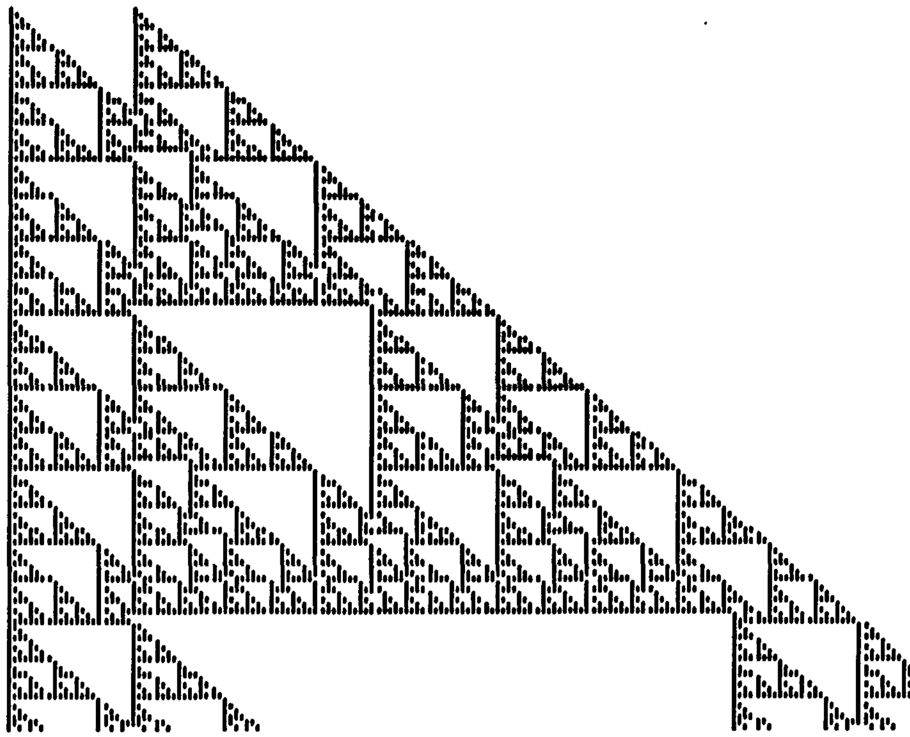


Figure 17

CCA Map coz(G) = {0,1}
card(coz(f)) = 2

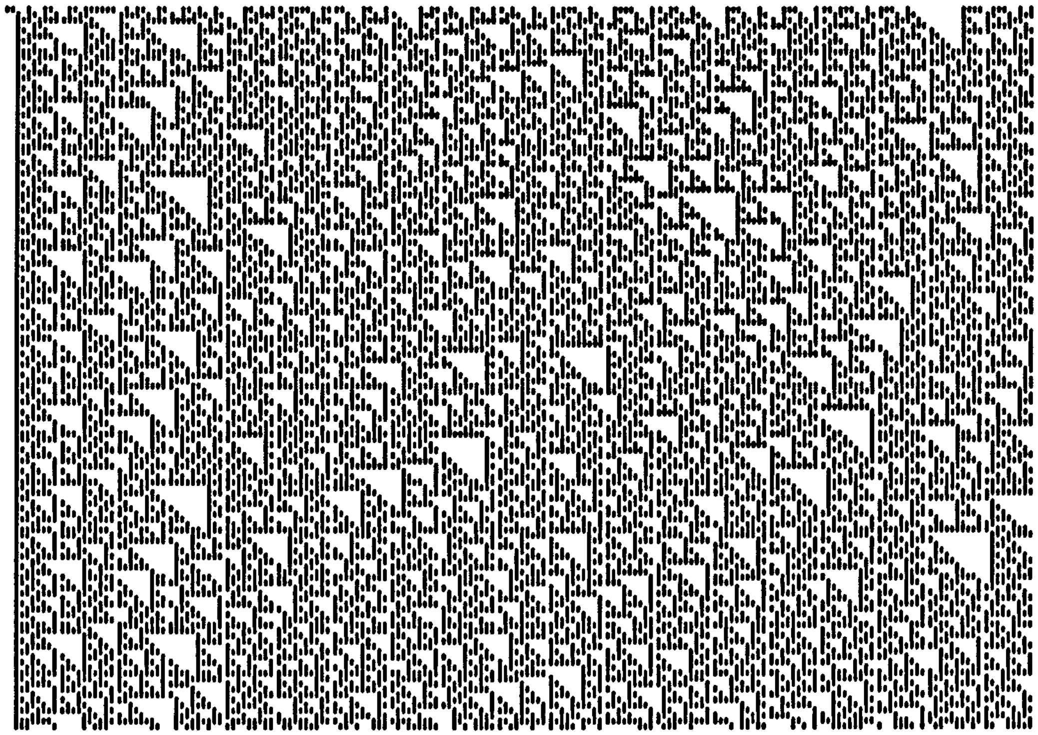


Figure 18

CCA Map coz(G) = {0,1}
coz(f) is Chosen Randomly

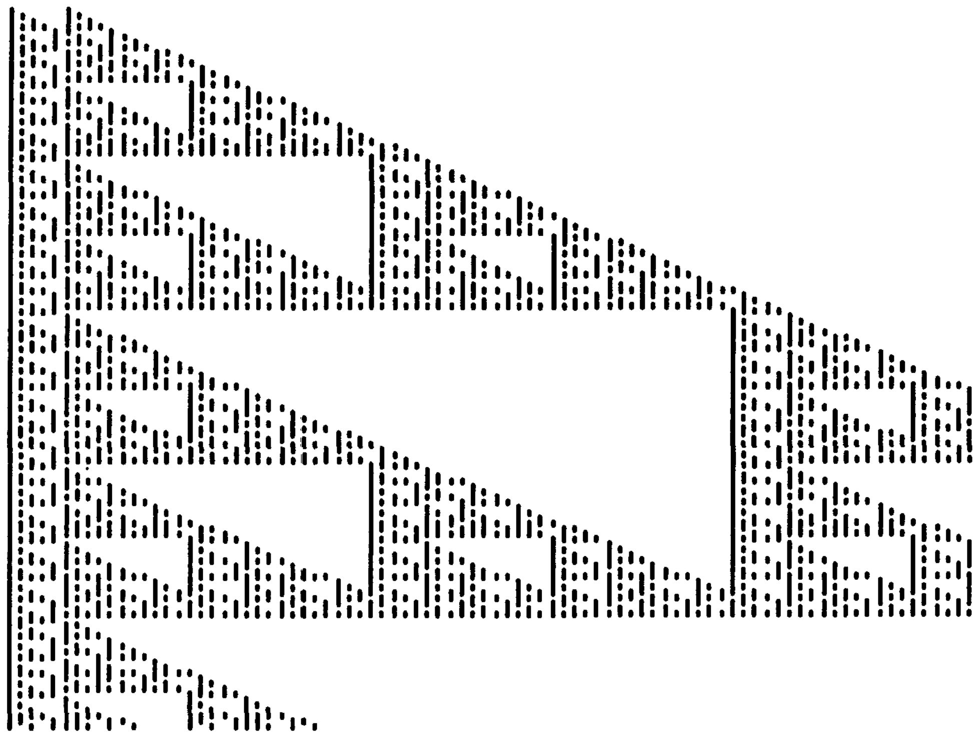


Figure 19

CCA Map coz(G) = {0,2}
card(coz(f)) = 2

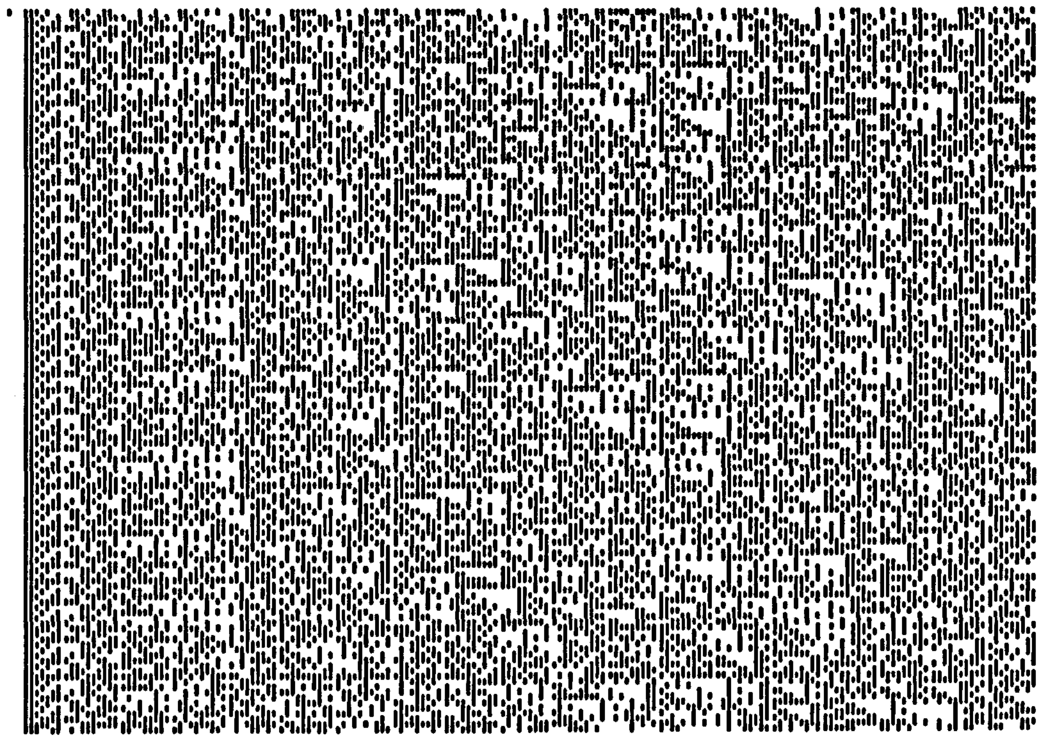


Figure 20

CCA Map coz(G) = {0,2}
coz(f) is Chosen Randomly

Another and much easier method for calculating convolutional cellular automata is to use the Z Transform.

Definition 7.12: For the digital signal f = $(\ldots a_{-2}\ a_{-1}\ a_0\ a_1\ a_2 \ldots)$ the Z Transform, denoted by Z(f), is the infinite series

$$\ldots + a_{-2}z^2 + a_{-1}z + a_0 + a_1z^{-1} + a_2z^{-2} + \ldots$$

$$= \sum_{n=-\infty}^{\infty} a_n z^{-n}$$

Example 7.6: f = $(\ldots 101010101 \ldots)$ where the alphabet = $\mathbb{Z}_2$

($\uparrow$ 0^{th})

then

$$Z(f) = \ldots + z^4 + z^2 + 1 + z^{-2} + z^{-4} + \ldots$$

$$= \sum_{n=-\infty}^{\infty} z^{-2n}$$

Note that the set of all formal infinite series utilizing integral powers of z with coefficients in S forms an R-module over the ring S and will be denoted by $S[z]$.

For every signal in S^Z there is a unique infinite series in $S[z]$ and for every series in $S[z]$ there is a unique signal in S^Z.

Hence, the map: $Z: S^Z \text{----}> S[z]$ is bijective.

Therefore, the inverse map $Z^{-1}: S[z] \text{----}> S^Z$ exists whereby the inverse mapping is defined as:

$$Z^{-1}\left(\sum_{n=-\infty}^{\infty} a_n z^{-n} \right) = (\ldots . a_{-2}\ a_{-1}\ a_0\ a_1\ a_2 \ldots .)$$

Note: $Z^{-1}(Z(f)) = f$

Properties of Z Transforms:

i.) Additive Property: $Z(f + y) = F(z) + Y(z)$

ii.) Homogeneous Property: $Z(af) = aF(z)$

iii.) Time Shifting Property: $Z(TRAN(f;n)) = z^{-n}F(z)$

These properties show that the two R-modules S^Z and $S[z]$ are isomorphic.

Convolution Theorem 7.1: Suppose $f \in S^Z$ and G is time limited. Then the Z transform Z(f) can formally be found and if Z(G) is the Z transform of G then

$$f * G = Z^{-1}[Z(f) \cdot Z(G)]$$

<u>Proof</u>: By the parallel convolution algorithm

$f * G = ADD[f_nG(n)]$, where

$f_nG(n) = SCALAR(TRAN(f;n), G(n))$

and by the above properties:

$$Z(f*G) = Z(ADD[f_nG(n)] = \sum_{n \in coz(G)} Z(f_nG(n)) = \sum_{n \in coz(G)} G(n) \cdot z^{-n} Z(f)$$

$$= Z(f) \sum_{n \in coz(G)} G(n) \cdot z^{-n} = Z(f) \cdot Z(g)$$

Hence $f * G = Z^{-1}[Z(f) \cdot Z(G)]$

<u>Examples 7.7</u>:

i.) $f = (11001)_0^0$ $G = (11)_0^0$ in the ring Z_2.

$$Z^{-1}[Z(f) \cdot Z(G)] = Z^{-1}[(1 + z^{-1} + z^{-4})(1 + z^{-1})]$$

$$= Z^{-1}(1 + 2z^{-1} + z^{-2} + z^{-4} + z^{-5})$$

$$= (1\ 0\ 1\ 0\ 1)_0^0$$

ii.) $f = (\ldots.a_{-1}\ a_0\ a_1\ a_2\ \ldots.)$ where $a_n \in Z_2$

$G = (1)_0^0 = \delta$ the identity operation

$$Z^{-1}[Z(f) \cdot Z(G)] = Z^{-1}[(\sum_{n=-\infty}^{\infty} a_n z^{-n})(1)]$$

$$= (\ldots.a_{-2}\ a_{-1}\ \underset{\substack{\uparrow \\ 0^{th}}}{a_0}\ a_1\ a_2.\ldots)$$

<u>Lemma 7.2</u>: For any signal f the CCA map $G = (1)^0_n$ is the right n-shift map.

<u>Proof</u>:

$$f = (a_i)_{i\epsilon Z} = (\ldots.a_{-2}\ a_{-1}\ \underset{\substack{\uparrow \\ 0^{\underline{th}}}}{a_0}\ a_1\ a_2\ldots.)$$

$$g = f*G = Z^{-1}(Z(f)\cdot Z(G)) = Z^{-1}[(\sum_{k=-\infty}^{\infty} a_k z^{-k})\cdot(z^{-n})]$$

$$= Z^{-1}[(z^{-n})(\ldots. + a_{-2}z^2 + a_{-1}z + a_0 + a_1z^{-1} + a_2z^{-2} + \ldots.)]$$

$$= Z^{-1}(\ldots+ a_{-n} +\ldots+ a_{-2}z^{2-n} + a_{-1}z^{1-n} + a_0z^{-n} + a_1z^{-1-n} + a_2z^{-2-n} +\ldots)$$

shifted n to right

$$= (\ldots.\ \underset{\substack{\uparrow \\ 0^{\underline{th}}}}{a_{-n}}\ \ldots.\ a_{-3}\ a_{-2}\ a_{-1}\ \underset{\substack{\uparrow \\ n^{\underline{th}}}}{a_0}\ a_1\ a_2\ \ldots.)$$

<u>Lemma 7.3</u>: For any signal f the CCA map $G = (1)^0_{-n}$ is the left n-shift map.

<u>Proof</u>:

$$f = (a_i)_{i\epsilon Z}$$

$$g = f*G = Z^{-1}(Z(f)\cdot Z(G)) = Z^{-1}[(\sum_{k=-\infty}^{\infty} a_k z^{-k})\cdot(z^{n})]$$

$$= Z^{-1}(\ldots+ a_{-2}z^{2+n} + a_{-1}z^{1+n} + a_0z^{-n} + a_1z^{-1+n} + a_2z^{-2+n} +\ldots+ a_n +\ldots)$$

shifted n to left

$$= (\ldots.\ a_{-2}\ a_{-1}\ \underset{\substack{\uparrow \\ -n^{\underline{th}}}}{a_0}\ a_1\ a_2\ a_3\ \ldots.\ \underset{\substack{\uparrow \\ 0^{\underline{th}}}}{a_n}\ \ldots.)$$

The Z Transform can be used to facilitate the computation of foward iterates of a convolutional cellular automata map. By the Convolution Theorem, for a time limited signal g and an arbitrary signal f,

$$g(f) = f * G = Z^{-1}(Z(f) \cdot Z(G))$$

Let $\xi = Z(f) \cdot Z(G)$ then $f * G = Z^{-1}(\xi)$

therefore $g^2(f) = (f*G)*G = Z^{-1}(Z[Z^{-1}(\xi)] \cdot Z(G))$

$$= Z^{-1}(\xi \cdot Z(G)) = Z^{-1}(Z(f) \cdot Z(G) \cdot Z(G)) = Z^{-1}(Z(G) \cdot Z^2(G))$$

and $f * g^3 = ((f*G)*G)*G = Z^{-1}(Z[Z^{-1}(Z[Z^{-1}(\xi)] \cdot Z(G))] \cdot Z(G))$

$$= Z^{-1}(Z(f) \cdot Z(G) \cdot Z(G) \cdot Z(G)) = Z^{-1}(Z(f) \cdot Z^3(G))$$

and, in general, the following iterative formula for any foward iterate i results:

$$g^i(f) = Z^{-1}(Z(f) \cdot Z^i(G))$$

The previous formula provides a direct method of computing any foward iterate of a convolutional cellular automata map.

<u>Example 7.8</u>: $f = (\ldots.101010101\ldots.)$ (with an arrow marking the 0^{th} position) $g = (110)^{0}_{-1}$

in the alphabet of $\mathbb{Z}_2$.

To compute the 2^{nd} iterate of the given CCA map:

$$g^2 = Z^{-1}(Z(f)\cdot Z^2(G))$$

$$Z(f) = \ldots.+ z^4 + z^2 + 1 + z^{-2} + z^{-4} +\ldots.$$

$$Z(G) = z + 1 \qquad Z^2(G) = z^2 + 2z + 1$$

$$Z(f)\cdot Z^2(G) = \ldots.+ 0z^6 + 0z^5 + 0z^4 + 0z^3 + 0z^2 + 0z + 0 + 0z^{-1} + 0z^{-2} + 0z^{-3} + 0z^{-4} +\ldots.$$

$$g^2 = Z^{-1}(Z(f)\cdot Z^2(G)) = (\ldots.00000000\ldots.) = (0)^{0}_{0}$$

The following section proposes a classification of convolution maps based on the classification presented in this dissertation. It will be shown that almost all (except for two maps) convolution maps belong to class Γ. The underlying reasoning is that, as seen from the parallel convolution algorithm, the convolution is a linear conbination of shift maps.

<u>Corollary 6.6.1</u>: If $G = \delta = (1)_0^0$, for $S = Z_2$, then $g \in A$.

<u>Proof</u>: It must be shown that there is a class $B_\varepsilon(f)^\circ \neq \emptyset$ for some $\varepsilon > 0$. That is $\exists$ an open ball, around f, which is contained in $B_\varepsilon(f)$.

Without loss of generality, assume $\lambda(0) = 1/2$ and $\lambda(1) = 1/2$.

Hence for $f = (1)_0^1$, $C_{1/8}(f) = \ldots * * * 1\ 1\ 1 * * * \ldots$

(here $*$ = 0 or 1)

Now, $B_{1/2}(f) = \{y | d(g^i(y), g^i(f)) \leq 1/2,\ i \in N_0\}$

That is, at least $(g^i(f))[0,0] = (g^i(y))[0,0]\ \forall\ i \in N_0$

and since $G = \delta$ is the idenity convolution map

$$C_{1/8}(f) \subset B_{1/2}(f)$$

and $B_{1/2}(f)^\circ \neq \emptyset$. Hence $g \in A$ by theorem 6.6.

<u>Corollary 6.6.2</u>: If $G = (0)_0^0$ for $S = \mathbb{Z}_2$, then $g \in A$.

<u>Proof</u>: Similiarly, without loss of generality, assume $\lambda(0) = 1/2$ and $\lambda(1) = 1/2$.

Hence for $f = (1)_0^1$, $C_{1/2}(f) = \ldots .* * * * 1 * * * * \ldots .$

```
                 0th
                  ↓
      f = ....1 1 1 1 1 1 1....
    f*G = ....0 0 0 0 0 0 0....
(f*G)*G = ....0 0 0 0 0 0 0....
      .           .
      .           .
      .           .
      .           .
```

the signal is $(0)_0^0$ for every foward iterate of f under g.

As seen, there is $C_{1/2}(f) \subset B_{1/2}(f)$. Hence $B_{1/2}(f)^\circ \neq \emptyset$ and by theorem 6.6 $g \in A$.

Definition 7.13: A convolution cellular automaton map, g, is **right permutive** if fixing all but the right most k^{th} entry of the signal f such that n - k = min(coz(G)) < 0 and allowing that entry to vary always gives a bijection from **S** to **S**.

Definition 7.14: A convolution cellular automaton map, g, is **left permutive** if fixing all but the left most k^{th} entry of the signal f such that n - k = max(coz(G)) > 0 and allowing that entry to vary always gives a bijection from **S** to **S**.

Definition 7.15: A convolution cellular automaton map, g, is said to be **doubly permutive** if it is both left and right permutive.

The following example shows the effect of a doubly permutive map on a binary signal:

Example 7.9: $f = (1\underset{\uparrow \atop 4^{th}}{0}1101)_{3}^{0}$ $\qquad G = (101)_{-1}^{0}$

$$\text{coz}(G) = \{-1,1\} \qquad \max(\text{coz}(G)) = 1$$

$$\min(\text{coz}(G)) = -1$$

$$g(5) = (f * G)(5) = \sum_{5-k\in \text{coz}(G)} f(k) \cdot G(5-k)$$

$$= \underset{k=4}{f(4)\cdot G(1)} + \underset{k=6}{f(6)\cdot G(-1)}$$

$$= 0\cdot 1 + 1\cdot 1 = 1$$

If $f = (11\underset{\substack{\uparrow \\ 4^{th}}}{1}101)_3^0$

$$g(5) = (f * G)(5) = f(4)\cdot G(1) + f(6)\cdot G(-1)$$
$$= 1\cdot 1 + 1\cdot 1 = 0$$

(Recall: for binary signals, arithmetic is performed using the binary ring operations in $\mathbb{Z}_2$ of addition and multiplication mod 2)

On the other hand, for $f = (1011\underset{\substack{\uparrow \\ 6^{th}}}{0}1)_3^0$ and $G = (101)_{-1}^0$

$$g(5) = (f * G)(5) = \sum_{5-k\in coz(G)} f(k) \cdot G(5-k)$$
$$= f(4)\cdot G(1) + f(6)\cdot G(-1)$$
$$= 0\cdot 1 + 1\cdot 1 = 1$$

If $f=(1010\underset{\substack{\uparrow \\ 6^{th}}}{0}1)_3^0$

$$g(5) = (f * G)(5) = f(4)\cdot G(1) + f(6)\cdot G(-1)$$
$$0\cdot 1 + 0\cdot 1 = 0$$

<u>Lemma 7.3</u>: For any signal f and non-zero CCA map g, taking values in $S \cong \mathbb{Z}_2$, the convolution cellular automaton map, g $\ni$ $\min(coz(G)) < 0$, is right permutive.

Proof: Construct the bijection between the the right most k^{th} entry of f for n - k = min(coz(G)).

Since g = f * G takes on values in $S \cong Z_2$ it is necessary to consider the following 2 cases:

case I.) suppose $\sum_{\substack{n-k \in coz(G) \\ n-k \neq min(coz(G))}} f(k) \cdot G(n-k) = 0$

$$\text{then} \quad g(n) = \begin{cases} 0 & \text{if } f(k)=0 \text{ for } n-k = \min(coz(G)) \\ 1 & \text{if } f(k)=1 \text{ for } n-k = \min(coz(G)) \end{cases}$$

case II.) suppose $\sum_{\substack{n-k \in coz(G) \\ n-k \neq min(coz(G))}} f(k) \cdot G(n-k) = 1$

$$\text{then} \quad g(n) = \begin{cases} 1 & \text{if } f(k)=0 \text{ for } n-k = \min(coz(G)) \\ 0 & \text{if } f(k)=1 \text{ for } n-k = \min(coz(G)) \end{cases}$$

Lemma 7.4: For any signal f and CCA map g, taking values in Z_2, the convolution cellular automaton map, g $\ni$ max(coz(G)) > 0, is left permutive.

Proof: Construct the bijection between the left most k^{th} entry of f for n - k = max(coz(G)). As in the previous lemma it is necessary to consider 2 cases:

case I: suppose $\sum_{\substack{n-k \in coz(G) \\ n-k \neq max(coz(G))}} f(k) \cdot G(n-k) = 0$

$$\text{then} \quad g(n) = \begin{cases} 0 & \text{if } f(k)=0 \text{ for } n-k = max(coz(G)) \\ 1 & \text{if } f(k)=1 \text{ for } n-k = max(coz(G)) \end{cases}$$

case II: suppose $\sum_{\substack{n-k \in coz(G) \\ n-k \neq max(coz(G))}} f(k) \cdot G(n-k) = 1$

$$\text{then} \quad g(n) = \begin{cases} 0 & \text{if } f(k)=1 \text{ for } n-k = max(coz(G)) \\ 1 & \text{if } f(k)=0 \text{ for } n-k = max(coz(G)) \end{cases}$$

Theorem 7.5: For any signal, taking values in Z_2, a convolution cellular automaton map g ∋ the local map G satisfies min(coz(G)) < 0 and max(coz(G)) > 0, is doubly permutive.

Proof: Directly follows from previous lemmas.

Corollary 7.5.1: For $Y = S^Z$, where $S = Z_2$, and g a convolution cellular automata map, if max(coz(G)) > 0 and min(coz(G)) < 0 then $g \in \Gamma$.

<u>Proof</u>: For any $f \in Y$ and $\varepsilon \leq (\min_{j \in S}\{\lambda(j)\})^{2n+1}$ for $1 \leq n \leq (r-1)/2$,

$$B_\varepsilon(f) = \{y | d(g^i(y), g^i(f)) \leq \varepsilon,\ i \in N_0\}$$

Consider the matrix $(a_{i,j})$ $i \in N_0$ and $j \in Z$ representation of the behavior of f. The $a_{i+1,j}$ element is computed by

(1) $$G(a_{i,j-r}, \ldots, a_{i,j+r})$$

Since g is doubly permutive there is at most one way to construct another signal that belongs to $B_\varepsilon(f)$ so that (1) holds. That is there is only onr way to construct a signal that will have the behavior of f on the infinite vertical jagged edge strip and signal will be identically f. Hence $B_\varepsilon(f) = \{f\}$ and $g \in \Gamma$.

<u>Corollary 7.5.2</u>: For $Y = S^Z$, where $S = Z_2$, and g a convolution cellular automata map, if $\max(coz(G)) > 0$ or $\min(coz(G)) < 0$ then $g \in \Gamma$.

<u>Proof</u>: For $f \in Y$, g is either right or left permutive. For $\varepsilon \leq (\min_{j \in S}\{\lambda(j)\})^{2n+1}$, where $n \leq (r-1)/2$, there is only one way to construct another signal belonging to $B_\varepsilon(f)$ that agrees with f on the right, or the left, depending on whether g is right or left permutive. Therefore, the set of signals y in $B_\varepsilon(f)$ must agree with f either on the right or left. Hence, for any probability distribution p on Z_2 and any $f \in Y$, $\mu(B_\varepsilon(f)) = 0$ and $g \in \Gamma$.

<u>Example 7.10</u>: For $G = (1100)^{0}_{-3}$ and $f = (1\ 0\ 1\ 1\ 1\ 1)^{0}_{0}$

```
                  0th
                  ↓
   f =      0 0 0 1 0 1 1 1 1 0 0 0 0 0 0....
 g (f)  =       1 0 0 0 0 1 0 0 0 0 0 0....
 g²(f)  =       0 0 1 1 0 0 0 0 0 0 0....
 g³(f)  =       0 1 0 0 0 0 0 0 0 0....
   .              .
   .              .
   .              .

                  0th
                  ↓
If   y     = 0 0 0 1 0 1 1 1 1 0 0 0 0 1 0 0 0 0....
   g (y)  =    1 0 0 0 0 1 0 0 0 1 1 0 0 0 0 0....
   g²(y)  =    0 0 1 1 0 0 1 0 1 0 0 0 0 0 0....
   g³(y)  =    0 1 0 1 1 1 1 0 0 0 0 0 0 0....
   g⁴(y)  =    1 0 0 0 1 0 0 0 0 0 0 0 0....
   g⁵(y)  =    0 1 1 0 0 0 0 0 0 0 0 0....
     .            .
     .            .
     .            .
```

Notice that y differs with f only in the 10^{th} place. However that single change propagates to the left of the entire signal.

Bibliography

[BS] S. Burris and H.P. Sankappanavar. *A Course in Universal Algebra.* NY: Springer-Verlag, 1981.

[CFS] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai. *Ergodic Theory.* NY: Springer-Verlag, 1982.

[CHY] K. Culik II, L.P. Hurd and S. Yu. Computation Theoretic Aspects of Cellular Automata. *Physica D.* 45 (1990) 357-378.

[D] R. Devaney. *Chaotic Dynamical Systems.* CA: Benjamin/Cummings Publishing Co.

[F] W. Feller. *An Introduction to Probability Theory and its Applications.* NY: John Wiley & Sons, 1968.

[GA] M. Gardner. "On cellular automata, self-reproduction, the Garden of Eden and the game 'life'," *Scientific American* 224(2), 112-117 (1971).

[CG] C. Giardina. *Parallel Digital Signal Processing.* NJ: Regency Publishing Co., 1991.

[G1] R. Gilman. Classes of Linear Automata. *Ergodic theory and dynamical systems.* (1987), **7**, 105-118.

[G2] R. Gilman. Periodic Behavior or Linear Automata. *Preprint.*

[KF] Kolmogorov and Fomin. *Functional Analysis* Vol. 1 Metric and Normed Spaces. NY: Graylock Press, 1957.

[PG] P. Grassberger. Chaos and Diffusion in Deterministic Cellular Automata. *Physica* **10 D** (1984) 52-58.

[PH] P. Halmos. *Measure Theory.* NY: Springer-Verlag, 1974.

[M] J. Munkres. *Topology.* NJ: Prentice-Hall, Inc., 1975.

[JM] J. Milnor. Notes on Surjective Cellular Automata-Maps. *Preprint.*

[ML] Saunders Mac Lane. *Categories for the Working Mathematician.* NY: Springer-Verlag, 1971.

[NBB] Narici, Beckenstein and Bachman. *Functional Analysis and Valuation Theory.* NY: Marcel Dekker, Inc., 1971

[N] I.P. Natanson. *Theory of Functions of a Real Variable.* NY: Frederick Unger Pub. Co. 1961.

[P] Karl Petersen. *Ergodic theory.* Great Britain: Cambridge University Press, 1983.

[PD] K. Preston and M. Duff. *Modern Cellular Automata.* NY: Plenum Press, 1984.

[W] P. Walters. *An Introduction to Ergodic Theory.* NY: Springer-Verlag, 1982.

[W1] S. Wolfram. Statistical Mechanics of Cellular Automata. *Reviews of Modern Physics.* Vol. 55, No. 3, July 1983, 601-644.

[W2] S. Wolfram. Universality and Complexity in Cellular Automata. *Physica* **10 D** (1984) 1-35.

[W3] S. Wolfram. Twenty Problems in the Theory of Cellular Automata. *Physica Scripta.* Vol. T9, (1985), 170-183.

[W4] S. Wolfram. Computation Theory of Cellular Automata. *Communications in Mathematical Physics.* 96, (1984), 15-57.