

THE UNIVERSITY OF MICHIGAN  
COLLEGE OF LITERATURE, SCIENCE, AND THE ARTS  
Computer and Communication Sciences Department

CELLULAR AUTOMATA AS MODELS  
OF NATURAL SYSTEMS

Andrew Gehret Barto

August 1975

THE UNIVERSITY OF MICHIGAN  
ENGINEERING LIBRARY

Technical Report No. 183

with assistance from:

National Science Foundation  
Grant No. DCR71-01997  
Washington, D.C.

en  
VM70268

ABSTRACT

CELLULAR AUTOMATA AS MODELS OF NATURAL SYSTEMS

by

Andrew Gehret Barto

Chairman: Bernard P. Zeigler

A view is taken of certain kinds of cellular automata that allows integration of some concepts and techniques from automata theory, linear systems theory, and the theory of partial differential equations. The emphasis is placed on the idea of modelling natural systems by by-passing continuous formulations. Elements of the abstract theory of harmonic analysis are used to relate solution techniques for linear partial differential equations with automata theoretic formulations of system simplification.

It is shown that classes of automaton-homomorphisms exist which can be given natural interpretations in the context of modeling distributed systems. Some of these homomorphisms justify usual modeling practices; some indicate where care must be taken in the model building process; and others suggest novel simplification techniques.

## ACKNOWLEDGEMENTS

I wish to thank the many people who have helped me during my doctoral studies. I am particularly indebted to Bernard P. Zeigler, the chairman of my doctoral committee, whose research has provided the framework in which this thesis was undertaken, and whose instruction, advice, and encouragement were crucial to its completion. I am indebted also to Arthur Burks, John Meyer, and Alan Wineman for their service on my doctoral committee and for their valuable participation in discussions of this research. I also thank Norman Foo who expressed an early interest in the problems considered here, and who has contributed in many ways to my education.

I am grateful to the Logic of Computers Group which provided excellent research facilities and generous financial support and to all its members for their help. I particularly want to thank John Holland and Richard Laing for many inspiring discussions and Phil Pilgrim whose graphics programs contributed a great deal to this thesis. I wish to express my deepest gratitude to my wife Monna, who not only typed the manuscript, but whose support and encouragement made this project possible.

The research contributing to this thesis was supported in part by National Science Foundation Grant No. DCR71-01997.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
CHAPTERS	
1. Introduction	1
2. Notational Preliminaries	
2.1 Introduction	10
2.2 Basic Notation	10
2.3 Linear Algebra	11
2.4 Group Theory	13
2.5 Automata Theory	16
3. Cellular Automata	
3.1 Introduction	22
3.2 Definition of a Cellular Automaton	23
3.3 Group Graphs	27
3.4 Connection with the Intuitive View	30
3.5 Automorphism Groups	32
3.6 Generalizations	34
4. Cellular Automata and Linearity	
4.1 Introduction	37
4.2 Linear Cellular Automata	37
4.3 Linear Networks and LCAs	42
4.4 LCAs and I/O Analysis of Linear Systems	47
4.5 Examples	49
5. Linear Cellular Automata and Harmonic Analysis	
5.1 Introduction	59
5.2 The Discrete Fourier Transform	61
5.3 The Convolution Theorem	66
5.4 Extension to Vector-valued Functions	67
5.5 Examples of the Discrete Fourier Transform and Convolution	70
5.6 Generalizations	78
5.7 Parallel Decomposition of Linear Cellular Automata	80
5.8 Examples of Parallel Decomposition	84
6. Homomorphisms of Linear Cellular Automata by Spatial Frequency Filtering	98
6.1 Introduction	98
6.2 Homomorphisms	99
6.3 Restrictions, Projections, and Embeddings	101
6.4 Distributed Realizations of Sets of LSMs	106

6.5	Homomorphisms of LCAs via Spatial Frequency Filtering	110
6.6	Complexity of Homomorphic Images	117
7.	Linear Cellular Automata and Sampling Theory	
7.1	Introduction	119
7.2	Homomorphisms between LCAs and having Different Structure Groups	120
7.3	Generalized Sampling Theory	123
7.4	Homomorphisms of LCAs via Filtering and Sampling	141
7.5	Homomorphisms of LCAs via Spatial Aggregation	154
7.6	Laminated LCAs	164
8.	Linear Cellular Automata as Models of Non-uniform Structures	
8.1	Introduction	169
8.2	Non-uniform Linear Cellular Automata	170
8.3	Harmonic Analysis of Non-uniform LCAs	174
8.4	Uniform Homomorphic Images of NULCAs	180
8.5	Homomorphisms of NULCAs produced by Filtering-Sampling and Aggregation	186
8.6	Discussion	196
9.	Conclusion	197
	REFERENCES	199

## Chapter 1

### Introduction

Many systems that occur in nature are traditionally viewed as very large collections of components distributed in space with interaction occurring between components that are near each other. Each component is usually considered to be essentially the same as every other and is often well understood when behaving in isolation. The usual way of modeling these kinds of systems is through partial differential equations (PDEs). Intuitively, this kind of model might be viewed as a network of an uncountably infinite number of components each interacting with infinitesimally close neighbors. When the components and their interactions are simple enough, many of these equations can be solved analytically using powerful results from functional analysis. However, for most equations one settles for computer generated approximations to specific solutions.

Models within this tradition are ubiquitous in the physical sciences and have been very successful in accounting for a wide variety of phenomena: the behavior of sound waves in air, probability waves in quantum mechanics, diffusion of heat or concentration of liquids and gases, diffusion of neutrons in atomic reactors, atmospheric or oceanographic flow and turbulence, and the entire range of electro-magnetic activity. In fact, a great proportion of scientific and technological progress can be directly linked to the success of modeling efforts based on the PDE modeling formalism. And there is little doubt that models formulated in this tradition will continue to play a significant role in contributing to our understanding of the natural world.

However, successful models tend to bias our perception of experience to the extent that the invention of other modes of explanation becomes less and less likely. Models and modeling formalisms that have been fruitful for long periods of time become regarded as necessities for the representation of experience and are sometimes confused with experience itself. Whether this is true in the case of PDEs is a question we cannot attempt to answer here, but it is against the background of these remarks that we discuss the cellular automaton formalism as an alternative means of formulating hypotheses about the operation of the real world.

The concept of a cellular automaton (also called a tessellation automaton or a homogeneous structure) originated with John von Neumann's design for a self-replicating machine. Although he planned to consider a model based on non-linear PDEs, only the conceptually simpler cellular model was completed (von Neumann [1966]). This modelling framework begins with a view of the components as actually separated in space and operating in discrete time. In fact, each component can be viewed as a computer - either a simple one as in John Conway's Game of Life which produces a 1 or 0 depending on how many neighbors are emitting 1s or 0s (Gardner [1970]), or a very complex one as in the computing elements of the ILLIAC IV computer.

More generally, a cellular automaton is an array of uniformly interconnected identical finite state automata. For an array of two dimensions, one can think of a finite state automaton situated at each point on a plane having integer valued coordinates. All of the automata are identical in their operating characteristics, and each receives input from automata having neighboring positions in the array.



Uniformity of these interconnections means that the network "looks the same" from any location: if an automaton is connected in a particular way to, say, its right-hand neighbor, then any other automaton is connected to its own right-hand neighbor in the same way. Each automaton operates in discrete time and in synchrony with its neighbors so that the state of each automaton in the array changes to a new state at the same time. An important point to understand is that the components of the network operate in parallel and without centralized control. At each time step, each automaton "reads" input from its neighbors and then computes a new state. The result is simultaneous computing activity over the entire network. It's customary to call each automaton in the array a cell and the state of the entire array a configuration of cell states.

This kind of conceptual framework has been elaborated and generalized and has provided a framework for modeling many kinds of systems (see Burks [1970], Yamada and Amoroso [1969,1971], or the recent survey by Aladyev [1974]). There is increasing interest in cellular automata as models of biological systems that exhibit a degree of spatial homogeneity such as certain areas of brain tissue (Mortimer [1970]), sections of heart muscle tissue (Foy [1974]), or colonies of bacteria (Goodman, Weinberg, and Laing [1971]).

These systems differ from those typically modeled by PDEs in several ways. First, they exhibit spatial homogeneity at a level higher than the "microscopic" homogeneity of a vibrating string or a heat conducting slab. Hence, the components are viewed as being actually discrete in space (sometimes with connections to fairly distant components). Secondly, the components are often assumed to

possess comparatively sophisticated computational power (e.g. the power of a muscle cell or neuron to process information). There are also important differences in the use made of the resulting models. While one may hope for a closed form solution with the PDE formulation, a cellular formulation is rarely expected to yield a "solution" in this classical sense. Computer simulation is usually the only means possible for gaining insight into a cellular automaton's behavior, and the simulation model is considered not as an approximation but as a realization of the cellular model. In fact, the formalism of a cellular automaton allows us to abandon the idea that modeling with discrete time and space necessarily represents an unfortunate and inaccurate discretization of "real" time and space. It is not necessary to feel that, all else being the same, a finer "mesh" would be better. Although many important things are lost when notions of continuous time and space are abandoned, other things are gained. We feel that in some modeling situations the use of PDEs presents unnecessary mathematical complications. This is especially true since often discrete approximations to the PDEs must be considered anyway. For some applications it may even be true that the great care taken in numerical approximation introduces an empty kind of precision into the problem since considerable abstraction may have been involved in the initial continuous formulation.

Thus, there are at least two approaches one can take in modeling what we have called distributed systems: the PDE approach and the cellular automaton approach. In some applications it is quite clear which kind of model is more appropriate. When processes appear to be continuous in space and time, the PDE formulation might be more

valuable. However, in many cases the choice is not clear and is usually decided by the researcher's background or by the tradition prevalent within a discipline. In particular, we feel that PDEs are often chosen because the modeller is unaware of or unfamiliar with other modeling techniques.

On the other hand, discrete formulations often appear to be disconnected from any rich body of lore such as that which exists for PDEs. They seem to be ad hoc constructions that neither benefit from well developed theory nor seem likely to contribute to it. This impression is only partly true. Cellular automata are related to a body of knowledge, but one that is motivated by an interest in these systems as models of computation. The more abstract of such studies deal with issues concerning the specification of cellular automata having universal computational power (e.g. Codd [1968], Thatcher [1964]), and the relationship between cellular automata and formal languages (e.g. Aladyev [1973]). Holland's iterative circuit computer (Holland [1959]) is the first of numerous uses of the cellular automaton framework to suggest forms of computer architecture capable of extensive parallel processing. Studies having more direct applications have resulted from the fact that logical circuits with uniform or iterative structure can be inexpensively constructed (see the survey by Minnick [1967]). However, investigations like these have historically had little direct relevance to issues which seem important in the context of modeling natural systems. PDEs, on the other hand, can provide useful formulations since enough structure is present to allow useful specification of behavioral questions.

Theoretical studies of cellular automata tend to be either about a single member of the class or about the class as a whole. The study of special single examples have established the existence of cellular automata with special capabilities (universal computational and constructional power) but provide nothing that can be generalized to larger classes of automata. On the other hand, studied in their full generality, cellular automata do not have enough formal structure to support results useful from a modeling point of view. Consequently, no results have been obtained concerning the behavioral possibilities of large but well structured classes of cellular automata. PDE's, however, can provide extremely useful formulations precisely because of the possibility of this kind of behavioral analysis.

Although it has been recognized that cellular automata are related to PDEs and numerical methods used for approximating their solutions (Aladyev [1974]), the PDE and cellular automaton traditions remain essentially disjoint, and no significant attempt has been made to investigate concepts of one tradition as they apply to the other. Several reasons for this lack of integration seem apparent. First, PDEs can be discussed in terms of well developed functional analytic theory so that they tend to be seen purely as structures within that theory. Second, systems typically modeled by cellular automata are so complex that formulation in terms of PDEs is very unnatural, and useful formulation in those terms is very difficult (if not impossible). A third reason for the conceptual separation of cellular automata and PDEs is simply that historically the ideas have come from different roots and the educational process has reflected this fact. We do not mean to imply that cellular automata and PDEs should be viewed as being

essentially the same. They clearly are not the same if only by virtue of the totally different kinds of questions that can be asked (and answered!) within each theory. They do, though, deal with the same sorts of phenomena, and some concepts and results from each approach can be interpreted in terms of the other.

In this thesis we attempt to provide one kind of link between these approaches so that a rigorous cross-fertilization of these very different modeling traditions can occur. In particular, we discuss a class of systems that can be formally related to both the PDE and the cellular automaton formalisms. This class of systems consists of cellular automata that have linear sequential machines as cells. These systems are related to linear PDEs in the same way that linear sequential machines are related to continuous time linear systems. Although this restriction limits behavioral possibilities, one can begin to deal with behavioral questions and arrive at results comparable to those that exist for linear PDEs but without the complexities of continuous mathematics. On the other hand, it becomes possible to examine concepts developed in automata theory as they are exemplified in this special class of systems and thereby make the corresponding connections with the continuous theory.

For purposes of modeling and simulation, perhaps the two most important links of this kind concern the relationship between known analytical means for solving certain PDEs and the conceptual framework, expressed in automata theoretic terms, which is developing for the process of model building and simulation (Zeigler [1972,1974]). In particular, we will discuss the separation of variables technique as it appears in a discrete situation and the automata theoretic technique

representing non-uniform structures as cellular automata. It is shown that in some instances spatial inhomogeneities can be ignored without producing misleading models of a system's operation.

Since our purpose is primarily to integrate aspects of the PDE and cellular automaton modeling formalisms, we make no attempt to present results in their most general form. It is felt that such abstraction tends to obscure the basic algebraic simplicity of the theory. We do, though, indicate directions of generalization that can be undertaken in the future.

## Chapter 2

### Notational Preliminaries

#### 2.1 Introduction

In this chapter we explain some of the notation and terminology used throughout this thesis. We have tried to use standard symbolism whenever possible, but since aspects of several formal traditions have been brought together, we have often found it necessary to make compromises. When each of several traditions has its own standard way of expressing the same ideas, we have tried to choose the most economical notation and the most suggestive terminology. Sec. 2.2 covers our usage of some basic mathematical notation. The other sections explain concepts from linear algebra, group theory, and automata theory which we will need later on. Notation that is very standard will be used with no explanation, and more specialized notation will be introduced as it is needed.

#### 2.2 Basic Notation

Throughout the text we will use the symbols  $\mathbb{R}$ ,  $\mathbb{C}$ , and  $\mathbb{Z}$  to denote the real numbers, complex numbers, and the integers respectively. The symbol  $\mathbb{R}$  will be reserved to identify an arbitrary ring. We will use  $\mathbb{Z}^+$  to indicate the set of non-negative integers, i.e.  $\{0, 1, \dots\}$ . For a set  $X$  with subset  $Y$ ,  $X - Y = \{x \mid x \in X \text{ and } x \notin Y\}$  is the complement of  $Y$  in  $X$ . The cardinality of a set  $X$  will be denoted  $|X|$ .

If  $\eta$  is a function from a set  $A$  into a set  $B$ , we will write  $\eta: A \rightarrow B$  and indicate its action on elements of  $A$  by  $\eta: a \mapsto b$  when  $a \in A$ ,  $b \in B$ , and  $\eta(a) = b$ . The image of  $A$  in  $B$  under  $\eta$  is the set

$\eta(A) = \{\eta(a) \mid a \in A\}$ . The function  $\eta$  is onto  $B$  when  $\eta(A) = B$  and one-one (or injective) when  $f(a) = f(a')$  implies  $a = a'$ . A one-one and onto function is called a bijection. We will write  $(\eta^{-1}(b))$  to indicate the set  $\{a \in A \mid \eta(a) = b\}$  whether  $\eta$  is injective or not.

Suppose  $A$ ,  $B$ , and  $C$  are sets,  $f:A \rightarrow B$ , and  $g:B \rightarrow C$ . The functions  $f$  and  $g$  can be composed to produce a function  $gf:A \rightarrow C$  defined by  $a \mapsto g(f(a))$  for all  $a \in A$ . Note the order used in writing this composition.

We will occasionally find it convenient to use the following simple terminology. Let  $Y$  be a subset of a set  $X$ . The map  $i:Y \rightarrow X$  given by  $i(y) = y$  for all  $y \in Y$ , is called the inclusion map of  $Y$  into  $X$ .

For two elements  $x$  and  $y$  of a group (or a ring or field), we will say that  $x$  divides  $y$  and write  $x \mid y$  when there is a positive integer  $d$  such that  $y = x+x+\dots+x$  with  $d$  terms on the right. We will also write  $y = x+x+\dots+x$  ( $d$  times).

The symbol  $\approx$  will be used in two ways: first, it indicates the isomorphism of various structures, and, secondly and less frequently, it indicates approximate equality. The context will make the meaning clear. Finally, we will abbreviate the phrase with respect to by writing simply w.r.t..

### 2.3 Linear Algebra

Let  $K$  be a field and let  $K^n$  denote the vector space of all  $n$ -tuples of elements of  $K$ . A typical vector will be written  $v = (v_0, \dots, v_{n-1})$ . Let  $\text{Mat}_K(n, m)$  denote the set of all  $n$  row  $\times$   $m$  column matrices over  $K$ . For a matrix  $A$ ,  $Av$  will denote the image of  $v$  where

$$(Av)_i = \sum_{j=0}^{m-1} a_{ij} v_j$$



where  $a_{ij}$  is the entry in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column of  $A$ . We will also write  $A = (a_{ij})$ . Recall that  $\text{Mat}_K(n,m)$  is also a vector space over  $K$ .

For any set  $X$ , if  $f: X \rightarrow K^n$  we define  $n$  scalar valued functions  $f_i: X \rightarrow K$ ,  $i=0, \dots, n-1$ , such that for all  $x \in X$

$$f(x) = (f_0(x), \dots, f_{n-1}(x)).$$

We will often find it necessary to define matrix valued functions of a set  $X$ . Suppose  $H: X \rightarrow \text{Mat}_K(n,m)$ . Then for every  $x \in X$ ,  $H(x)$  is a matrix which we will denote by  $(h_{ij}(x))$  where  $h_{ij}: X \rightarrow K$  for  $i=0, \dots, n-1$ ;  $j=0, \dots, m-1$ . We will use capital script letters to indicate matrix valued functions.

Almost all of the theory to follow deals exclusively with finite dimensional vector spaces. For some parts of the theory, however, we have chosen to use notation and terminology that is unconventional in the finite dimensional case. The reason for this is that we would like the notation to suggest extensions to the infinite dimensional theory that is the subject of functional analysis. To prevent this notation from giving rise to misconceptions, we will briefly describe its relationship to the more familiar symbolism of finite dimensional linear algebra.

Let  $V$  be a finite dimensional vector space over  $K$  with basis  $\{e_0, \dots, e_{n-1}\}$ . Then, as we have indicated, a vector  $v \in V$  is usually written  $v = (v_0, \dots, v_{n-1})$  when

$$v = \sum_{i=0}^{n-1} v_i e_i$$

with the  $v_i \in K$ . Another way to write this vector is as a function. Let

$I = \{0, \dots, n-1\}$ . Then  $\{e_0, \dots, e_{n-1}\} = \{e_i \mid i \in I\}$  and we can view a vector  $v \in V$  as a function of  $I$ , i.e.  $v: I \rightarrow K$  is given by  $v(i) = v_i$ .

This slight notational bias makes it easier to consider index sets that are more than just sets. In particular, the index set will often be a finite group (actually it will be the underlying set of a finite group, a distinction we will make clear below). Functional analysis deals with the structures arising when the index set  $I$  is the set  $\mathbb{R}$ . Needless to say, the uncountability of  $\mathbb{R}$  adds numerous subtleties to the theory.

## 2.4 Group Theory

We will use some elementary concepts from the theory of finite abelian groups that can be found in any introductory text on group theory (e.g. Rotman [1973]). Let  $G$  be a non-empty set. If there is a binary operation on  $G$  that associates to each pair of elements  $x, y$  in  $G$ , the element denoted by  $x+y$  in  $G$  such that:

- i) for all  $x, y, z \in G$ ,  $(x+y)+z = x+(y+z)$ ;
- ii) there is an element  $e \in G$  called the identity element such that  $e+x = x+e = x$  for all  $x \in G$ ; and
- iii) for every  $x \in G$ , there is an element  $-x$  of  $G$  called the inverse of  $x$  such that  $x+(-x) = (-x)+x = e$ ;

then  $G$  is said to be a group. We will consistently use the additive notation.

For conceptual clarity, it will sometimes be necessary to distinguish between a group  $G$  and the set of elements of  $G$ . We will refer to this set as the underlying set of  $G$  - it has the same elements as  $G$  but there is no rule of composition. Usually, however, the symbol  $G$  will denote both the group  $G$  and its underlying set.

A group is abelian or commutative if  $x+y = y+x$  for all  $x, y \in G$ . The order of  $G$ , denoted by  $|G|$ , is the number of elements in  $G$ . If the order of  $G$  is finite, we call  $G$  a finite group.

A non-empty subset  $D$  of (the underlying set of) a group  $G$  is a subgroup if the elements of  $D$  form a group under the operation defined for  $G$ . By Lagrange's theorem, if  $D$  is a subgroup of  $G$  and  $G$  is finite, then  $|D|$  divides  $|G|$ . The subgroup of  $G$  consisting of the identity element alone is called the trivial subgroup of  $G$ .

If  $X$  is a subset of  $G$  (not necessarily a subgroup), for every  $s \in G$  we will let  $s+X$  denote the set  $\{s+x \mid x \in X\}$ . Similarly,  $s-X = \{s-x \mid x \in X\}$  and should not be confused with our notation for set complement (Sec. 5.2) in which both symbols refer to sets. If  $X$  is a subgroup of  $G$ ,  $s+X$  is a (left) coset of  $X$  in  $G$ .

Let  $X$  be a subset of  $G$ . The smallest subgroup of  $G$  containing  $X$  is the subgroup generated by  $X$  denoted by  $\langle X \rangle$ . When  $X$  has only one element  $x$ ,  $\langle X \rangle$  is called the cyclic group generated by  $x$ . If this cyclic group has order  $N$ , then we call it the cyclic group of order  $N$  and denote it by  $Z_N$ . This is justified by the fact that any two cyclic groups of the same finite order  $N$  are isomorphic to each other and hence isomorphic to the group formed by the set of integers  $\{0, \dots, N-1\}$  with modulo  $N$  addition as the group operation.

If  $G$  and  $H$  are groups, then the set of all ordered pairs  $(g, h)$ ,  $g \in G$  and  $h \in H$ , forms a group under the operation

$$(g, h) + (g', h') = (g+g', h+h') .$$

This group is denoted  $G \oplus H$  and is called the (external) direct sum of  $G$  and  $H$ . For example,  $Z^d = Z \oplus \dots \oplus Z$  ( $d$  times) is the group of all  $d$ -tuples of integers with component-wise integer addition as the group

operation.

A mapping  $\eta: G \rightarrow H$  between the two groups  $G$  and  $H$  is a group-homomorphism if

$$\eta(g+g') = \eta(g) + \eta(g') \quad \text{for all } g, g' \in G.$$

The addition on the left is the operation of  $G$  and the one on the right is the operation of  $H$ . If  $\eta$  is also one-one and onto (a bijection) then it is a group-isomorphism and we'll write  $G \approx H$ .

If  $\eta: G \rightarrow H$  is a group-homomorphism, the set  $K = \{g \in G \mid \eta(g) = e\}$ , where  $e$  is the identity element of  $H$ , is called the kernel of  $\eta$  and is a subgroup of  $G$ . The set  $\eta(G)$  is a subgroup of  $H$ .

Our use of group theory does not go very far beyond the use of the elementary concepts described above. Our reason for using the idea of a group is that it precisely and compactly characterizes the intuitive idea of structural symmetry or uniformity. To see what group theory has to do with cellular automata, recall that we described a cellular automaton as a set of components, all identical, arranged so that each component is located at a point in space with integral coordinates. Moreover, the components are interconnected in a "uniform" way. This uniformity means that if the entire structure of the cellular automaton is shifted appropriately in space, then it looks exactly the same as it did before the shift. These shifts can be thought of as the elements of a group with the "sum" of two shifts given by the following: first do one shift, then, starting from the resulting position, do the other shift. It's easy to see that with this kind of "addition" the set of shifts does indeed form a group. In fact, for a cellular automaton having a component at each point in space with integer coordinates, the set of shifts that leave the structure unchanged is isomorphic to  $Z^3$

(for 3 dimensional space), i.e. the group consisting of the coordinates of the cells. This group is the group of symmetries of the cellular automaton's structure which we'll later call its structure group.

Now, as Holland [1960] first pointed out, other groups do as well as  $Z^3$  and describe other kinds of cellular automata. Some of these have non-abelian groups of symmetries consisting of various reflections and rotations in addition to simple shifts. Since our usage of these terms coincides with their usage in studies of crystals, it is quite literally true that cellular automata are crystals of automata. We will present these facts more precisely in Chapter 3. For more information on groups as models of real structures we refer the reader to Budden [1972] and Grossman and Magnus [1964].

In the chapters to follow, several sections are digressions into aspects of the theory of harmonic analysis on groups. In these sections, as in the current one, we will use  $G$  to denote a group. We will always use the symbol  $S$ , however, to denote a group that underlies the structure of a cellular automata. The intent is to suggest the words "space" or "structure".

## 2.5 Automata Theory

The concept of a finite automaton or, what is the same thing, a finite state sequential machine has been developed to provide a model of discrete, deterministic, computational processes requiring only finite memory. The theory presented in the following chapters is formulated entirely within an automata theoretic framework in order to provide a connection between models of computational processes and more classical models based on differential equations. We have, however, dropped the

finite memory restriction in order to include models closely related to these more traditional modeling formalisms. Accordingly, the structures that are studied in this thesis are examples of automata or sequential machines having state sets that are possibly infinite. All of the major results, however, also apply to a non-trivial class of finite automata.

We will use the terms automaton and sequential machine interchangeably to refer to a Moore type\* sequential machine as defined, for example, by Hartmanis and Sterns [1966], but with the finiteness restrictions removed. An automaton, therefore, is a quintuple  $M = \langle X, Q, Y, \delta, \lambda \rangle$  where

- i)  $X$  is a non-empty set of inputs;
- ii)  $Q$  is a non-empty set of states;
- iii)  $Y$  is a non-empty set of outputs;
- iv)  $\delta: Q \times X \rightarrow Q$  is the transition function; and
- v)  $\lambda: Q \rightarrow Y$  is the output function.

$M$  is a finite automaton when  $X$ ,  $Q$ , and  $Y$  are finite sets.

This mathematical structure is intended to model a discrete, deterministic process in the following sense: at any time  $t \in \mathbb{Z}^+$ , the process (machine) is entirely characterized by a state  $q \in Q$  so that if it receives an input  $x \in X$  at time  $t$ , then its state at time  $t+1$  is given by  $\delta(q, x)$ . Its output at time  $t+1$  is  $\lambda(\delta(q, x))$ . Given a starting state and a sequence of inputs, by successively applying the transition function, it is possible to generate the sequence of states through which the process passes. Similarly, it is possible to generate the output sequence.

---

\*As opposed to a Mealy type sequential machine in which  $\lambda$  is a function of  $Q \times X$  instead of just  $Q$ . See Hartmanis and Sterns [1966].

One of the crucial problems faced in applying this model to any actual process is in deciding what aspects of the process are to be taken as states. The choice is not arbitrary since, according to the definition of an automaton, a state must be rich enough so that it can be used, together with input, to predict the next state. Of course another problem arises since it's not always obvious what is appropriately considered input to the system. Keeping these problems in mind, the reader can think of the prototypical example of modeling a digital computer as a (finite) automaton: a state at any time of its operation can be taken to be the contents of its entire memory including such things as data registers and instruction location counter. See Minsky [1967] for a careful development of the concept of state.

In the next chapter we define a cellular automaton as a special kind of autonomous automaton or state-transition system. An autonomous automaton is simply a pair  $M = \langle Q, \delta \rangle$  where  $Q$  is a non-empty set of states and  $\delta: Q \rightarrow Q$  is the transition function. This structure represents a process in which there is no input and in which the output at time  $t$  can be considered to be the state at time  $t$ , i.e.  $\lambda(q) = q$ . If at time 0,  $M$  is in state  $q$ , we will write

$$q_t = \delta^t(q) = \delta(\delta(\dots\delta(q)))$$

(where  $\delta$  is applied  $t$  times) to indicate the state of  $M$  at time  $t \in \mathbb{Z}^+$ .

Although most of the theory we develop can be extended without difficulty to the case having input and output, we will stay almost entirely within the framework of autonomous sequential machines.\*

---

\*The exception to this is that the cells of a cellular automaton are sequential machines with input and output (from and to their neighbors). Viewed as a whole, however, a cellular automaton (as we will define one) is autonomous.

While this restriction makes it impossible to talk about such important ideas as controllability, observability, and minimality (see, for example, Kalman, Falb, and Arbib [1969]), it removes a great deal of notational complexity. In addition, our desire is simply to provide points of contact between automata theory and other formal traditions, and we feel that for this purpose a study of the input-free case suffices. In the conclusion we indicate the appropriate directions of generalization and mention some results that can be obtained.

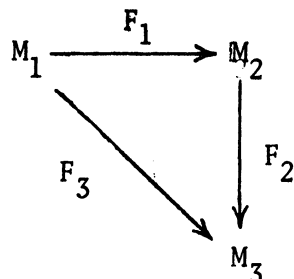
The automata theoretic concepts to be described below are therefore given for the autonomous case only. The complete definitions can be found in Hartmanis and Stearns [1966]. When it is clear that the term automaton (or sequential machine) refers to a pair instead of a quintuple, we will drop the word autonomous.

Let  $M_i = \langle Q_i, \delta_i \rangle$ ,  $i=1,2$  be two sequential machines.  $M_2$  is a subautomaton or submachine of  $M_1$  if  $Q_2 \subseteq Q_1$  and  $\delta_2(q) = \delta_1(q)$  for all  $q \in Q_2$ . A map  $F: Q_1 \rightarrow Q_2$  that is onto  $Q_2$  is an automaton-homomorphism if  $F(\delta_1(q)) = \delta_2(F(q))$  for all  $q \in Q_1$ . We will say that  $M_2$  is a homomorphic image of  $M_1$  and write  $M_1 \xrightarrow{F} M_2$  or just  $M_1 \rightarrow M_2$ . If  $F$  is also one-one, then it is an automaton-isomorphism and we'll write  $M_1 \xrightleftharpoons{F} M_2$ ,  $M_1 \xrightleftharpoons[F^{-1]} M_2$ , or just  $M_1 \approx M_2$ . When there is no danger of confusion between homomorphisms (isomorphisms) of groups and automata, we will omit the identifying prefix.

A fact that will play an important role in the proofs of many theorems is that when automaton-homomorphisms are composed, the result is also an automaton-homomorphism, i.e. if  $M_i = \langle Q_i, \delta_i \rangle$ ,  $i=1,2,3$ , are automata, with  $M_1 \xrightarrow{F_1} M_2$ , and  $M_2 \xrightarrow{F_2} M_3$ , then  $M_1 \xrightarrow{F_1 F_2} M_3$ . If both  $F_1$  and  $F_2$  are isomorphisms, then so is  $F_2 F_1$ . It will often be convenient



to use the standard commutative diagram conventions to express facts of this kind. For example, the diagram below conveys all of these facts: i)  $F_1$  is a homomorphism from  $M_1$  to  $M_2$ ; ii)  $F_2$  is a homomorphism from  $M_2$  to  $M_3$ ; iii)  $F_3$  is a homomorphism from  $M_1$  to  $M_3$  and  $F_3 = F_2F_1$ .



Suppose  $M_1$  and  $M_2$  are two automata and  $M'_1$  is a submachine of  $M_1$ . If  $M'_1 \xrightarrow{F} M_2$ , we say that  $M_1$  realizes  $M_2$ . The map  $F^{-1}$  is called an assignment of  $M_2$  into  $M_1$  (c.f. Hartmanis and Sterns [1966]). Intuitively,  $M_1$  can do everything  $M_2$  can do and more. For any starting state  $q$  of  $M_2$  there is a starting state  $F^{-1}(q)$  of  $M'_1$ , such that the behaviors of  $M_1$  and  $M_2$  will be the same (with relabelling by  $F$ ) for all time. When  $M_1$  is started in a state not in the state set of  $M'_1$ , it may behave in a way entirely unrelated to any possible behavior of  $M_2$ . This concept is intended to model the situation which occurs when it's necessary to construct ("realize") a device (e.g. logical circuit) that behaves like  $M_2$  but its simpler for some reason to build one that behaves like  $M_1$ . If  $M_1$  realizes  $M_2$ , building the simpler device suffices.

Finally, we define a subclass of sequential machines which provides other points of contact between different modeling traditions. This class consists of sequential machines whose transition functions and output functions can be represented by linear operators. Although we will mostly need to consider only the autonomous case, we start with the definition from Harrison [1969] which includes input and output.

A (Moore type) linear sequential machine (or LSM) is a sequential machine  $M = \langle X, Q, Y, \delta, \lambda \rangle$  such that there is a field  $K$  and integers  $k, n, \ell \in \mathbb{Z}^+$  such that  $X = K^k$ ,  $Q = K^n$ ,  $Y = K^\ell$ , and such that there are matrices  $A$ ,  $B$ , and  $C$  over  $K$  (of appropriate sizes) so that  $\delta$  and  $\lambda$  can be given as follows:

$$\delta(q, x) = Aq + Bx$$

$$\lambda(q) = Cq \quad \text{for all } q \in Q \text{ and } x \in X.$$

Similarly, an autonomous LSM is a special kind of autonomous sequential machine and corresponds to an LSM with  $k = \ell = 0$ . We will denote an autonomous LSM by writing  $M = \langle K^n, A \rangle$ . When  $K$  is a finite field, an LSM is a special kind of finite automaton.

The concept of an LSM represents a formulation in the language of sequential machines of a class of models that goes by several different names. In the more mathematically oriented literature, e.g. the literature of numerical analysis, LSMs are called linear difference equations. From the systems theory point of view, they are discrete time, time invariant, finite dimensional linear systems (e.g. Kalman, Falb, and Arbib [1969]). In the more applications oriented areas involving signal processing, they are sometimes known as recursive digital filters (see Gold and Rader [1969]).

## Chapter 3

### Cellular Automata

#### 3.1 Introduction

The idea of a cellular automaton was first developed by John von Neumann, at the suggestion of S. Ulam, as the framework for a demonstration that no logical problems were involved in the idea that a "machine" could, at least in principle, reproduce itself. He began with the description of a structure (which most people could comfortably call a machine) that would be able to construct a copy of itself. This logical construction was left unfinished at his death but was completed by A. W. Burks and published as the Theory of Self-reproducing Automata [1966]. The self-reproducing structure was imbedded in a particular two-dimensional example of what von Neumann called a cellular automaton. He defined the general notion of a cellular automaton as a uniform interconnection of identical finite state automata and viewed the theory of these networks as part of a more general theory of automata.

From the beginning, then, cellular automata have been thought of in terms of the discrete mathematics of automata theory. Although von Neumann hoped to develop a continuous model of self-reproduction formalized by non-linear PDEs of the sort used to model fluid dynamics, only the discrete cellular automaton version was completed (von Neumann [1966]). However, investigations inspired by von Neumann's construction continued in the tradition of automata theory where the technical context is largely that of digital computation: computational and constructional universality, formal grammar theory, and analysis and synthesis of cellular hardware. Aladyev [1974] surveys a broad range of these studies. Minnick [1967] surveys the more applications oriented work on cellular logic.

The emphasis in this thesis, however, is on the view of cellular automata that is more closely allied with continuous mathematics. We give a formal definition of a cellular automaton that uses terminology and the conceptual view of functional analysis but is, nevertheless, essentially the same as the definition familiar to automata theorists. We then use this framework to indicate how the idea of a cellular automaton can be altered to make contact with various non-automata theoretic formalisms.

### 3.2 Definition of a Cellular Automaton

Before a formal definition of a cellular automaton is given, some notational preliminaries are required since we take a functional analytic point of view that is unusual in discussions of cellular automata. Instead of talking about  $n$ -tuples and  $d$ -tuples of  $n$ -tuples etc., we'll talk about functions and operators on function spaces. The purpose of using this terminology and the accompanying notation is to indicate connections to traditional mathematical structures that are obscure within the usual automata theoretic framework.

Let  $S$  be a countable set and let  $V$  be a set with a distinguished element denoted by  $0$ . The support of a function  $f: S \rightarrow V$  is the set  $X \subseteq S$  such that  $f(s) \neq 0$  if and only if  $s \in X$ . A function  $f: S \rightarrow V$  has finite support if its support is a finite set. Let  $F(S, V)$  denote the set all functions from  $S$  to  $V$  that have finite support. For any subset  $X$  of  $S$  and any  $f \in F(S, V)$ , let  $f|X$  denote the restriction of  $f$  to  $X$ , i.e.  $f|X : X \rightarrow V$  is given by  $(f|X)(s) = f(s)$  for  $s \in X$ .

Now suppose  $S$  is a group under an operation denoted by the symbol '+'. For any  $X \subseteq S$  and any  $s \in S$ , let  $s+X = \{s+x | x \in X\}$  and  $s-X = \{s-x | x \in X\}$ .

For every  $s \in S$ , it is convenient to define a shift operator  $L_s$  on the set  $F(S, V)$  such that  $(L_s f)(r) = f(s+r)$  for every  $f \in F(S, V)$  and  $r \in S$ . Under operator composition the set of shift operators forms a group isomorphic to  $S$  via the map that takes  $s$  into  $L_{-s}$ . This is true since  $L_s L_t = L_{t+s}$  or  $L_{-s} L_{-t} = L_{-t-s} = L_{-(s+t)}$  (as can be verified using the definition of function composition). In case  $S$  is abelian,  $S$  and the set of shifts are also isomorphic via the map  $s \mapsto L_s$ .<sup>\*</sup> Ignoring for a moment our countability restriction of  $S$ , suppose  $S = V = \mathbb{R}$  and '+' is the usual real addition. Fig. 3.1 indicates the action of a particular  $L_s$ .

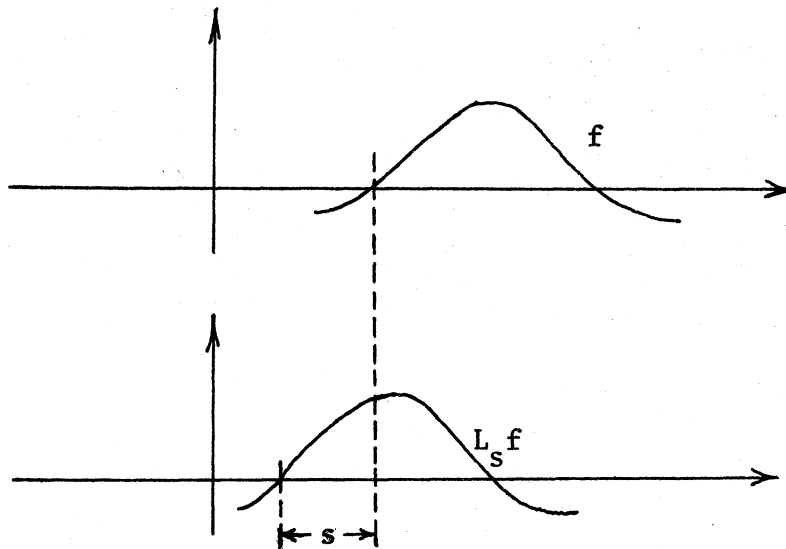


Figure 3.1: The effect of a shift operator  $L_s$ ,  $s > 0$ .

---

<sup>\*</sup>If we defined a shift operator so that  $(L_s f)(r) = f(r+s)$ , then the map given by  $s \mapsto L_s$  would be an isomorphism for both abelian and non-abelian groups. The definition we have chosen, however, offers certain notational conveniences later on.

We define a cellular automaton as a special kind of sequential machine (see Chapter 2). Although it is possible to consider cellular automata with input and output, for simplicity only the autonomous case will be considered here. In this case a cellular automaton is a special kind of autonomous sequential machine (state-transition system)  $M = \langle Q, \delta \rangle$  where  $\delta: Q \rightarrow Q$ .

### Definition 3.1

A cellular automaton is an autonomous sequential machine  $M = \langle Q, \delta \rangle$  with the following special properties:

- i) There is a countable group  $S$  (additive) called the structure group and a set  $V$  called the cell state set with a distinguished element  $0$  called the quiescent state such that  $Q = F(S, V)$ . Elements of  $Q$  are called configurations; and for every  $q \in Q$  and every  $s \in S$ ,  $q(s)$  is called the state of cell  $s$  in configuration  $q$ .
- ii) There is a finite subset  $X$  of  $S$  called the neighborhood template of  $M$ . The set  $s+X$  is the neighborhood of cell  $s$ .
- iii) There is a mapping  $\sigma: Q \rightarrow V$  such that if  $q \in Q$  is identically  $0$ , then  $\sigma q = 0$ ; and such that for  $q_1, q_2 \in Q$  with  $q_1|_X = q_2|_X$ ,  $\sigma q_1 = \sigma q_2$ . The map  $\sigma$  is called the local transition function of  $M$ .
- iv)  $\delta: Q \rightarrow Q$  is called the global transition function of  $M$  and is defined as follows:  $\delta(q) = q'$  where  $q'$  is given by  $q'(s) = \sigma L_s q$  for all  $s \in S$ .

For convenience we will denote the cellular automata  $M$  by  $M = \langle S, V, X, \sigma \rangle$  but also by  $M = \langle Q, \delta \rangle$  when it is to be considered as an unstructured sequential machine.

In order for this definition to make sense it must be true that the

global transition function  $\delta$  does indeed map  $Q$  into  $Q$ , i.e. it must be true that if  $q$  has finite support, then so does  $\delta(q)$ . This is true as the following argument shows. Suppose  $q \in Q$  has finite support  $Y$ . We'll show that the support of  $\delta(q)$  is a subset of  $A = \bigcup_{y \in Y} (y-X)$  and therefore, being contained in a finite union of finite sets, is also finite. To see this, suppose  $s \in S$  but  $s \notin A$ . This means that for all  $x \in X$  and all  $y \in Y$ ,  $s \neq y-x$ ; or, for all  $x \in X$ , that  $s+x \notin Y$ . Thus for any  $x \in X$ ,  $(L_s q)(x) = q(s+x) = 0$  since the support of  $q$  is  $Y$ . If  $f \in Q$  is the configuration identically equal to zero, we have that  $(L_s q)|_X = f|_X$  so that by part (iii) of the definition  $\sigma L_s q = \sigma f = 0$ . Thus  $(\delta(q))(s) = \sigma L_s q = 0$ , and we've shown that the support of  $\delta(q)$  is contained in  $A$  and is therefore finite.

Our definition of a cellular automaton, then, is internally consistent; but is it consistent with the various other forms of the definition that are found in the literature of cellular automata? Although the functional notation we use perhaps obscures the fact, our concept of a cellular automaton is essentially identical to that used elsewhere in the literature with the exceptions that we have allowed the structure group  $S$  to be any countable group instead of just the group  $Z^d$ , and we have allowed the cell state set  $V$  to be infinite. Our more general definition is given for two reasons: first, allowing the structure group to be any countable group lets us explore the relation of these structures to classical symmetrical structures which we do in the next section, and second, we want to include cases where  $V$  is a finite-dimensional vector space - also to include more classical structures (Chapter 4).

Except for these two differences, we have defined what has been called an algorithmic cellular automaton (Burks [1970]). Even though the array of cells may be infinite, if all but a finite number of

cells are quiescent initially, then all but a finite number of cells will be quiescent at any time during the cellular automaton's operation. This property is the same as requiring the state set  $Q$  to consist only of configurations with finite support and showing, as we have done, that  $Q$  is closed under the action of the global transition function  $\delta$ . The least standard part of our definition is (iii), the description of the local transition function  $\sigma$ . The conditions on  $\sigma$  given there imply that 1) if a cell's neighborhood contains only quiescent cells, then its next state will be quiescent, and 2) that  $\sigma$  is indeed local since its action does not depend on a function's values outside of the finite set  $X$ . The specification of the global transition function  $\delta$  that we give in part (iv) differs from that given in other definitions (e.g. Yamada and Amoroso [1969]) only because we use the notion of a shift operator. A cell's next state is obtained by viewing it as located at the origin (the identity of  $S$ ) and operating with  $\sigma$ . Thus, at least for  $S = Z^d$ , our cellular automata have the usual kind of uniformity. In the next section we discuss the meaning of uniformity with respect to an arbitrary countable group.

### 3.3 Group Graphs

As we have indicated, the concept of structural uniformity is usually defined with respect to  $Z^d$  as in Aladyev [1974]. However, the more basic notion is structural symmetry which group theory precisely characterizes as mentioned in Chapter 2. Def. 3.1 allows the structure group  $S$  to be finite or even non-abelian in order to place structural uniformity in this more basic context.

It is necessary, however, to show that the idea of a cellular automaton with a non abelian and/or finite structure group makes intuitive



sense. It was shown by Wagner [1966] that uniform interconnection of components can be characterized by group graphs based on any kind of group. The integral lattice and neighborhood template of a cellular automaton as it is usually defined can be regarded as a group graph based on  $Z^d$ .

A group graph (slightly different from the graph of a group) is a graph whose vertices correspond to and are labeled with the elements of a countable group  $S$ , and for which there is a finite subset  $X$  of  $S$  such that there is an edge between vertices corresponding to  $s$  and  $r$  if and only if there is an  $x \in X$  such that  $r = s+x$ . We will say the graph is a group graph of  $S$  w.r.t  $X$ . This graph is completely connected if and only if  $X$  generates  $S$ . (The set  $X$  will correspond to the neighborhood template called  $X$  in Def. 3.1.) Wagner has shown that this concept precisely characterizes that we mean when we say that a network "looks the same" from each component. Such a network corresponds to a group graph and every group graph gives rise to such a network.

### Example 3.1

Consider, for example, the group  $S_3$  of all the permutations of 3 objects. This group is non-abelian and has six elements, say  $a, b, c, d, e$ , and  $f$ ; and the operation given by this table:

+	a	b	c	d	e	f
a	a	b	c	d	e	f
b	b	a	f	e	d	c
c	c	e	a	f	b	d
d	d	f	e	a	c	b
e	e	c	d	b	f	a
f	f	d	b	c	a	e

Let  $X = \{b, e\}$ . The group graph of  $S_3$  w.r.t.  $X$  is shown in Fig. 3.2.

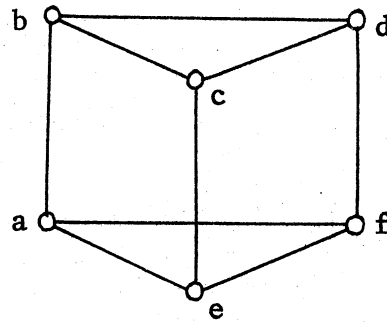


Figure 3.2: The group graph of  $S_3$  w.r.t.  $X = \{b, e\}$ .

The crystal-like structure in Fig. 3.2 could very well underlie a cellular automaton by our definition. The shift operators would perform shifts of configurations according to the operation of the group  $S_3$ . Fig. 3.3 illustrates the action of  $L_b$ .

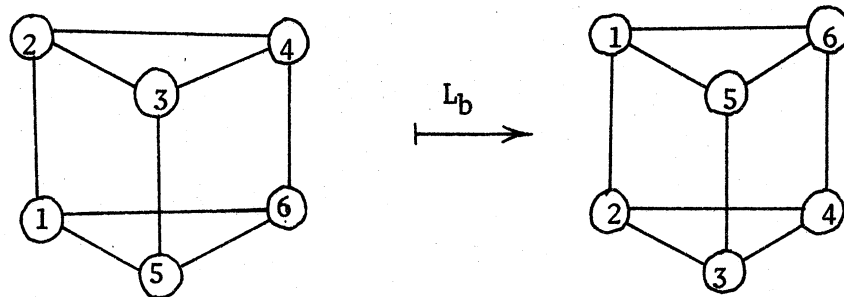


Figure 3.3: The action of a shift operator on a function of  $S_3$ .

Intuitively, a translation with respect to a non-abelian group requires a rotation in addition to a shift. This is more easily seen by studying the non-abelian tessellation of the entire plane shown in Fig. 3.4. We have indicated two sets,  $s_1+X$  and  $s_2+X$ , that would correspond to the neighborhoods of cells  $s_1$  and  $s_2$  respectively if the graph

in Fig. 3.4 were associated with a cellular automaton.

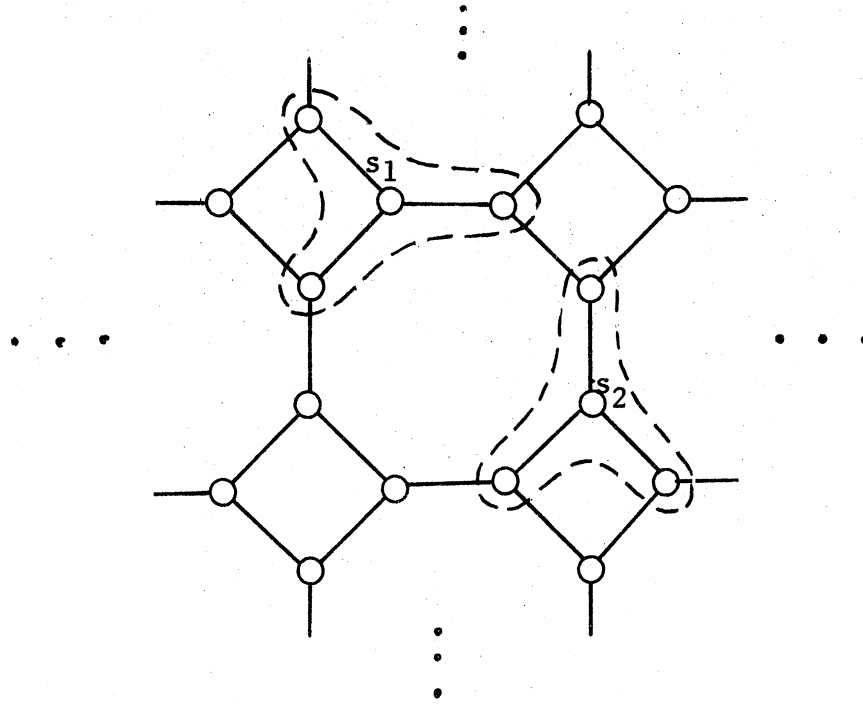


Figure 3.4: A non-abelian tessellation of the plane.

The reader is referred to Grossman and Magnus [1964] for an elementary presentation of these concepts. In what follows, however, we will deal almost exclusively with abelian groups.

### 3.4 Connection with the Intuitive View

We can now clearly connect Def. 3.1 and the intuitive notion of a cellular automaton given in Chapter 1. The network of interconnected components associated with the cellular automaton  $M = \langle S, V, X, \sigma \rangle$  would be based on the group graph of  $S$  w.r.t.  $X$ . On each node of the graph would be located a copy of the sequential machine  $M' = \langle Q', I', S' \rangle$  with state set  $Q' = V$ , input set  $I'$ , and transition function  $\delta'$  given according to which of the following two cases holds:

case i)  $e \notin X$ , where  $e$  is the identity element of  $S$ .

Let  $X = \{x_0, \dots, x_{n-1}\}$ . Then the input set  $I' = V^n$ . For any  $v \in V$  and  $(v_0, \dots, v_{n-1}) \in V^n$ , let  $q$  be any configuration in  $Q$  with  $q(x_i) = v_i, i=0, \dots, n-1$ . Then  $\delta'(v, (v_0, \dots, v_{n-1})) = \sigma q$ .

case ii)  $e \in X$ . Say  $x_0 = e$ . Then  $I' = V^{n-1}$ , and for any  $q \in Q$  as in case i, we have  $\delta'(v_0, (v_1, \dots, v_{n-1})) = \sigma q$ .

It follows from the definition of  $\sigma$  that in either case  $\delta'$  is well defined. The details of this specification are somewhat awkward given our definition but not complicated. Case i corresponds to the situation in which cells are not in their own neighborhood and thus do not immediately depend on themselves. Case ii provides for this immediate dependence.

These sequential machines are interconnected so that at any time  $t$  the component  $v_i$  of the current input vector to the machine at node  $s$  is the current state of the machine at node  $s+x_i$ . The interconnection graph of the network is a graph like the group graph of  $S$  w.r.t.  $X$  but the edges are directed and labeled. For any two group elements  $s$  and  $r$  and any  $x_i \in X$  there is an edge labelled  $x_i$  directed from  $r$  to  $s$  if and only if  $r = s+x_i$ . Fig. 3.5 shows the labelled and directed group graph corresponding to the  $S$  and  $X$  that generated Fig. 3.2. This graph is the same as the diagram of immediate effects (cf. Ashby [1963]) of a cellular automaton  $M = \langle S_3, V, \{b, e\}, \sigma \rangle$ .

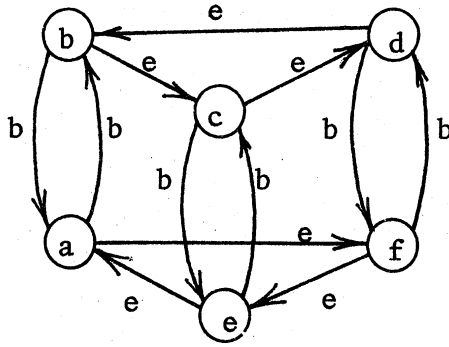


Figure 3.5: Labeled and directed group graph of  $S_3$  w.r.t.  $X = \{b, e\}$ .

Thus, the cells of the cellular automata can be regarded as sequential machines. If a cellular automaton is in configuration  $q$  at time  $t$ , then  $q(s)$  gives the state at time  $t$  of the sequential machine located at node  $s$  of the group graph.

### 3.5 Automorphism Groups

In this section we point out an alternative view of cellular automata that is more closely related to the non-automata theoretic approach than that provided by Def. 3.1. Again, only the autonomous case will be considered even though the results below can be extended to the more general situation. An automorphism of an automaton  $M$  with no input is simply an isomorphism from  $M$  to  $M$ , i.e. it is a bijection on  $Q$  that commutes with the transition function  $\delta$ . The composition of two automorphisms is again an automorphism, and the set of all automorphisms of  $M$  forms a group (Fleck [1962]) with function composition as the group

operation. We'll call this group  $G(M)$ .

Theorem 3.1

If  $M = \langle S, V, X, \sigma \rangle$  is a cellular automaton, then  $S$  is isomorphic to a subgroup of  $G(M)$ .

Proof:

For any  $s \in S$  the shift operator  $L_s$  is an automorphism of  $M$ : each  $L_s$  is a bijection since  $L_s L_{-s}$  is the identity on  $Q$ , and  $L_s \delta = \delta L_s$  since

$$\begin{aligned} (L_s \delta(q))(r) &= (\delta(q))(s+r) \\ &= L_{s+r} q = \sigma L_r L_s q = \sigma L_r (L_s q) \\ &= (\delta(L_s q))(r). \end{aligned}$$

Since the  $L_s$ ,  $s \in S$ , form a group isomorphic to  $S$  ( $s \mapsto L_s$ ), we are done.

Q.E.D.

The following theorem is a version of a result proved by Jump [1969].

Theorem 3.2 (Jump)

If  $M = \langle Q, \delta \rangle$  is an autonomous finite state automaton with automorphism group  $G(M)$ , then it can be realized by a cellular automaton with structure group  $S$  where  $S$  is any subgroup of  $G(M)$ .

Proof:

See Jump [1969].

An example given by Jump shows that the realization concept is necessary in this theorem. For an automaton  $M = \langle Q, \delta \rangle$  with automorphism group  $G(M)$ , it is not necessarily true that  $M$  is isomorphic to a cellular automaton with any subgroup of  $G(M)$  as structure group.  $M$  is isomorphic to a submachine of the cellular automaton. With this caveat, however, the structure group of a cellular automaton is the same as a group of automorphisms of an (autonomous) automaton.

Finally, we emphasize that Def. 3.1 does not require the structure group of a cellular automaton to be maximal in any sense. Any sequential machine can be viewed as a cellular automaton with a trivial structure group. We have not, therefore, delineated a subclass of automata to examine but have added additional structure to the idea of an automaton.

### 3.6 Generalizations

Yamada and Amoroso [1969] indicate that if the group  $S$  and template  $X$  are allowed to be uncountably infinite then a cellular automaton (or tessellation automaton) might be viewed as a model of a continuous physical system with the local transition function being an integration. The notation used in our definition of a cellular automaton was chosen so that this kind of generalization might be made more precise using standard concepts from functional analysis.

However, we do not adopt the more general framework in this thesis for two reasons. First, requiring  $S$  to be countable avoids topological considerations which tend to obscure the basic simplicity of our results. Second, our emphasis is on the modeling and simulation aspects of the cellular automaton formalism. We will not need to "discretize" a cellular automaton (according to our definition) in order to simulate it on a computer. The generalization to continuous cellular space does, though, serve to indicate the place occupied by the theory of cellular automata with respect to more classical mathematical formalisms.

The natural generalization (and the one that allows the harmonic analysis results described in Chapter 5 to go through) is to allow  $S$  to be any locally compact topological group and the template  $X$  to be a

compact subset of  $S$ . A cellular automaton could then have structure group  $\mathbb{R}^d$  and any closed and bounded region  $X$  as a neighborhood template. The natural generalization of  $Q$ , the set of configurations, would be to let it consist of functions with compact support. This definition would specialize to ours since any countable group with discrete topology is locally compact and, for the discrete topology, compactness and finiteness mean the same thing.

It is possible to go further however. The idea of a finite neighborhood or region of influence is important when cellular automata are considered from a computational point of view. Computers cannot handle an infinite amount of input at each time step. Viewed from the point of view of continuous mathematics, however, this problem takes a different form. The analogous concern is with whether the local operators are bounded or not. In the linear theory, compact neighborhoods insure boundedness but are not necessary for it. Systems described by hyperbolic PDEs have compact regions of influence that determine a finite upper limit on propagation speed (e.g. wave equation). Systems described by parabolic equations, on the other hand, do not. The heat equation, for example, has unbounded neighborhoods but they have bounded influence. Thus, it is natural to define a norm on the set of functions from  $S$  to  $V$  and require  $Q$ , the set of configurations, to consist only of functions with finite norm. The local transition function  $\delta$  would then be required to be bounded with respect to this norm. For example, if  $V = \mathbb{R}$ ,  $Q$  might be  $L^2(\mathbb{R})$ , the set of square-integrable (in the Lebesgue sense) real valued functions of  $\mathbb{R}$ . A local transition function might be given by



$$\sigma q = \int_{-\infty}^{\infty} w(s)q(s)ds$$

where  $w \in L^2(\mathbb{R})$ . The support of  $w$  would be the neighborhood template, but the integral will converge (in the norm of  $L^2(\mathbb{R})$ ) even if the support of  $w$  is all of  $\mathbb{R}$ . This follows from the fact that  $L^2(\mathbb{R})$  is a Hilbert space and hence reflexive. See, for example Schechter [1971]. In this framework, then, the idea of a neighborhood loses significance.

Much more general frameworks are possible. For example, we will be considering operators on spaces of group functions that are invariant with respect to various group generators and causal with respect to at least one (the generator for time). Our desire, however, is to stay within the sequential machine orientation where the concept of state plays a central role

## Chapter 4

### Cellular Automata and Linearity

#### 4.1 Introduction

In this chapter we begin the discussion of cellular automata that have the additional structure of linearity, or, what is the same thing, that are composed of cells which are linear sequential machines (LSMs). The resulting mathematical structures are very well studied and have applications in many different areas. It is basic knowledge within many disciplines that structural symmetry together with linearity gives rise to the notion of convolution of functions and hence to analysis techniques based on the Fourier transform. It is not clear, however, that cellular automata researchers realize that some of the objects of their study have close relatives within classical mathematics. It is apparently also true that few people trained in the more mathematical tradition are aware of the non-standard modeling formalism provided by the cellular automaton idea.

In the following sections we present, with as little mathematical abstraction as is practical, basic definitions and equations that will be used in later chapters. We also include a discussion of iterative linear networks and numerous examples. The discussion following Example 4.5 contains a proof of a result by Amoroso and Cooper [1971] and Ostrand [1971] that is quite natural within the framework developed.

#### 4.2 Linear Cellular Automata

In this section we formally define a linear cellular automaton (LCA) by adding appropriate structure to Def. 3.1. Our use of the

term "linear" is not the same as the use that occasionally appears in cellular automata literature where it describes arrays of cells that are linear strings, i.e. one-dimensional or associated with the group  $Z$ . We use the term LCA to refer to the subclass of cellular automata that are also linear systems because they have a vector space structure and a global transition function that is a linear operator. Some of these are one-dimensional arrays, but not all of them are.

Definition 4.1

A linear cellular automaton (LCA) is a cellular automaton

$M = \langle S, V, X, \sigma \rangle$  such that:

- i) The cell state set  $V = K^n$  for some non-negative integer  $n$  and field  $K$ .
- ii) There is a function  $W: S \rightarrow \text{Mat}_K(n, n)$  with finite support  $X$  such that the local transition function  $\sigma: Q \rightarrow K^n$  is given by:

$$\sigma q = \sum_{r \in S} W(r) q(r) . \quad (4.2-1)$$

We will denote the LCA  $M$  by  $\langle S, K^n, W \rangle$ . Its state set will be denoted by  $Q$  and its global transition function by  $\delta$ . The neighborhood  $X$  is implicit in the specification of  $W$ .

Note that  $\sigma$  is indeed local since if  $q_1, q_2 \in Q$  with  $q_1(r) = q_2(r)$  for  $r \in X$ , then

$$\sigma q_1 = \sum_{r \in S} W(r) q_1(r) = \sum_{r \in X} W(r) q_1(r) = \sum_{r \in X} W(r) q_2(r) = \sigma q_2 .$$

This is true because  $W(r)$  is the zero matrix for  $r \in S - X$ . We also point out that the zero vector of  $K^n$  is a quiescent state since  $\sigma$  is a linear operator. Thus Def. 3.1 part (iii) is satisfied.

It follows from Def. 3.1 (iv) that the global transition function  $\delta$  of an LCA is given by  $\delta(q) = q'$  where

$$q'(s) = \sigma L_s q = \sum_{r \in S} \omega(r) q(s+r) \quad \text{for all } s \in S. \quad (4.2-2)$$

If we define a function  $H: S \rightarrow \text{Mat}_K(n, n)$  by letting  $H(s) = W(-s)$ , for all  $s \in S$ , we can write

$$q'(s) = \sum_{r \in S} H(-s) q(s+r) = \sum_{-r \in S} H(r) q(s-r) = \sum_{r \in S} H(r) q(s-r). \quad (4.2-3)$$

The change of variable in the last step simply re-orders the summation due to the group structure of  $S$ . Equation (4.2-3) is true even when  $S$  is non-abelian. The function  $H$  is the reflection of  $W$  and has support  $\bar{X} = \{-x | x \in X\}$ . The set  $\bar{X}$  indicates the set of all cells that are influenced by the cell at the origin (instead of those that influence the cell at the origin). Therefore, consistent with terminology for PDEs (see Duff and Naylor [1966]), the set  $s + \bar{X} = s - X$  is called the region of influence of cell  $s$ .

Letting  $H(s) = (h_{ij}(s))$ , for all  $s \in S$ , (4.2-3) can be written in the expanded form

$$\begin{bmatrix} q'_0(s) \\ \vdots \\ q'_i(s) \\ \vdots \\ q'_{n-1}(s) \end{bmatrix} = \sum_{r \in S} \begin{bmatrix} h_{0,0}(r) & \dots & h_{0,n}(r) \\ \vdots & \cdot & \vdots \\ h_{n,0}(r) & \dots & h_{n-1,n-1}(r) \end{bmatrix} \begin{bmatrix} q_0(s-r) \\ \vdots \\ q_i(s-r) \\ \vdots \\ q_{n-1}(s-r) \end{bmatrix} \quad (4.2-4)$$

for all  $s \in S$ .

The component functions  $h_{ij}$  will be called spatial impulse reponse functions since if  $q = (q_0, \dots, q_{n-1})$  is such that

$$q_k(s) = \begin{cases} 1 & \text{for } k = j \text{ and } s = e \text{ (the identity of } S) \\ 0 & \text{otherwise} \end{cases}$$

then  $q_i' = h_{ij}$ . Letting  $W(s) = (w_{ij}(s))$ , for all  $s \in S$ , we call the  $w_{ij}$  the weighting functions of  $M$ . Observe that since the support of  $W$  is  $X$  we know that  $X_{ij} \subseteq X$ ,  $i, j = 0, \dots, n-1$ , where  $X_{ij}$  is the support of  $w_{ij}$ . Similarly for  $H$ .

The operators given by equations (4.2-2) and (4.2-3) involve the natural generalization of cross-correlation and convolution respectively. For  $f, g: S \rightarrow K$ , i.e. for scalar valued functions, we will occasionally write  $f *_S g$  for the function given by

$$(f *_S g)(s) = \sum_{r \in S} f(r)g(s-r) \quad \text{for all } s \in S.$$

This scalar convolution is commutative if and only if  $S$  is abelian.

Since the function  $H$  will often provide the most convenient form for specifying an LCA, we will sometimes denote an LCA  $M$  by  $\langle S, V, H \rangle$  with the understanding that the symbol  $H$  is always interpreted as the reflection of the function  $W$  found in Def. 4.1.

The function  $H$  corresponds to what is called the Green's function for linear PDEs with constant coefficients (Duff and Naylor [1966] or Roach [1970]). For PDEs, however,  $H$  is a function of time as well as space:  $H(s, t)$  gives the response at  $s$  (place) and  $t$  (time) to an impulse applied at the origin. In the case of discrete time, it is possible to specify this response for all time  $t \in \mathbb{Z}^+$  by specifying it for  $t = 1$ . Then  $H(s, t)$  for  $t > 1$  is determined by the composition property of a state determined system (Kalman, Falb, and Arbib [1969]). Thus, if we were to let  $H_G$  be the discrete analog of a Green's function for an LCA, we would have that  $H(s) = H_G(s, 1)$ , for all  $s \in S$ .

It is convenient to make several other observations about Def. 4.1. The set  $Q$  of configurations of an LCA is a vector space of vector-valued

functions with the usual component-wise operations of function addition and scalar multiplication. Each shift operator  $L_s$ ,  $s \in S$ , is a linear operator on  $Q$ . The local transition function  $\sigma: Q \rightarrow K^n$  is also a linear operator that is bounded since  $W$  has finite support. Finally, the global transition function  $\delta$  is a linear operator on  $Q$  and can be written in operator form as follows:

$$\delta = \sum_{r \in X} W(r) L_r. \quad (4.2-5)$$

This is verified by the following manipulations: for  $q \in Q$ ,

$$\delta(q) = \left( \sum_{r \in X} W(r) L_r \right) q = \sum_{r \in X} (W(r) L_r q)$$

since the  $L_r$  are linear. Thus, for any  $s \in S$ ,

$$(\delta(q))(s) = \sum_{r \in X} W(r) [(L_r q)(s)] = \sum_{r \in X} W(r) q(s+r) = \sum_{r \in S} W(r) q(s+r)$$

which is formula (4.2-2).

Within the framework provided by Def. 3.1, i.e.  $S$  countable and  $X$  finite, part (ii) of Def. 4.1 is the same as simply requiring the local transition function to be a linear operator. We have chosen to explicitly provide the representation  $W$  of the local transition function in order to simplify our presentation. In a more abstract theory formalized in terms of Hilbert space, the representation of  $\sigma$  is given in terms of an inner product which, in some cases, is a convergent integration (not necessarily over the set  $X$ ) as indicated in Sec. 3.6. For our purposes, however, Def. 4.1 suffices and we leave it for the more mathematically inclined reader to provide the appropriate generalizations.

### 4.3 Linear Networks and LCAs

Intuitively, an LCA is a cellular automaton in which each cell is a Moore type linear sequential machine (see Chapter 2) with state space dimension  $n$ . Moreover, the interconnections between the cells are linear so that the LCA as a whole is a linear sequential machine (although not necessarily finite dimensional). We will make this view explicit and provide an intuitive basis for understanding the various parts of Def. 4.1 by showing how an LCA can be realized by a linear network.

A sequential linear network is an interconnection of three kinds of components: unit delays, summers, and coefficient devices. Fig. 4.1 illustrates the function of these primitive components.

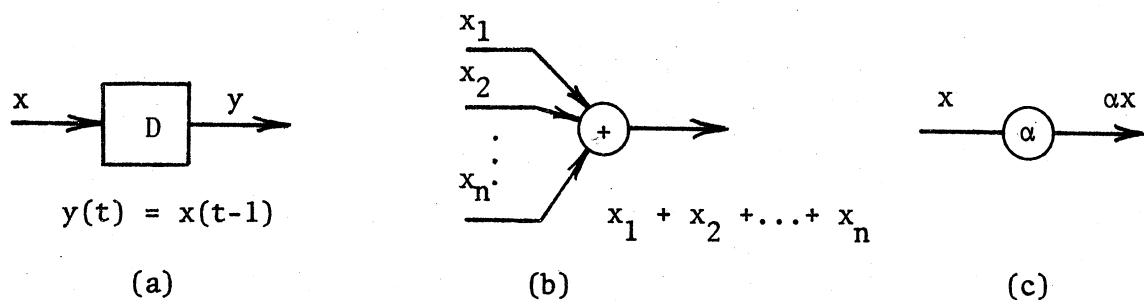


Figure 4.1: Sequential linear network components:  
(a) Unit delay, (b) Summer, (c) Coefficient device

The arithmetic operations shown are those of an underlying field  $K$ , i.e. the summation in Fig. 4.1(b) is the operation of  $K$  and  $\alpha$  in (c) is any element of  $K$ . A network represents an ideal computational process where a unit delay can store any field element, a summer performs exact multiplication by an arbitrary field element even if the field is infinite. Of course any real implementation of a network for an infinite field

will involve round-off error, but for the moment we will ignore this issue. It is of course also impossible to construct or simulate under all circumstances a network with an infinite number of components. For LCA, however, we can illustrate such infinite networks by drawing just the structure in the neighborhood of one cell. This structure is, by definition, finite and can be mentally iterated in space to give the complete structure. Elipses in our diagrams will indicate this iteration.

The following examples illustrate the network interpretation of Def. 4.1. Since it is very difficult to illustrate in detail the linear networks associated with LCAs having large  $S$  and  $n$ , we will take this opportunity to establish a diagram convention for use in the following chapters.

#### Example 4.1

Let  $M = \langle Z, R, W \rangle$  where  $W$  has support  $\{-1, 0, 1\}$ , i.e.  $M$  is an LCA with structure group  $Z$ , cell state set  $R$ , neighborhood template  $\{-1, 0, 1\}$  and local transition function  $W: Z \rightarrow \text{Mat}_R(1, 1) \approx R$ .

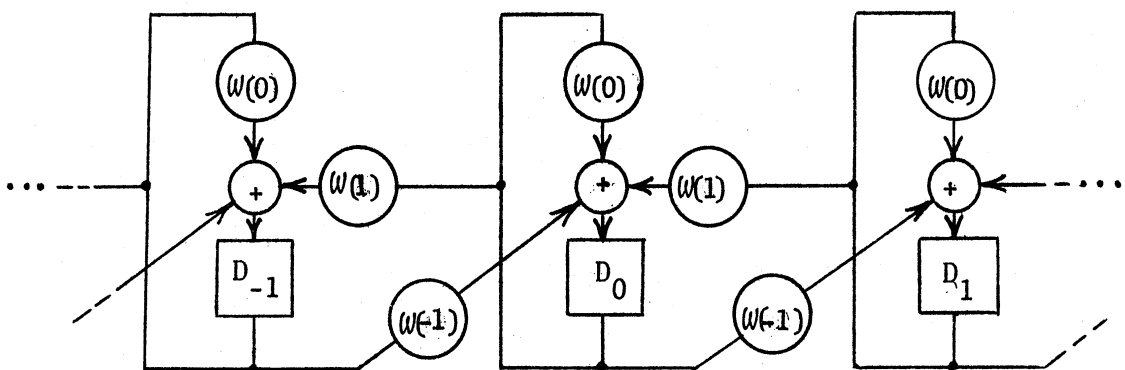


Figure 4.2: Linear network for a LCA with structure group  $Z$  and neighborhood template  $\{-1, 0, 1\}$ .



Figure 4.2 shows a section of an infinite linear network which realizes  $M$ . Note that 1) we have labeled the delays with the elements of the structure group  $Z$ ; 2) we have not drawn the coefficient devices for  $w(s)$ ,  $s \in Z - X$ , since, by definition, the coefficient values are all zero; and 3) it is clear that the neighborhood template  $X$  gives the labels of delays that have wires leading to  $D_0$  and that a wire from  $D_s$  goes through a coefficient device with coefficient value  $w(s)$ .

The more schematic representation shown in Fig. 4.3 is produced by drawing the labeled and directed group graph of  $Z$  w.r.t.  $X = \{-1, 0, 1\}$  as described in Sec. 3.4. Here, however, we label an edge arising from  $s \in X$  by a circle containing  $w(s)$  in order to suggest a coefficient device with value  $w(s)$ . The nodes are square to suggest unit delays and labeled with elements of  $Z$ . In the sequel diagram of this form will always be intended to represent an underlying linear network of unit delays, summers, and coefficient devices.

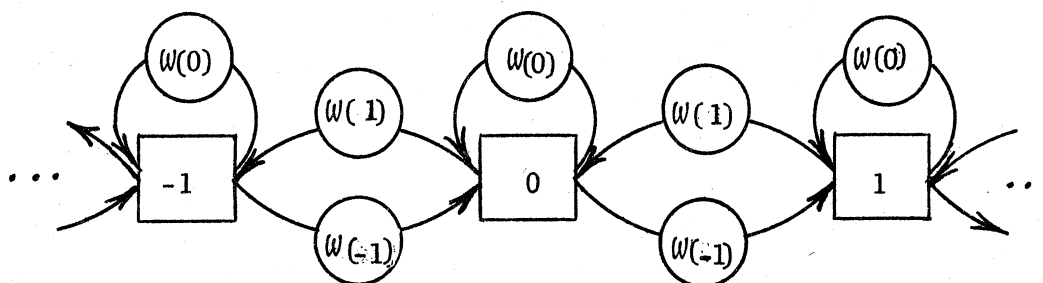


Figure 4.3: A schematic representation of an interative linear network.

A final simplification is shown in Fig. 4.4(a) where only the edges directed into one node are shown. By the definition of an LCA, this is

the information necessary to uniquely specify the complete network.

Fig. 4.4(b) similarly specifies a network labeled with the values of  $H$ .

Since  $H(s) = W(-s)$ , these networks are exactly the same.

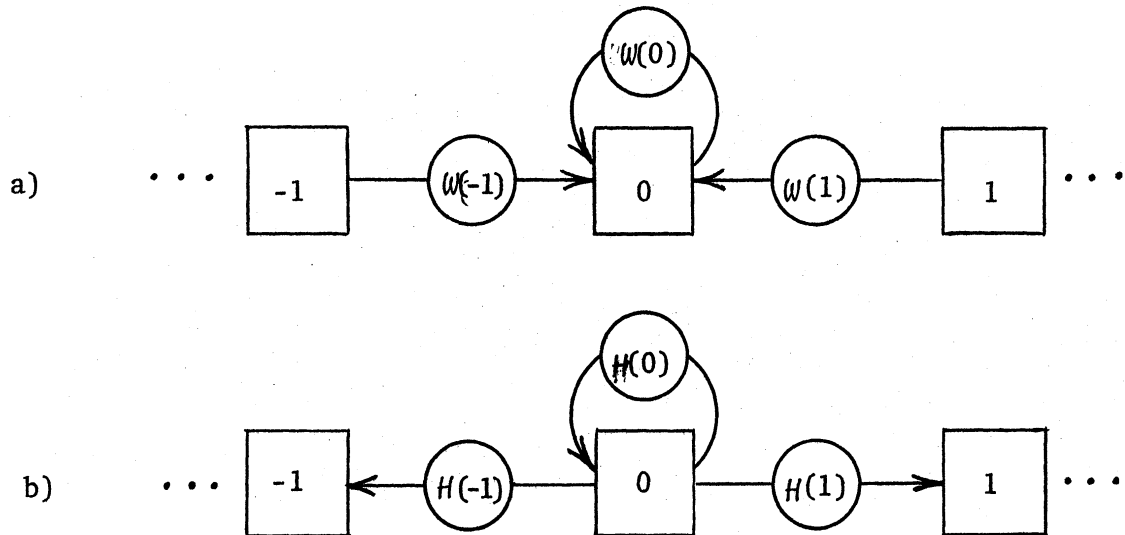


Figure 4.4: a) A minimal representation of an iterative network,  
b) the same network represented by values of  $H$ .

#### Example 4.2

Fig. 4.5 shows a simplified diagram for an LCA  $M = \langle \mathbb{Z}, \mathbb{R}^2, W \rangle$  where the support of  $W$  is  $\{-1, 0, 1\}$ .

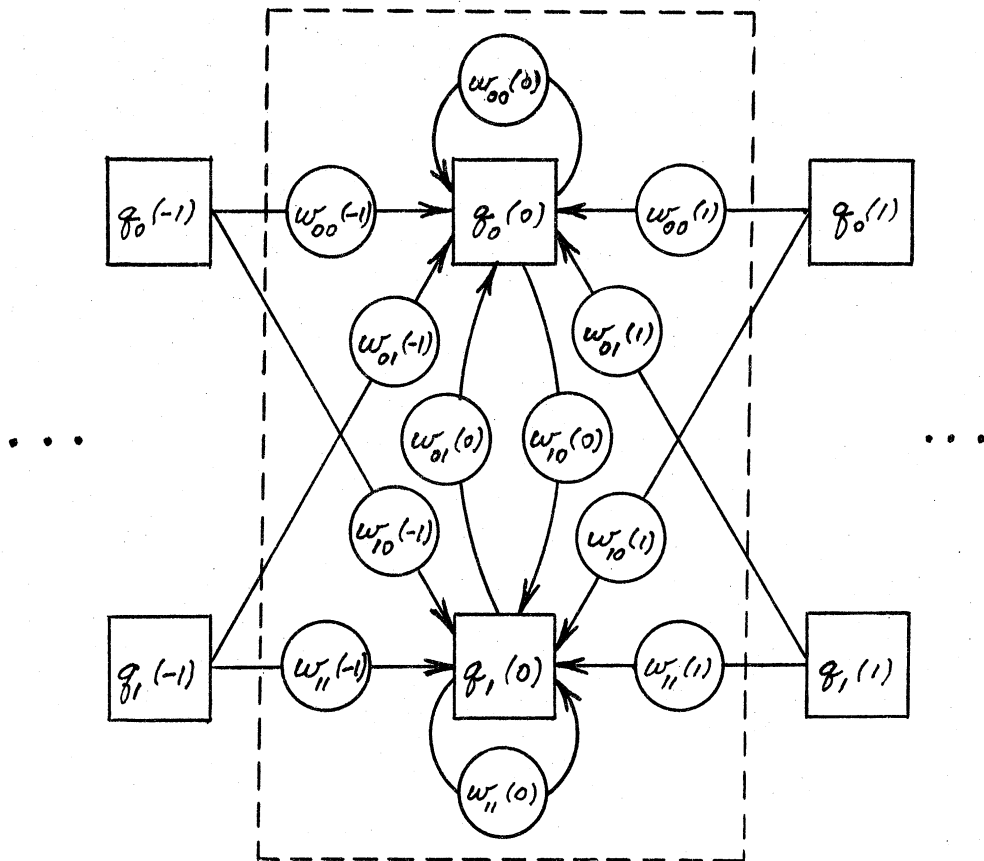


Figure 4.5: Diagram for  $M = \langle \mathbb{Z}, \mathbb{R}^2, W \rangle$ .

This network can be viewed as consisting of two interconnected layers each like the network in Fig. 4.4. The weighting functions  $w_{ij}$  specify the coefficients in the following way:  $w_{ij}(s)$  gives the coefficient for the connection to cell 0 of layer  $i$  from cell  $s$  of layer  $j$ . Alternatively, the structure within the dotted lines can be viewed as a cell following Sec. 3.4. It is an LSM with state set  $\mathbb{R}^2$ , input set  $\mathbb{R}^4$ , and transition function given by

$$\delta(q) = W(0)q + [W(-1), W(1)] x$$

for  $q \in \mathbb{R}^2$ ,  $x \in \mathbb{R}^4$  and  $[W(-1), W(1)] \in \text{Mat}_{\mathbb{R}}(2,4)$  obtained by writing the matrices  $W(-1)$  and  $W(1)$  next to each other.

It's interesting to note for this example that if each matrix  $W(s)$ ,  $s \in S$ , happens to be symmetric, then the cells of  $M$  are internally uniform and connected in such a way that Fig. 4.5 has an axis of symmetry running horizontally between layers 0 and 1. In this case the LCA  $M = \langle Z, \mathbb{R}^2, W \rangle$  realized by that network is isomorphic to an LCA  $M' = \langle Z \oplus Z_2, \mathbb{R}, W' \rangle$  for  $W'$  appropriately defined. This is true since the standard term symmetric applied to a matrix means that the operator represented by that matrix has a group of automorphisms isomorphic to  $Z_2$ . In general, if  $S'$  is a group of automorphisms of a single cell, then the LCA with structure group  $S$  composed of those cells is also an LCA with structure group  $S \oplus S'$  and cells of lower dimension.

#### 4.4 LCAs and I/O analysis of linear systems

The reader familiar with linear systems theory as presented in Chen [1970], or Schwarz and Friedland [1965], for example, will have recognized that the formal view of LCA developed above is essentially the same as that for the I/O analysis of a relaxed time invariant linear system. We have given the underlying group the interpretation of space instead of time and have not restricted it to  $\mathbb{R}$  or  $Z$ .

The spatial uniformity required of a cellular automaton is formally identical to time invariance. Theorem 3.1 says that the global transition function of a cellular automata commutes with a group of shift operators. Similarly the definition of a time-invariant system requires the system operator to commute with a set shift operator that forms a group isomorphic to  $\mathbb{R}$  or  $Z$ . Combining this with linearity as we do in

Def. 4.1 gives rise to convolution in exactly the same way that it usually arises.  $H$  is formally the same as the usual temporal impulse response matrix. The spatial analog of causality would have the neighborhood  $s+X$  of cells extending only on "one side" of cell  $s$ , e.g., for  $S = Z$ , if  $x \in X$  then  $x \leq 0$ .

This formal identity of concepts, however, should not obscure the different interpretations given to them. Mathematically, of course, whether  $S$  is treated as space or time or anything makes no difference (except that usually  $S$  is required to be singly generated and the neighborhood one-sided for the temporal interpretation). Any theorem is true for any interpretation. It is common practice, in fact, not to distinguish between space and time and to study "instantaneous" operators. Hennie [1961] does this for some cellular automata by looking at "space-time diagrams". We could formulate an LCA, for example, as a single convolution operation based on the group  $S \times Z$ .

Our view of an LCA as an autonomous sequential machine, however, means that we view its operation as an iteration in time of a convolution operation. At each time step, formula (4.2-3) is computed. (It should be clear by now that iteration means invariance with respect to certain shift operators: either spatial as in "iterative network" or temporal as in "iterative process.") Thus we have already taken advantage of the time invariance by incorporating it into a state space realization. The advantage of this view is twofold: 1) the explicit appearance of time is eliminated and 2) concepts of automata theory are immediately applicable.

In light of the observations above it is interesting to note that the "neighborhood standardization" technique presented by Yamada and

and Amoroso [1971] can be viewed as a kind of "spatial realization". They show that a cellular automaton with an arbitrary neighborhood template is isomorphic to a cellular automata in which each cell has only immediately adjacent neighbors. In the images of this analogy, the usual idea of realization involves reducing the "temporal neighborhood" to the immediately adjacent "earlier" cell so that the resulting state transition function is very local in time.

#### 4.5 Examples

This section consists of examples of LCAs that have been studied in other contexts or have been used as models of natural systems. The purpose is to demonstrate the diversity of these kinds of structures and to indicate some of the many connections this study has with other areas of interest.

##### Example 4.3

Finite difference approximation to the heat equation.

$$\begin{aligned} \text{The heat equation: } \frac{\partial u}{\partial t} &= k \frac{\partial^2 u}{\partial x^2} & u &= u(x,t) \\ u(x,0) &= f(x) \text{ for } -\infty < x < \infty & & (4.5-1) \\ 0 &\leq t & & \end{aligned}$$

This equation is the classical formalism for modeling diffusion on a one-dimensional strip. In this case the strip is infinitely long since we have not imposed any boundary conditions. The variables  $x$  and  $t$  are in  $\mathbb{R}$  and have the interpretation of space and time respectively, and  $u(x,t)$  gives the model's value for the temperature at point  $x$  and time  $t$ .

To approximate the solution using a finite difference scheme these variables are discretized and the solution computed at equally spaced points  $x_i$ ,  $i \in \mathbb{Z}$ , where  $x_{i+1} - x_i = \Delta x$ . (Of course in practice only a finite

number of points can be considered). Time is discretized also so the solution is considered only at points  $t_i, i \in \mathbb{Z}^+$  where  $t_{i+1} - t_i = \Delta t$ .

The second derivative can be approximated by

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u(x+\Delta x, t) - 2u(x, t) + u(x-\Delta x, t)}{\Delta x^2}$$

and the first derivative by

$$\frac{\partial u}{\partial t} \approx \frac{u(x, t+\Delta t) - u(x, t)}{\Delta t} .$$

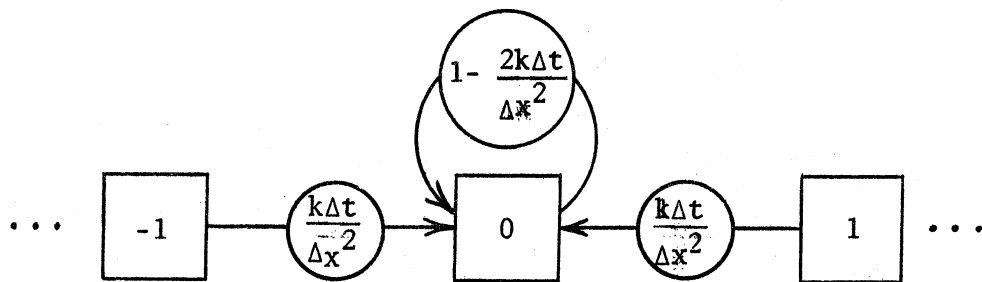
Thus (4.5-1) is written

$$\frac{u(x, t+\Delta t) - u(x, t)}{\Delta t} \approx k \left[ \frac{u(x+\Delta x, t) - 2u(x, t) + u(x-\Delta x, t)}{\Delta x^2} \right]$$

or

$$u(x, t+\Delta t) \approx \frac{k\Delta t}{\Delta x^2} u(x+\Delta x, t) + \left( 1 - \frac{2k\Delta t}{\Delta x^2} \right) u(x, t) + \frac{k\Delta t}{\Delta x^2} u(x-\Delta x, t) .$$

This is a two dimensional linear difference equation and can be thought of as specifying the LCA  $M = \langle \mathbb{Z}, \mathbb{R}, \omega \rangle$  shown in Fig. 4.6.



**Figure 4.6:** An LCA for a finite difference approximation to the heat equation.

If  $M$  is started with initial configuration  $q_0$  where  $q_0(s) = f(s\Delta x)$ , then the state of any cell  $s \in S$  at any time  $t \in \mathbb{Z}$  will be an approximation for  $u(s\Delta x, t\Delta t)$  of the initial value problem (4.5-1). Under many

circumstances this will not be a good approximation since care must be taken in choosing  $\Delta x$  and  $\Delta t$ . We will return to this point in Chapter 5.

#### Example 4.4

Finite difference approximation to a simplified wave equation  
(discussed by Gary [1972])

$$\begin{aligned} \frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} &= 0 & u &= u(x,t) \\ u(x,0) &= f(x) & \text{for } -\infty < x < \infty & \quad (4.5-2) \\ 0 &\leq t & & \end{aligned}$$

The exact solution of this equation is simply  $u(x,t) = f(x-ct)$  which represents a wave propagating to the right (toward higher values of  $x$ ) with velocity  $c$ . The shape of  $f$  is not changed as it moves. A difference scheme with no truncation error is given by

$$u(x, t+\Delta t) = u(x-c\Delta t, t) .$$

The LCA  $M = \langle Z, R, W \rangle$  indicated by Fig. 4.7 corresponds to this difference scheme. Letting the initial configuration  $q_0$  be given by  $q_0(s) = f(s\Delta x)$ , we have that  $q_t(s) = u(sc\Delta t, t\Delta t)$ . Or, letting  $c\Delta t = 1$ ,  $q_t(s) = u(s, t\Delta t)$ . In other words, the structure group  $Z$  of  $M$  is scaled so that the correct velocity results.  $M$  can be thought of as a simple shift register.

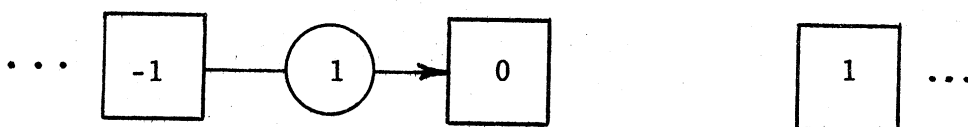


Figure 4.7: LCA for the hyperbolic PDE  $\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0$ .

In this case, the discretization produces no error. We will return to this example in Sec. 7.6 when we discuss laminated cellular automata.



Example 4.5 Circulents

Let  $S = Z_4$ , the cyclic group of order 4. An LCA  $M = \langle Z_4, K, W \rangle$  has a network like the one shown in Fig. 4.8 (we have let  $w(1) = w(2) = 0$  to make the diagram simpler).

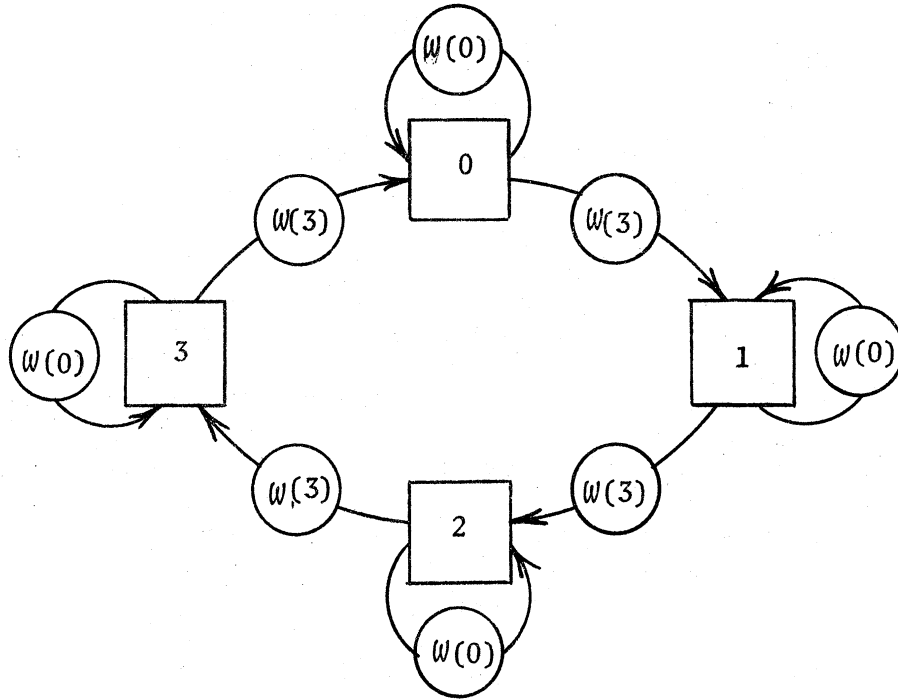


Figure 4.8: An LCA  $M = \langle Z_4, K, W \rangle$ .

If the global transition function  $\delta$  is written in matrix form, it looks like this:

$$\delta \begin{bmatrix} q(0) \\ q(1) \\ q(2) \\ q(3) \end{bmatrix} = \begin{bmatrix} w(0) & w(1) & w(2) & w(3) \\ w(3) & w(0) & w(1) & w(2) \\ w(2) & w(3) & w(0) & w(1) \\ w(1) & w(2) & w(3) & w(0) \end{bmatrix} \begin{bmatrix} q(0) \\ q(1) \\ q(2) \\ q(3) \end{bmatrix}$$

The first row of the matrix is the weighting function and the first column is the impulse response function (i.e.  $H(0) = W(0)$ ,  $H(1) = W(3)$ , etc.) A matrix having this kind of cyclical structure is commonly called a circulant or a discrete convolution matrix in the engineering literature (e.g. Waltz [1966]). Mathematical approaches formalize this structure as the regular matrix representation of an element of the group algebra of  $Z_4$  (see Curtis and Reiner [1962] or Boerner [1970]).

Example 4.5 Cellular automata that can reproduce arbitrary patterns

Amoroso and Cooper [1971] show that cellular automata can be designed which reproduce patterns in a sense to be defined below. Their systems are examples of LCAs which have an underlying finite field  $K$ . We'll consider only the one-dimensional case.

Let  $M$  be an LCA with structure group  $Z$  and let  $q$  be any configuration of  $M$  with finite support  $Y \subseteq Z$  with  $|Y| = n$ . A configuration  $q'$  contains  $k$  copies of the pattern found in  $q$  if there is  $k$ -tuple of integers  $(s_1, \dots, s_k)$  such that

$$\bigcap_{i=1}^k (s_i + Y) = \emptyset$$

and  $(L_{s_i} q')|_Y = q|_Y$  for  $i=1, \dots, k$ , i.e.,  $q(y) = q'(s_i + y)$  for all  $y \in Y$ .

The  $k$  copies are contained in a quiescent environment if the support of

$q'$  is

$$\bigcup_{i=1}^k (s_i + Y) .$$

$M$  reproduces the pattern found in  $q$  if, from initial configuration  $q$ , a configuration is reached that contains two or more copies of the pattern.\*

---

\*These definitions are from Amoroso & Cooper [1971], p. 459.

One result of Amoroso and Cooper can be stated as follows:

The LCA  $M = \langle Z, Z_p, W \rangle$  shown in Fig. 4.9, where, for  $p$  a prime,  $Z_p$  is the field of integers modulo  $p$ , will reproduce the pattern found in any configuration  $q$ . The resulting copies will be in quiescent environments.

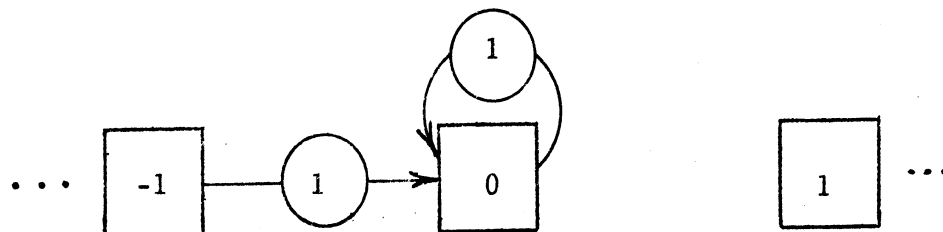


Figure 4.9: An LCA that reproduces arbitrary patterns

We give a proof of this result that is essentially the same as theirs but is based on the following more general ideas.

Suppose  $M$  is an LCA  $\langle S, K, W \rangle$  where  $W: S \rightarrow \text{Mat}_K(1,1) \approx K$ , i.e. the cells have one-dimensional state spaces. Let  $X$  be the neighborhood template of  $M$ , i.e.  $X$  is the support of  $W$ . Define, for each  $x \in X$ ,  $w_x: S \rightarrow K$  by

$$w_x(s) = \begin{cases} W(x) & \text{for } s = x \\ 0 & \text{otherwise .} \end{cases}$$

In other words, each  $w_x$  has just one non-zero value which is at point  $x$  and equal to  $W(x)$ . For each  $x \in X$ , let  $M_x$  be the LCA  $\langle S, K, w_x \rangle$  with global transition function denoted by  $\delta_x$ . Each  $M_x$  is an LCA in which each cell has only a single neighbor. Fig. 4.10 shows  $M_0$  and  $M_{-1}$  for the LCA in Fig. 4.9.

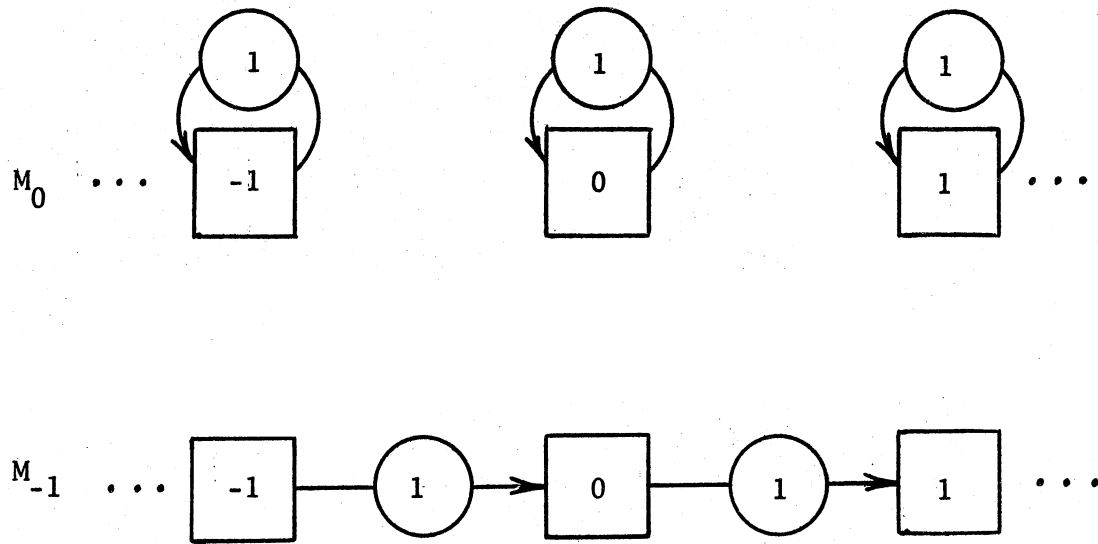


Figure 4.10: The LCAs  $M_0$  and  $M_{-1}$  associated with the LCA shown in Fig. 4.9.

It is clearly not true in general that

$$\delta^t(q) = \sum_{x \in X} \delta_x^t(q)$$

for any configuration  $q$  and time  $t$ . The linearity of  $M$  implies that this is true for  $t = 1$ , but it is not necessarily true, even with linearity, for  $t > 1$ . It's not generally possible simply to run the  $M_x$  separately and then superimpose the results to find what  $M$  would have computed. However, under certain restrictions it is possible.

#### Theorem 4.1

Let  $M$  be an LCA  $\langle S, K, \omega \rangle$  where  $S$  is abelian and  $K$  has characteristic  $p > 0$ . With LCA  $M_x$ ,  $x \in X$ , defined as above it is true that

$$\delta^{p^m}(q) = \sum_{x \in X} \delta_x^{p^m}(q) \quad \text{for any } q \in Q \text{ and } m \in \mathbb{Z}^+.$$

**Proof:**

It's well known (e.g. Rotman, p. 154) that the binomial theorem leads to the fact that for  $a, b \in K$ , where  $K$  has characteristic  $p > 0$ ,

$(a+b)^p = a^p + b^p$  for  $m > 0$ . This is clearly also true for any finite sum. Recall that the global transition function  $\delta$  of  $M$  can be written in operator form as

$$\delta = \sum_{x \in X} w(x) L_x.$$

Since  $S$  is abelian  $L_r L_s = L_s L_r$  and we can apply the fact stated above so that

$$\delta^{p^m} = \left( \sum_{x \in X} w(x) L_x \right)^{p^m} = \sum_{x \in X} (w(x) L_x)^{p^m} = \sum_{x \in X} \delta_x^{p^m}.$$

Thus

$$\delta^{p^m}(q) = \left( \sum_{x \in X} \delta_x^{p^m} \right)(q) = \sum_{x \in X} \delta_x^{p^m}(q)$$

since the  $\delta_x^{p^m}$  are linear.

Q.E.D.

The Amoroso-Cooper result can be seen to follow from Theorem 4.1 as follows: since  $Z_p$  has characteristic  $p > 0$ , the LCAs  $M_0$  and  $M_{-1}$  shown in Fig. 4.10 can be run separately each starting with configuration  $q$ . For any  $m > 0$ , their configurations after  $p^m$  time steps can be superimposed to produce what  $M$  produces in  $p^m$  time steps.  $M_0$  does nothing to  $q$  since  $\delta_0$  is the identity operator.  $M_{-1}$  computes  $L_{-1}^{(p^m)}$ , i.e. it shifts  $q$  to the right  $p^m$  cells. Superimposing these results for  $m$  large enough to eliminate any pattern overlap produces a configuration containing two copies, each in a quiescent environment, of the pattern in  $q$ .

Amoroso and Cooper show a similar result for  $S = Z^2$  and conjecture that it's possible for any finite dimension, i.e. for  $S = Z^d$ . This

conjecture was proven by Ostrand [1971]. By Theorem 4.1 and an argument similar to the one above these results are clearly true for an arbitrary countable abelian group  $S$ . Theorem 4.1 expresses the essential fact underlying all of these results.

Finally we note that pattern reproduction for arbitrary finite dimension can be accomplished using a simpler scheme than that presented by Ostrand. Simply construct a laminated LCA (see Sec. 7.6 below) using the LCA shown in Fig. 4.9. Fig. 4.11 shows the scheme for  $S = \mathbb{Z}^2$ .

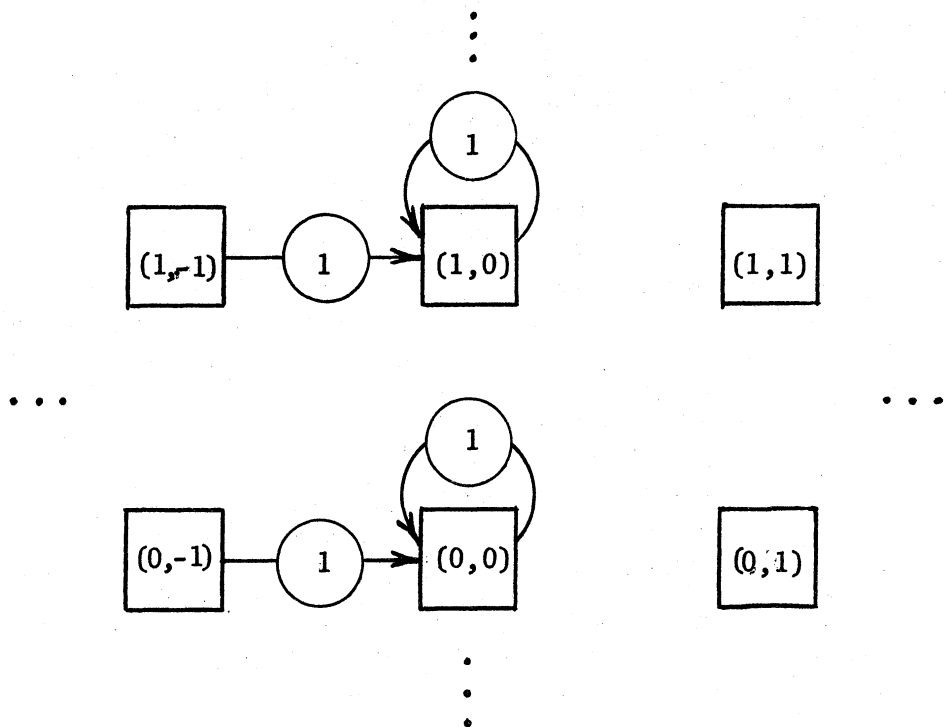


Figure 4.11: A laminated LCA that reproduces arbitrary patterns for  $S = \mathbb{Z}^2$ .

Since there is no interference between horizontal slices in this space, any two-dimensional pattern will produce a copy of itself that

will appear to the right of the original position. The scheme presented by Ostrand produces two copies - one to the right and one above.

A similar lamination of the LCA of Fig. 4.11 will produce a scheme for

$$S = Z^3.$$

## Chapter 5

### Linear Cellular Automata and Harmonic Analysis

#### 5.1 Introduction

In the case of linear cellular automata the simplicity with which the structure is specified, i.e. the fact that the corresponding networks are iterative, can be translated into simplicity of behavior. This is accomplished by using Fourier analysis which is applicable in this case for the same reasons it is in other contexts: homogeneity of structure (e.g. time invariance and/or spatial invariance) together with linearity gives rise to the operation of convolution which can be uncoupled by the Fourier transform. This situation is found in the I/O analysis of time invariant linear systems (e.g. signal processing), in the study of linear PDEs with constant coefficients, and in models of lenses and imaging systems in optics (see Andrews [1970]).

Often, though, harmonic analysis techniques are learned in a specific context and an understanding of the principles involved becomes tied to a specific application. Among the misconceptions arising in this way is the belief that functions need to be real or complex valued functions of one or several real variables in order for harmonic analysis techniques to apply. Less standard applications tend to be viewed as approximations to the continuous case or as necessarily derived from the continuous case. As a result, it is often felt that no real working knowledge and understanding can be achieved unless one feels comfortable with integrals, continuity, convergence properties, distributions, and other concepts necessary for dealing with functions of a real variable.

That these beliefs are unjustified is becoming clear as elements of abstract harmonic analysis become known in applications oriented areas.



The chief reason for interest in this approach is the greatly increased interest in harmonic analysis due to the introduction of the Fast Fourier Transform (FFT) algorithm in 1965 for computing the Discrete Fourier Transform (DFT). The DFT can be understood completely within a purely algebraic framework that is a simple special case of a theory in which functions of locally compact topological groups are considered. The Fourier transform, Fourier series, Z-transform, the DFT, and even the Walsh transform are special cases in a unifying theory in which the structure of abelian groups plays a central role.

In this chapter we will precisely state the conditions under which harmonic analysis is applicable to LCAs and interpret the results in automata theoretic terms. Our theorems will be illustrated by several examples which serve to unify previously known techniques.

The literature pertaining to harmonic analysis and its applications is very extensive since uniform, linear structures are ubiquitous as models of natural systems. Bracewell [1965] provides an intuitively satisfying view of Fourier series and transforms. The DFT and FFT are well covered in engineering literature such as the IEEE special issues on the FFT [1967] and [1969], Bergland [1969] or Cooley, Lewis, and Welch [1967]. For the more computer oriented literature on digital signal processing see Gold and Rader [1969] or Andrews [1970]. The Walsh Transform is covered by Harmuth [1972]. Nicholson [1971] provides a concise algebraic development of the more abstract approach and gives an algebraic view of the FFT as does Cairns [1971]. The mathematical literature is represented by Rudin [1962] and Hewitt and Ross [1963].

## 5.2 The Discrete Fourier Transform

In this section we will define in a general form what has come to be known as the discrete or finite Fourier transform (DFT). Although the DFT is usually defined over the complex numbers and for simple cyclic groups, we give a definition that allows certain finite fields and arbitrary finite abelian groups. Sec. 5.3 specializes the definition to familiar cases. We follow the presentation given by Nicholson [1971] and Aho, Hopcroft, and Ullman [1974].

Let  $G$  be a finite abelian group, with  $|G| = N$ . We will denote the group operation by '+'. The well known fundamental structure theorem for abelian groups (Rotman [1973]) says that  $G$  is isomorphic to a direct sum of cyclic groups:

$$G \approx Z_{n_1} \oplus \dots \oplus Z_{n_d} \quad \text{where} \quad \prod_{j=1}^d n_j = N \quad \text{and} \quad n_i | n_{i+1}. \quad (5.2-1)$$

Here  $Z_{n_i}$  denotes the cyclic group of order  $n_i$  which can be conveniently thought of as represented by the integers with modulo  $n_i$  addition. We will call a decomposition as in (5.2-1) a canonical decomposition of  $G$ .

Hence  $G$  can be identified with the set of  $k$ -tuples of integers  $\{(s_1, \dots, s_d) \mid 0 \leq s_i < n_i\}$  with addition modulo  $(n_1, \dots, n_d)$ . It will occasionally be convenient to denote an element  $s \in G$  as  $(s_1, \dots, s_d)$ . The integer  $m$  which is the least integer such that  $m \cdot s = s + \dots + s$  ( $m$  times)  $= 0$  for every  $s \in G$  is the exponent of  $G$ . For  $G$  decomposed according to (5.2-1),  $m = n_d$ . A primitive or principle  $p^{\text{th}}$  root of unity is an element  $\omega \in R$  which generates a multiplicative cyclic group of order  $p$ . In other words  $\omega$  is such that

- 1)  $\omega^p = 1$  and
- 2)  $\omega^j \neq 1$  for  $0 < j < p$ .

The most familiar example of a principle  $p^{\text{th}}$  root of unity is  $e^{2\pi i/p}$  where  $i = \sqrt{-1}$  in the ring of complex numbers.

The complex exponential functions  $\phi_x(t) = e^{ixt}$  into which a large class of complex valued functions of  $\mathbb{R}$  can be resolved by the usual Fourier transform are special examples of functions called group characters. For the discrete case considered here, the group characters can be defined using only the decomposition (5.2-1) of  $G$  and roots of unity in the integral domain  $R$ . Suppose that  $R$  contains a principle  $m^{\text{th}}$  root of unity where  $m$  is the exponent of  $G$  (here  $m = n_d$ ). Let  $\omega$  be this root and recall that  $F(G, R) = \{f | f: G \rightarrow R\}$ . A set of group characters is defined as follows:

Definition 5.1

The set of group characters of  $G$  in  $R$  is the set

$$\Phi(G, R) = \{\phi_s | s \in G\} \subseteq F(G, R) \quad \text{where}$$

$$\phi_{(s_1, \dots, s_d)}((r_1, \dots, r_d)) = \prod_{j=1}^d \omega_j^{s_j r_j n_d / n_j} \quad (5.2-2)$$

for all  $(s_1, \dots, s_d), (r_1, \dots, r_d) \in G$ .

If we further assume that  $N^{-1} \in R$  where we take  $N = 1+1+\dots+1$  ( $N$  times), then the following properties, which are well known for the case  $R = \mathbb{C}$ , can be proven (see Cairns [1971], Nicholson [1971], and Aho, Hopcroft, and Ullman [1974]):

- 1) Each  $\phi_s \in \Phi(G, R)$  is a homomorphism from  $G$  into the multiplicative monoid of  $R$ , i.e.

$$\phi_s(\alpha + \beta) = \phi_s(\alpha) \phi_s(\beta) \quad \text{for all } s, \alpha, \beta \in G.$$

In fact,  $\Phi(G, R)$  is the set of all such homomorphisms excluding the one identically equal to 0. See Hewitt and Ross [1963] p. 345.

Nicholson [1971] proves this for integral domains.

- 2)  $\Phi(G, R)$  is a group with group operation '.' given by pointwise multiplication of functions, i.e.

$$(\phi_s \cdot \phi_r)(\alpha) = \phi_s(\alpha)\phi_r(\alpha) \quad \text{for } s, r, \alpha \in G.$$

- 3)  $\Phi(G, R)$  with operation given in 2) is isomorphic to  $G$  under the isomorphism  $s \mapsto \phi_s$ . Thus  $\phi_s(\alpha)\phi_r(\alpha) = \phi_{s+r}(\alpha)$  for all  $s, r, \alpha \in G$ .

- 4)  $\phi_s(r) = \phi_r(s)$  for all  $r, s \in G$ .

- 5) Each character is an eigenfunction of every shift operator

$L_s$ ,  $s \in G$  (see Sec. 3.2). In particular  $L_s \phi_r = \phi_r(s)\phi_r$  for all  $s, r \in G$ .

- 6) The characters are orthogonal in the sense that

$$\frac{1}{N} \sum_{\alpha \in G} \phi_{-s}(\alpha)\phi_r(\alpha) = \begin{cases} 1 & \text{if } r = s \\ 0 & \text{otherwise} \end{cases}$$

- 7) Any  $f \in F(G, R)$  can be written as a unique linear combination of the characters as follows:

$$f = \frac{1}{N} \sum_{\alpha \in G} \hat{f}(\alpha)\phi_{-\alpha}$$

where

$$\hat{f}(\alpha) = \sum_{\beta \in G} f(\beta)\phi_{\alpha}(\beta) \quad \text{for all } \alpha \in G.$$

The equations in 7) can be taken to define an invertible operator on  $F(G, R)$  which we will call the discrete Fourier transform with respect to  $G$ . We make this fact precise by stating it as a theorem. For a

proof of the existence and invertibility of the DFT in this general setting see Nicholson [1971].

Theorem 5.1

Let  $G$  be a finite abelian group with  $|G| = N$  and exponent  $m$ . Let  $R$  be an integral domain containing  $N^{-1}$  and a principle  $m^{\text{th}}$  root of unity. There is an invertible operator  $F_G$  on  $F(G, R)$  called the discrete Fourier transform w.r.t.  $G$  given by

$$F_G f = g \quad \text{where} \quad g(\alpha) = \sum_{\beta \in G} f(\beta) \phi_{\alpha}(\beta) \quad \text{for all } \alpha \in G \quad (5.2-3)$$

where the  $\phi_{\alpha}$  are the characters of  $G$  in  $R$ . It's inverse, denoted by  $F_G^{-1}$ , is given by

$$F_G^{-1} g = f \quad \text{where} \quad f(\beta) = \frac{1}{N} \sum_{\alpha \in G} g(\alpha) \phi_{-\alpha}(\beta) \quad \text{for } \beta \in G. \quad (5.2-4)$$

This theorem gives one of several forms that the DFT can take. Cooley, Lewis, and Welch [1969], for example, normalize the group characters so that the factor  $1/N$  appears in the forward transform rather than the inverse. Note also that from fact 3) above we can con-

clude that for every  $s, \alpha \in G$ ,  $\phi_s(\alpha)$  has a multiplicative inverse in  $R$ . This is because  $\phi_s(\alpha) \phi_{-s}(\alpha) = \phi_e(\alpha)$  where  $e$  is the identity of  $S$ . From Def. 5.1 it can be verified that  $\phi_e(\alpha) = 1$  for all  $\alpha \in S$ . Thus  $(\phi_s(\alpha))^{-1} = \phi_{-s}(\alpha)$  and, using fact 4),  $(\phi_{\alpha}(s))^{-1} = \phi_{-\alpha}(s)$ . Expression (5.2-4) is therefore equivalent to

$$F_G^{-1} g = f \quad \text{where} \quad f(\beta) = \frac{1}{N} \sum_{\alpha \in G} g(\alpha) (\phi_{\beta}(\alpha))^{-1}$$

Finally note that  $F(G, R)$  is a module with the usual operations of function

addition and multiplication by constants, and that  $F_G$  is a module homomorphism or, in case  $R$  is a field, a linear operator. The discrete Fourier transform given in Theorem 5.1 has the same intuitive meaning as the more familiar Fourier transform and Fourier series. The group characters correspond to the complex exponential functions. Every function in  $F(G, R)$  is a superposition of the group characters where each character in the sum is multiplied by an element of  $R$ . A Fourier coefficient  $\hat{f}(\alpha)$  gives the "amplitude" of the character  $\phi_{-\alpha}$  in the sum (in other formulations of the DFT,  $\hat{f}(\alpha)$  gives the amplitude of  $\phi_{\alpha}$ ). We will often say that  $\hat{f}$  is the "frequency domain" representation of  $f$ . While the term "frequency" may not retain its literal significance in the more abstract cases, it can be given the following interpretation. Each group character  $\phi_s$  is a homomorphism from  $G$  into the multiplicative structure of  $R$ . Therefore the kernel of  $\phi_s$  is the subset of  $G$  on which  $\phi_s$  is 1, and, since  $\phi_s$  is a group homomorphism, this subset is a subgroup of  $G$ . Subgroups of an abelian group  $G$  consist of "regularly spaced" elements of  $G$ . One can therefore think of the frequency of a character in terms of the "spacing" of its kernel - low frequencies have widely spaced kernels and vice versa.

In what follows we will always use one of two notational conversions to indicate the DFT w.r.t.  $G$ : 1) as above,  $F_G$  and  $F_G^{-1}$  will denote the operator and its inverse; 2) for any function  $f$ ,  $\hat{f}$  will denote  $F_G f$  in cases where it is clear what group is involved. For longer expressions the hat will be placed following the expression, e.g.  $(f *_S g)^{\hat{}}$  means  $F_G(f *_S g)$ . The integral domain  $R$  underlying these operations will be clear from the context.

### 5.3 The Convolution Theorem

The significant fact about the Fourier transform, whether the continuous or discrete one, is that it simplifies convolution. Letting  $G$  and  $R$  satisfy the hypothesis of theorem 5.1, the convolution with respect to  $G$  of two functions  $f, g \in F(G, R)$  is the function  $f *_G g \in F(G, R)$  given by

$$(f *_G g)(s) = \sum_{r \in G} f(r)g(s-r), \quad \text{for all } s \in G.$$

Convolution, then, is a way to "multiply" functions in  $F(G, R)$ . Another way to multiply functions is to pointwise multiply them. We'll use the symbol  $\cdot$  to indicate this operation on  $F(G, R)$ , i.e. for  $f, g \in F(G, R)$ ,  $f \cdot g$  is the function given by

$$(f \cdot g)(s) = f(s)g(s), \quad \text{for all } s \in G.$$

Each of these operations makes  $F(G, R)$  into a (different) commutative algebra: the operation given by  $*_G$  or  $\cdot$  in addition to the usual module structure of pointwise addition and scalar multiplication. The algebra with the convolution operation is called the group algebra or group ring. The convolution theorem says that  $F_G$  is an isomorphism from the group algebra to the point-wise multiplication algebra. Since we know that  $F_G$  is an isomorphism of modules, we need only that it preserve the "multiplication" operation of the algebras. That this is true is well-known for  $R = \mathbb{C}$  and is proven in the general form by Nicholson [1971].

#### Theorem 5.2 (Convolution theorem)

Let  $R$  and  $G$  satisfy the hypothesis of Theorem 5.1. Then for any  $f, g \in F(G, R)$  we have that

$$(f *_G g)^\wedge = \hat{f} \cdot \hat{g} \quad .$$

In other words, for convolution and the DFT both based on the same group  $G$  and integral domain  $R$ , the DFT of the convolution of two functions is the pointwise product of the DFTs of the functions.

#### 5.4 Extension to Vector-valued Functions

It is convenient at this point to describe the well-known extensions of the DFT and the convolution theorem to vector-valued functions.

Consider the space of functions  $F(G, R^n)$  where  $R^n$  is a module with the usual component-wise operations. Assume that  $G$  and  $R$  satisfy the hypotheses of Theorem 5.1.

The DFT can be extended to  $F(G, R^n)$  by applying the operator in (5.2-3) to component functions. For any  $f \in F(G, R^n)$  recall the convention of writing  $f(s) = (f_0(s), \dots, f_{n-1}(s))$  where  $f_i \in F(G, R)$ . Define

$$F_G^n : F(G, R^n) \rightarrow F(G, R^n)$$

by

$$F_G^n(f) = (F_G f_0, \dots, F_G f_{n-1}) = (\hat{f}_0, \dots, \hat{f}_{n-1}) = \hat{f} \quad (5.4-1)$$

Since  $F_G$  is a bijection on  $F(G, R)$ ,  $F_G^n$  is a bijection on  $F(G, R^n)$ , and its inverse is defined in the obvious way, i.e.

$$(F_G^n)^{-1}(f) = (F_G^{-1} f_0, \dots, F_G^{-1} f_{n-1}). \quad (5.4-2)$$

Let  $\text{Mat}_R(n, n)$  denote the set of  $n \times n$  matrices with entries in  $R$  and let  $A : G \rightarrow \text{Mat}_R(n, n)$ . For any  $f \in F(G, R^n)$  let  $A *_G f : G \rightarrow R^n$  be given by

$$(A *_G f)(s) = \sum_{r \in G} A(r) f(s-r), \text{ for all } s \in G. \quad (5.4-3)$$

This is what we'll mean by convolution for  $F(G, R^n)$ . Note that it is



not a way to "multiply" functions in  $F(G, R^n)$  as it is for  $n = 1$  since  $A$  is a matrix valued function. The extended DFT, however, simplifies this generalization of convolution if we define the DFT of a matrix and the analog of pointwise multiplication of functions in the following ways.

For any  $A: G \rightarrow \text{Mat}_R(n, n)$ , we will write  $A(s) = (a_{ij}(s))$ ,  $s \in G$ . Letting  $\hat{a}_{ij} = F_G a_{ij}$ ,  $0 \leq i, j \leq n$ , define  $\hat{A}: G \rightarrow \text{Mat}_R(n, n)$  where  $\hat{A} = (\hat{a}_{ij})$ . For  $f \in F(G, R^n)$ , let  $A \cdot f$  denote the function in  $F(G, R^n)$  given by

$$(A \cdot f)(s) = A(s)f(s), \quad \text{for all } s \in G. \quad (5.4-4)$$

The generalization of pointwise multiplication, then, is pointwise action of matrices on vectors.

In this framework, the convolution theorem appears in the following form. Since a concise proof of this generalization seems to be lacking in the elementary literature, we provide one here even though the result is well-known (at least for  $R = \mathbf{C}$ ).

### Theorem 5.3

Let  $R$  and  $G$  satisfy the hypotheses of Theorem 5.1. Then for any  $f \in F(G, R^n)$  and  $A: G \rightarrow \text{Mat}_R(n, n)$ , we have that

$$(A *_G f)^\wedge = \hat{A} \cdot \hat{f}.$$

Proof:

According to (5.4-1) we need only show that

$$F_G[(A *_G f)_i] = (\hat{A} \cdot \hat{f})_i \quad \text{for } i=0, \dots, n-1.$$

From (5.4-3) we have for  $s \in G$  and  $i=0, \dots, n-1$  that

$$\begin{aligned}
(A *_G f)_i(s) &= \sum_{r \in G} [A(r)f(s-r)]_i \\
&= \sum_{r \in G} \left[ \sum_{j=0}^{n-1} a_{ij}(r) f_j(s-r) \right] \\
&= \sum_{j=0}^{n-1} \left[ \sum_{r \in G} a_{ij}(r) f_j(s-r) \right] \\
&= \sum_{j=0}^{n-1} (a_{ij} *_G f_j)(s) = \left( \sum_{j=0}^{n-1} a_{ij} *_G f_j \right)(s) .
\end{aligned}$$

$$\text{Hence } (A *_G f)_i = \sum_{j=0}^{n-1} a_{ij} *_G f_j, \quad i=0, \dots, n-1 .$$

Now, since  $F_G$  is a module homomorphism, one can write

$$\begin{aligned}
F_G[(A *_G f)_i] &= F_G \left[ \sum_{j=0}^{n-1} a_{ij} *_G f_j \right] \\
&= \sum_{j=0}^{n-1} F_G(a_{ij} *_G f_j) \\
&= \sum_{j=0}^{n-1} \hat{a}_{ij} \cdot \hat{f}_j \quad \text{by Theorem 5.2.}
\end{aligned}$$

Making the argument  $s$  explicit, we have that

$$\begin{aligned}
\sum_{j=0}^{n-1} (\hat{a}_{ij} \cdot \hat{f}_j)(s) &= \sum_{j=0}^{n-1} \hat{a}_{ij}(s) \hat{f}_j(s) \\
&= (\hat{A}(s) \hat{f}(s))_i \\
&= (\hat{A} \cdot \hat{f})_i(s) \quad \text{for all } s \in G \text{ by } (
\end{aligned}$$

Therefore,  $F_G[(A *_G f)_i] = (\hat{A} \cdot \hat{f})_i, \quad i=0, \dots, n-1.$

Q.E.D.

The reader may note that for  $n=1$  Theorems 5.2 and 5.3 are identical. In fact, in a more abstract presentation the more general situation would be discussed directly in terms of the theory of multilinear operators and Theorem 5.2 would be presented as a special case of Theorem 5.3. We will therefore use the notation  $'*_G'$  and  $'.'$  for both the scalar-valued and vector-valued cases. We will also write  $F_G$  for  $F_G^n$  and  $F_G^{-1}$  for  $(F_G^n)^{-1}$  when it is clear from the context that an operator on  $F(G, R^n)$  is required.

### 5.5 Examples of the Discrete Fourier Transform and Convolution

In order to emphasize the generality of Theorems 5.1 and 5.2 we present several examples. The first is the most commonly described form of the DFT and convolution:  $R = \mathbb{C}$  and  $G$  is a cyclic group. The last example is very unfamiliar to most people as has intriguing possibilities for modeling certain natural systems.

#### Example 5.1

The most common form of the DFT is for the case when  $G = Z_N$ , the finite group of integers with modulo  $N$  addition, and  $R = \mathbb{C}$ , the complex numbers. The exponent of  $Z_N$  is  $N$  and a principle  $N^{\text{th}}$  root of unity in  $\mathbb{C}$  is  $e^{2\pi i/N}$ . Following Def. 5.1, we have that  $\phi_k(j) = e^{2\pi i k j / N}$  for all  $j, k \in Z_N$ . Thus we have that

$$\hat{f}(k) = \sum_{j=0}^{N-1} f(j) e^{2\pi i k j / N}, \quad k=0, \dots, N-1$$

and

$$f(j) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}(k) e^{-2\pi i k j / N}, \quad j=0, \dots, N-1.$$

In this case, it can be seen that  $\phi_{-s}(\alpha) = \overline{\phi_s(\alpha)}$ , for all  $s, \alpha \in Z_N$ , where the bar indicates complex conjugation. The orthogonality relation given by fact 6) of Sec. 5.2 can therefore be stated in the more familiar form

$$\frac{1}{N} \sum_{\alpha \in G} \overline{\phi_s(\alpha)} \phi_r(\alpha) = \begin{cases} 1 & \text{if } r=s \\ 0 & \text{otherwise.} \end{cases}$$

This is true whenever the underlying integral domain is  $\mathbf{C}$ .

Convolution with respect to  $Z_N$  takes the form

$$(f *_{Z_N} g)(k) = \sum_{j=0}^{N-1} f(j)g(k-j), \quad k=0, \dots, N-1.$$

The minus sign indicates subtraction modulo  $N$ , i.e. refers to the operation of the underlying group. This is often called circular or wrapped convolution since  $Z_N$  is a simple cyclic group (c.f. Fig. 4.8).

A matrix interpretation of this DFT and convolution may help to clarify the concepts involved. While it is possible to represent the operator  $F_G$  as a matrix for any finite abelian group  $G$ , we will only show what happens for this special case. A function  $f: Z_N \rightarrow \mathbf{C}$  is (naturally isomorphic to) a vector  $(f(0), f(1), \dots, f(N-1))$  in  $\mathbf{C}^N$ .  $F_{Z_N}$  can therefore be expressed as an  $N \times N$  matrix which we'll call  $U$  having entry  $e^{2\pi i k j / N}$  in the  $k^{\text{th}}$  row and  $j^{\text{th}}$  column. The inverse transform is represented by  $\frac{1}{N} U^\dagger$  where  $U^\dagger$  is the conjugate transpose of  $U$ . For example we have for  $N = 2$ :

$$U = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{and} \quad \frac{1}{2} U^\dagger = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} ;$$

and for  $N = 4$ :

$$U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix} \quad \text{and} \quad \frac{1}{4} U^\dagger = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix} .$$

If the rows and columns of  $U$  are labeled so that by the  $s^{\text{th}}$  row of  $U$  we mean the same thing as the  $s^{\text{th}}$  coordinate of a vector in  $\mathbb{C}^N$ , then  $U$  is an  $N \times N$  matrix whose  $s^{\text{th}}$  row is the character of  $\phi_s$  viewed as a vector in  $\mathbb{C}^N$ . The matrix  $\frac{1}{N} U^{-1}$  represents a change of basis for  $\mathbb{C}^N$  in which the new basis vectors are the columns of  $\frac{1}{N} U^{-1}$ . Thus, the  $s^{\text{th}}$  basis element is the normalized group character  $\frac{1}{N} \bar{\phi}_s$ .

We now give a concrete matrix interpretation of the convolution theorem which makes that result particularly easy to understand.

Suppose  $N = 4$ ,  $f = (f(0), \dots, f(3))$ ,  $g = (g(0), \dots, g(3))$ , and  $h = f *_{\mathbb{Z}_4} g$ . Convoluting  $g$  with any fixed  $f$  is a linear operation on  $\mathbb{C}^4$  and thus has a matrix representation which we'll call  $A_f$ . It's easy to check that in this case  $A_f$  is a  $4 \times 4$  circulant (see Ex. 4.5):

$$A_f = \begin{bmatrix} f(0) & f(3) & f(2) & f(1) \\ f(1) & f(0) & f(3) & f(2) \\ f(2) & f(1) & f(0) & f(3) \\ f(3) & f(2) & f(1) & f(0) \end{bmatrix} .$$

We then have that  $h = A_f g$ . Theorem 5.2 says that  $U$  that pointwise multiplication by any fixed  $Uf$  is all on  $\mathbb{C}^4$  and has a diagonal matrix representation  $D_f$  with the function (vector)  $Uf$  along the diagonal, i.e.

$$D_f = \begin{bmatrix} (Uf)(0) & 0 & 0 & 0 \\ 0 & (Uf)(1) & 0 & 0 \\ 0 & 0 & (Uf)(2) & 0 \\ 0 & 0 & 0 & (Uf)(3) \end{bmatrix} .$$

Using this observation, Theorem 5.2 says that  $UA_f g = D_f U g$  or; recalling that  $U$  is invertible,  $A_f g = U^{-1} D_f U g = \frac{1}{4} U^\dagger D_f U g$ .

Thus, the convolution theorem in this case says that for any  $f \in \mathbb{C}^4$ ,  $A_f$  is similar to a diagonal matrix  $D_f$  under the similarity transformation represented by  $U$ :  $A_f = U^{-1} D_f U$ . The significant point, however, is that the same change of basis diagonalizes the matrix  $A_f$  for all vectors  $f$ . In other words, all matrices that have a structure like  $A_f$  can be diagonalized by the same change of basis. The particular entries in the matrix are not important, only their arrangement or pattern.

### Example 5.2

Let  $R = \mathbb{C}$  and let  $G$  be finite abelian group which decomposes so that

$$G \approx \mathbb{Z}_{n_1} \oplus \dots \oplus \mathbb{Z}_{n_d} .$$

Writing elements of  $G$  as  $d$ -tuples of integers, we have that

$$\hat{f}(k_1, \dots, k_d) = \sum_{j_1=0}^{n_1-1} \dots \sum_{j_d=0}^{n_d-1} f(j_1, \dots, j_d) e^{2\pi i \sum_{\ell=1}^d k_\ell j_\ell / n_\ell}$$

for all  $(k_1, \dots, k_d) \in G$

and

$$f(j_1, \dots, j_d) = \frac{1}{\prod_{\ell=1}^d n_\ell} \sum_{k_1=0}^{n_1-1} \dots \sum_{k_d=0}^{n_d-1} \hat{f}(k_1, \dots, k_d) e^{-2\pi i \sum_{\ell=1}^d k_\ell j_\ell / n_\ell}$$

for all  $(j_1, \dots, j_d) \in G$ .

This is the general form of what is usually called the d-dimensional DFT. Convolution in this case takes the following form:

$$[f *_G g](k_1, \dots, k_d) = \sum_{j_1=0}^{n_1-1} \dots \sum_{j_d=0}^{n_d-1} f(j_1, \dots, j_d) g(k_1-j_1, \dots, k_d-j_d)$$

for all  $f, g \in F(F, \mathbb{C})$  and  $(k_1, \dots, k_d) \in G$

where  $k_i - j_i$  is subtraction modulo  $n_i$ .

### Example 5.3

The previous three examples involved the infinite field  $\mathbb{C}$ . However, the requirement that  $R$  contain an  $m^{\text{th}}$  principle root of unity does not imply that  $R$  must be infinite. Let  $R$  be the finite field of characteristic  $p$  where  $p$  is prime. Then  $R \approx \mathbb{Z}_p$  and contains a principle  $(p-1)^{\text{th}}$  root of unity which we'll call  $\omega$  (see Nicholson [1971]).

It's easy to see that if  $m$  divides  $p-1$ ,  $\mathbb{Z}_p$  contains a principle  $m^{\text{th}}$  root of unity, namely  $\omega^{(p-1)/m}$ . Thus,  $F_G$  and  $F_G^{-1}$  exist when  $R = \mathbb{Z}_p$  and  $G$  has exponent  $m$  that divides  $p-1$ . In particular, for  $R = \mathbb{Z}_p$  and  $G = \mathbb{Z}_n$  Fourier analysis is applicable when  $n$  divides  $p-1$ . Nicholson [1971] points out the computational significance of these results. In some cases, it is possible to compute the DFT and convolution exactly using integer arithmetic. This fact together with the fast algorithm for computing the DFT underlies the Schönhage-Strassen fast integer multiplication algorithm (see Knuth [1969] or Aho, Hopcroft, and Ullman [1974]).

For example, let  $R = \mathbb{Z}_5$  and  $G = \mathbb{Z}_4$ . It can be verified that 2 is a principle  $4^{\text{th}}$  root of unity in  $\mathbb{Z}_5$  so that the matrix representation

of the DFT (see Ex. 5.1) is as follows:

$$U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 3 \\ 1 & 4 & 1 & 4 \\ 1 & 3 & 4 & 2 \end{bmatrix}.$$

Its inverse  $U^{-1}$  is given by

$$U^{-1} = 4 \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 4 & 2 \\ 1 & 4 & 1 & 4 \\ 1 & 2 & 4 & 3 \end{bmatrix}$$

since  $4^{-1} = 4$  in  $Z_5$ .

#### Example 5.4

Suppose  $R = \mathbb{C}$  and  $G = (Z_2)^n$ . In this case the DFT is called the finite Walsh transform and is receiving more and more attention due to its very interesting properties (see Harmuth [1972]). The most suggestive fact about it is that each group character  $\phi_s$ , or Walsh wave, has a range consisting only of +1 and -1. This is true because the exponent of  $G$  is 2 and +1 and -1 are two distinct roots of unity. Since these functions can be generated easily by digital means (e.g. switching), standard signal processing techniques applied in such areas as radio, TV, radar, and spectroscopy are being recast in terms of this new underlying group with the hope that digital technology may be more easily applied.

The usual way of presenting the finite Walsh transform and the corresponding convolution theorem is through a binary representation of  $G$ . Letting  $Z_2 = \{0,1\}$  with module 2 addition (or, what is the same



thing, the logical operation of exclusive-or), the group  $G$  can be represented as the set of  $n$ -tuples of 0s and 1s with bit-wise exclusive-or as the group operation. This operation is often called dyadic addition and the group  $G$  is often called the dyadic group (Harmuth [1971]). For  $s, r \in G$ , let  $s \oplus r$  denote the dyadic sum of  $s$  and  $r$  and note that dyadic subtraction is the same as dyadic addition. Then dyadic convolution is the same as dyadic correlation and can be written as follows: for  $f, g \in F(G, \mathbb{C})$

$$(f *_G g)(s) = \sum_{r \in G} f(r)g(s \oplus r), \text{ for all } s \in G.$$

Ordering the elements of  $G$  as if they were binary representations of integers, the matrix representation of the Walsh transform for  $n = 2$  is

$$U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} .$$

For any  $n$ , the entries in the matrix representation of the finite Walsh transform are all 1s or -1s and hence the matrix is real. Since the matrix for any DFT is symmetric, we have that

$$U^{-1} = \frac{1}{|G|} U^\dagger = \frac{1}{2^n} U .$$

A forward Walsh transform is the same (except for scale) as a backward one.

In most applications, the elements of  $G$  are viewed as encodings of the integers so that elements of  $F(G, \mathbb{C})$  are thought of as functions of the set  $\{0, 1, \dots, n-1\}$  with its usual interpretation as a set of time or space samples. The clearest example of this "straight-line" represen-

tation occurs in the continuous case. Although they are not instances of the results we have stated above, the following facts follow from the general theory of harmonic analysis on locally compact topological groups. Let  $L^2[0,1]$  denote the space of all square integrable functions (in the Lebesgue sense) on the real interval  $[0,1]$ . Each such function can be expanded as a (perhaps infinite) sum of Walsh functions. Convergence is according to the usual  $L^2$  norm. The first 4 Walsh functions are graphed in Fig. 5.1.

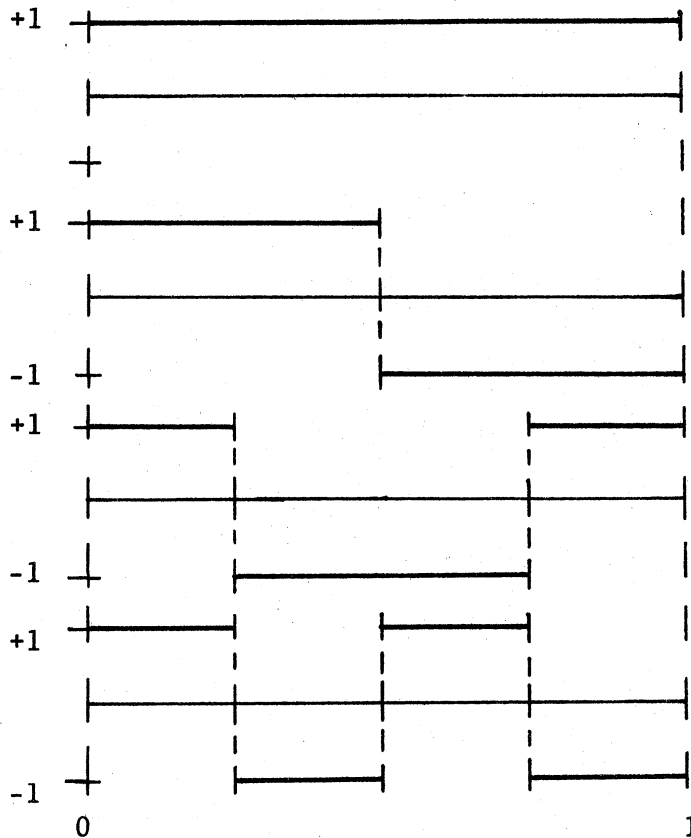


Figure 5.1: First 4 Walsh functions (of  $[0,1] \subset \mathbb{R}$ ).

Compare these graphs with the columns (and rows) of  $U$ . Although we graph these functions in the familiar way, they are functions of the group  $(\mathbb{Z}_2)^n$ . This group can be identified naturally with the interval  $[0,1]$  but has a different topology.

Since the usual group structure of the integers (or reals) often successfully models our notions of time or space, this "straight-line" view can lead to confusion. In particular, it might be felt that there is no convolution theorem for the Walsh transform. There is, of course, but the operation of dyadic addition does not model our usual view of translation in time or space. This does not mean, however, that the dyadic group, and hence dyadic convolution, may not provide a natural and useful model for real world phenomena.

Two other facts about the finite Walsh transform should be mentioned. First, there is a fast algorithm for computing it which is a special case of the usual FFT. Since the characters only have values  $+1$  and  $-1$  the algorithm need not involve complex arithmetic and is faster than the usual FFT. Second, although the term Walsh transform usually refers to cases in which the underlying integral domain is  $\mathbb{C}$ , the DFT on  $(\mathbb{Z}_2)^n$  works for other integral domains. In particular, it works for  $\mathbb{R}$  and, in accordance with the remarks in Ex. 5.3, for the field  $\mathbb{Z}_3$  consisting of only 3 elements.

## 5.6 Generalizations

It is possible to relax both the finiteness and commutativity conditions on the group  $G$  and still obtain results similar to those stated above. In fact, the theory of the DFT for  $R = \mathbb{C}$  is the simplest instance of a theory of harmonic analysis on locally compact topological

groups (see Hewitt and Ross [1963]). While we will make no attempt to precisely outline these generalizations, it is possible to mention two major differences that can occur in more complex situations.

First, removing the finiteness condition on  $G$  permits  $G$  and the character group to be non-isomorphic. In general, the character group is isomorphic to a group  $G^*$  known as the dual group of  $G$ . In case  $G$  is finite and abelian,  $G = G^*$ . This has important consequences since the Fourier transform (defined appropriately) of a function of  $G$  is a function of  $G^*$  (and, it turns out, vice versa, i.e.  $(G^*)^* = G$ ). An example is the case of the usual Fourier series for periodic functions of  $\mathbb{R}$ . Suppose, for simplicity, the period is 1. Then each function can be viewed as a function of  $\mathbb{R}/\mathbb{Z} = \{x+\mathbb{Z} \mid x \in \mathbb{R}\}$  which is sometimes called the circle group (see Lang [1967]). The group  $G$  is then  $\mathbb{R}/\mathbb{Z}$  but  $G^* \approx \mathbb{Z}$ . This corresponds to the fact that the spectrum of a periodic function must consist of discrete multiples of the fundamental frequency. In other words, the Fourier transform of a periodic function of  $\mathbb{R}$  is a function of  $\mathbb{Z}$  and hence the term "series" applies. Although  $\mathbb{R}/\mathbb{Z}$  is uncountable, its dual group is countable. (dually, the Fourier transforms of suitable functions of  $\mathbb{Z}$  are functions of  $\mathbb{R}/\mathbb{Z}$ ).

The second kind of complication encountered in the more general theory is that for non-abelian groups the Fourier transform of a scalar valued function is in general operator valued. In the matrix framework described in Ex. 5.1, this means that the similarity transformation given by  $U$  and  $U^{-1}$  does not necessarily diagonalize all the convolution matrices  $A_f$ . Instead of producing a matrix with single elements along the diagonal, it produces one with "blocks" along the diagonal. Therefore, the Fourier transform for non-abelian groups does not necessarily convert convolution into pointwise multiplication.

## 5.7 Parallel Decomposition of Linear Cellular Automata

In this section we interpret the application of Theorem 5.3 to linear cellular automata in the language of automata theory. This requires us to restrict attention to LCAs with finite abelian structure groups. These LCAs have finite dimensional state spaces and are therefore linear sequential machines (LSMs) in the usual sense (Harrison [1969]). Now, since most studies of LSMs do not contain discussions of module theoretic generalizations (i.e. where the field requirement is relaxed and rings considered instead), we also restrict attention to the application of the DFT based on an underlying field. Thus we will be discussing a subclass of LCAs as defined by Def. 4.1. The reader should note that the existence of the DFT and convolution theorem for certain rings as discussed in Sec. 5.2 and Sec. 5.3 leads to natural generalization of the ideas presented here. However, since our purpose is primarily integrative, we will not pursue the theory in its most general form.

A finite set of autonomous sequential machines,  $M_i = \langle Q_i, \delta_i \rangle$ ,  $i=1, \dots, N$ , can be regarded as a single sequential machine

$$M = \langle \prod_{i=1}^N Q_i, \prod_{i=1}^N \delta_i \rangle$$

where  $\prod_{i=1}^N \delta_i(q_1, \dots, q_N) = (\delta_1(q_1), \dots, \delta_N(q_N))$

for all  $(q_1, \dots, q_N) \in \prod_{i=1}^N Q_i$ .

We say that  $M$  is a parallel composition of the  $M_i$ . The machines  $M_i$  are just viewed as synchronously operating in parallel, i.e. they all change state at the same time and there are no connections between

them. We will write

$$M = \prod_{i=1}^N M_i .$$

These notions are usually defined for nonautonomous machines and our results are generalizable to such cases but will only consider the autonomous case here (see Hartmanis and Sterns [1966] or Harrison [1969]).

Since each state transition of an LCA is a convolution as indicated by (4.3-3), the convolution theorem can be applied (in the form given by Theorem 5.3). The resulting pointwise multiplication of functions is the same as the transition function of a parallel composition of simple automata. "Pointwise" and "parallel" mean the same thing: there are no interconnections between the components. This is a simple observation, but we state it as a theorem to precisely connect the terminology of the functional view with that of automata theory.

#### Theorem 5.4

Let  $M = \langle S, K^n, H \rangle$  be an LCA where  $S$  is a finite abelian group with exponent  $m$  and  $K$  is a field containing a principle  $M^{\text{th}}$  root of unity. Then  $M$  is isomorphic to a parallel composition  $\prod_{s \in S} M_s$  where for each  $s \in S$ ,  $M_s = \langle Q_s, \delta_s \rangle$  with  $Q_s = K^n$  and  $\delta_s = \hat{H}(s)$ . That is,  $M_s$  is an autonomous LSM with an  $n \times n$  transition matrix given by  $(\hat{h}_{ij}(s))$  where the  $h_{ij}$  are the impulse reponse functions of  $M$  (see 4.2-4).

Proof:

If the elements of  $S$  are put into a standard order, say  $S = \{s_1, \dots, s_N\}$  where  $|S| = N$ , then the state set of  $\prod_{s \in S} M_s$  can be naturally identified with  $F(S, K^n)$  by the correspondence  $(q_1, \dots, q_N) \mapsto q \in F(S, K^n)$  where  $q(s_i) = q_i \in K^n$  for  $i=1, \dots, N$ . Under this identification the transition

function  $\hat{\delta} = \prod_{s \in S} \delta_s$  of  $\prod_{s \in S} M_s$  is given by  $(\hat{\delta}q)(s) = \delta_s(q(s)) = \hat{H}(s)q(s)$ , for all  $s \in S$ , or simply  $\hat{\delta}(q) = \hat{H} \cdot q$ .

If  $\delta$  is the global transition function of  $M$ , we know from (4.2-3) that

$$\delta(q) = H *_S q \quad \text{for all } q \in F(S, K^n).$$

Since  $S$  and  $K$  are such that the DFT  $F_S$  exists, the convolution theorem (for vector-valued functions) says that

$$(H *_S q)^\wedge = \hat{H} \cdot \hat{q} \quad \text{for all } q \in F(S, K^n).$$

In other words

$$F_S \delta(q) = \hat{\delta}(F_S q) \quad \text{for all } q \in F(S, K^n)$$

so that  $F_S$  is an automaton-homomorphism. Since we also know that the DFT is invertible,  $F_S$  (for vector valued functions) is the required isomorphism.

Q.E.D.

This result is formally identical to the corresponding theorem in which the group  $S$  is taken to represent time instead of space. This theorem forms the basis of frequency domain analysis for time invariant linear systems with  $n$  input and  $n$  output wires (see for example, Chen [1970] or Schwartz and Friedland [1965]). It should be noted, however, that the usual application of these results in linear systems theory produces an isomorphism of systems in the I/O sense. In our application, the Fourier transform is a state isomorphism. This difference allows concepts from systems theory and facts from harmonic analysis to make contact in novel ways. For example, in the next chapter we will discuss how filtering techniques can be used to produce homomorphic models of LCAs.

For the sake of further integrating automata theoretic terminology, we note that the parallel composition  $\prod_{s \in S} M_s$  given in Theorem 5.4 realizes the LCA  $M$  through the assignment  $F_S$ . Therefore it is consistent with the usage of the term decomposition in automata theory (Hartmanis and Sterns [1966]) to say that  $\prod_{s \in S} M_s$  is a parallel decomposition of the LCA  $M$ . The following two definitions provide some notational conveniences we will find useful in the succeeding chapters.

Definition 5.2

An LCA  $M = \langle S, K^n, H \rangle$  is decomposable if it satisfies the hypotheses of Theorem 5.4.

Definition 5.3

If  $M = \langle S, K^n, H \rangle$  is a decomposable LCA, then the parallel decomposition given by Theorem 5.4 will be denoted by  $\hat{M} = [S, K^n, \hat{H}]$ .

In what following chapters, we will assume the natural correspondence described in the proof of Theorem 5.4 and take the state set of  $\hat{M}$  to be  $F(S, K^n)$ . Also, for  $W$  any subset of  $S$ , we will take the state set of  $\prod_{s \in W} M_s$  to be  $F(W, K^n)$  so that its transition function can be given by  $(\hat{H}|_W) \cdot q$  for  $q \in F(W, K^n)$ .



## 5.8 Examples of Parallel Decomposition

Three examples of the application of Theorem 5.4 follow. The first two are based on the LCAs of Ex. 4.3 and Ex. 4.4, i.e. on the finite difference approximations to the heat equation and simplified wave equation. In these cases, the parallel decomposition of the LCAs corresponds to a well-known technique for stability analysis of finite difference schemes.

The third example is a simple system that has been used to model an aspect of morphogenesis. We present a well-known observation about this system as a special case of our theory of LCAs. The general context thus provided allows interesting connections to be made.

### Example 5.5 Heat equation

The LCA  $M = \langle Z, \mathbb{R}, \omega \rangle$  given in Ex. 4.3 as a representation of a finite difference scheme for the heat equation is not decomposable since  $Z$  is infinite and  $\mathbb{R}$  contains only two principle roots of unity:  $+1$  and  $-1$ . Therefore, consider the LCA  $M' = \langle Z_N, \mathbb{C}, \omega' \rangle$  where  $\omega'(s \bmod N) = \omega(s)$  for all  $s \in Z$ , i.e.  $M'$  has a cyclic structure group, a complex cell state set, and the same weighting functions as  $M$  but viewed as functions of  $Z_N$ . Except for their complex state sets, the cells of  $M'$  are the same as the cells of  $M$ , but  $M'$  consists of  $N$  of them arranged in a loop. Since  $\mathbb{C}$  contains a principle  $N^{\text{th}}$  root of unity for any  $N$  we know that  $M'$  is decomposable.

If we relax, for the moment, the requirement that the state set of  $M$  consist only of configurations with finite support, then we can formalize the relationship between  $M'$  and  $M$  as follows: a submachine of  $M'$  is isomorphic to a submachine of  $M$ . Since  $\omega$  has real values

$$\frac{k\Delta t}{\Delta x^2}, 1 - \frac{2k\Delta t}{\Delta x^2}, \frac{k\Delta t}{\Delta x^2}, \text{ and } 0,$$

a real valued initial configuration of  $M'$  will generate only cell states with zero imaginary part. Thus there is a submachine of  $M'$  with global state set  $F(Z_N, \mathbb{R})$ . Call this submachine  $\text{Re}M'$ .

The submachine of  $M$  to which  $\text{Re}M'$  is isomorphic has state set consisting of configurations (i.e. functions of  $Z$ ) that are periodic with period  $N$ . To see this consider an operator from  $F(Z_N, \mathbb{R})$  into  $F(Z, \mathbb{R})$  that periodically extends a function of  $Z_N$  into a function of  $Z$ . Think of "unrolling" the cyclic group so that a function's values on  $Z_N$  are repeated in a periodic way for all of  $Z$ . Clearly, then, the range of this operator consists of all functions of  $Z$  with period  $N$ . Since  $M$  and  $M'$  are spatially uniform and have the same kind of cells, it's not hard to convince yourself that this periodic extension operator is an isomorphism from  $\text{Re}M'$  to the submachine of  $M$  whose state set consists of configurations with period  $N$ . Call this submachine  $P_N M$ . We have that  $\text{Re}M' \xrightarrow{\cong} P_N M$ . We've relaxed the finite support assumption since the only function of  $Z$  with finite support and period  $N$  is identically equal to 0.

With this relationship in mind, the behavior of  $M$  can be studied for certain initial configurations by examining the behavior of  $M'$ . Since  $M'$  is decomposable, information about its behavior can be obtained by studying the behavior of its parallel decomposition  $\hat{M}' = [Z_N, \mathbb{C}, \hat{H}']$ . The importance of these relationships is that  $\hat{M}'$  is very simple to analyze since there are no interconnections between its components. Fig. 5.2 shows the LCA  $M'$  and its parallel decomposition  $\hat{M}'$  for the case  $N = 4$ .

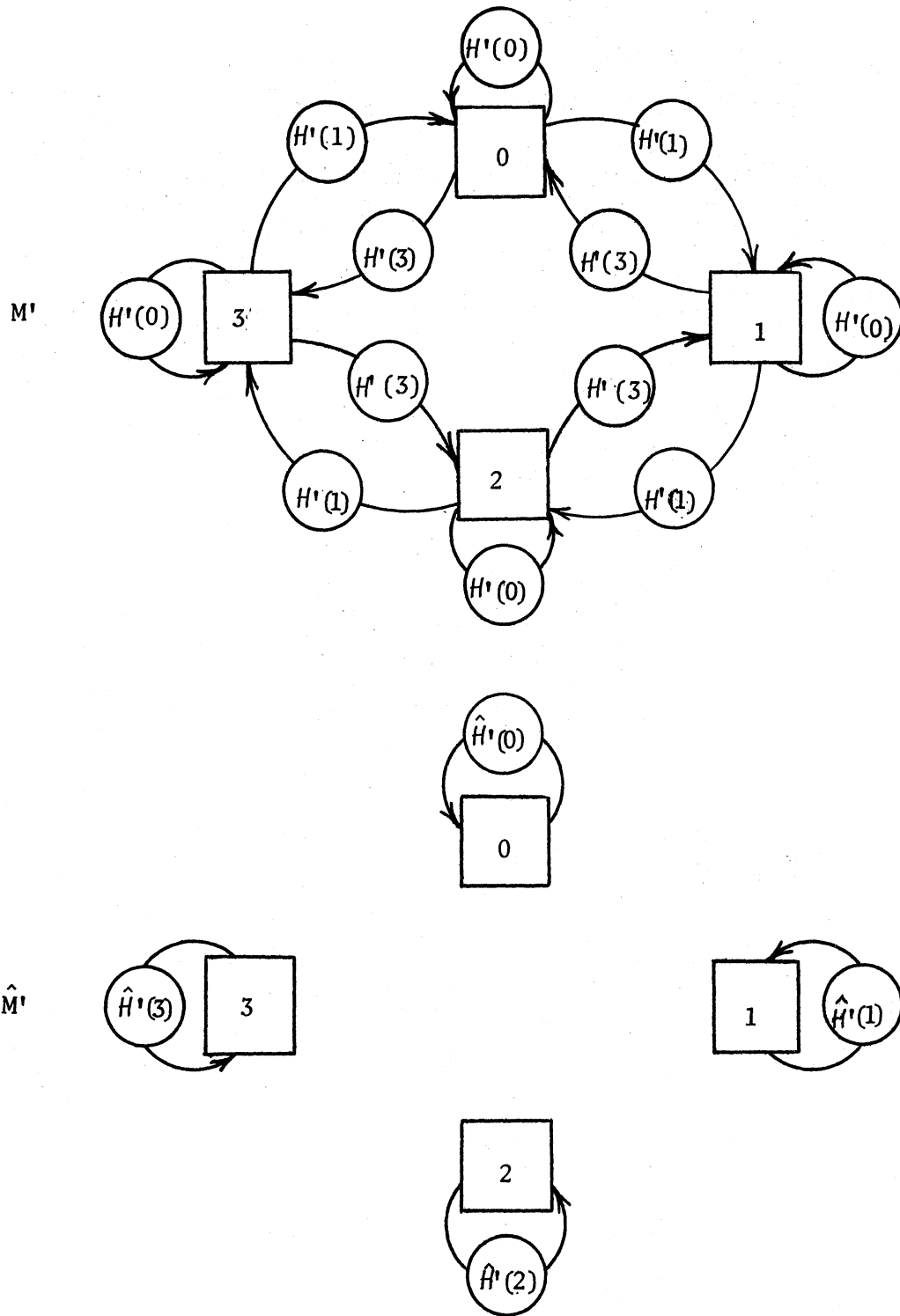


Figure 5.2: The LCA  $M'$  and its parallel decomposition  $\hat{M}'$  for  $N = 4$ .

Notice that even though we've arranged its components in a circle,  $\hat{M}'$  is not an LCA since the values of  $\hat{H}'$  will in general all be different, i.e. it is not uniform. The circular arrangement is used only to suggest the fact that the function  $\hat{H}'$  and the states of  $\hat{M}'$  are functions of  $Z_4$ .

From the formula for the DFT given in Ex. 5.1, the function  $\hat{H}'$  can be computed. Using some elementary facts about complex exponentials, the following expression is obtained:

$$\hat{H}'(s) = 1 - \frac{2k\Delta t}{\Delta x^2} + \frac{2k\Delta t}{\Delta x^2} \cos(2\pi s/N) \quad \text{for all } s \in Z_N. \quad (5.8-1)$$

The cosine is the usual one evaluated at the integers. Although the function  $H'$  (which is the same as  $W'$  since it is an even function) and  $\hat{H}'$  technically have complex values, their imaginary parts are zero.

Fig. 5.3 shows the real parts of these functions for  $N = 16$  and  $\frac{k\Delta t}{\Delta x^2} = \frac{1}{2}$ . We've graphed these functions using a convention similar to that used for functions of  $\mathbb{R}$ , but it should be remembered that they are functions of a discrete variable and need not be thought of as related in any way to continuous functions. In particular, they are not discontinuous functions of  $\mathbb{R}$ .

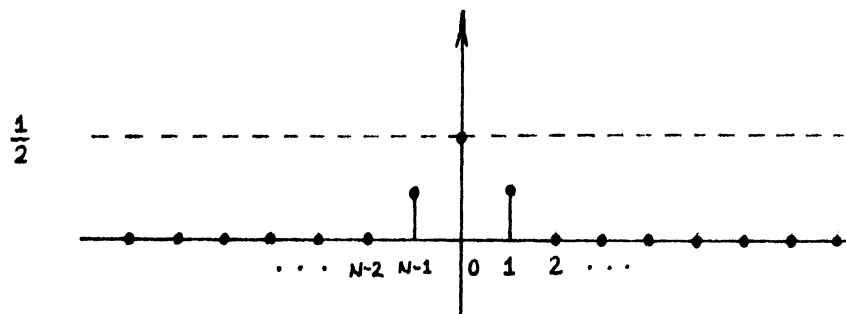
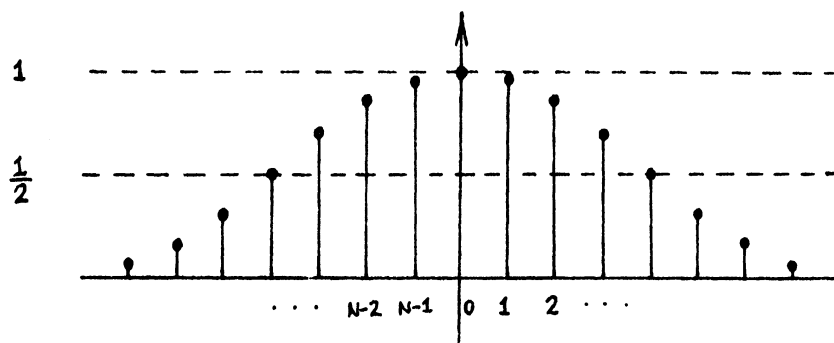
a) Real part of  $H'$ .b) Real part of  $\hat{H}'$ .

Figure 5.3: The functions  $H'$  and  $\hat{H}'$  for  $N = 16$  and  $\frac{k\Delta t}{2} = \frac{1}{2}$ .

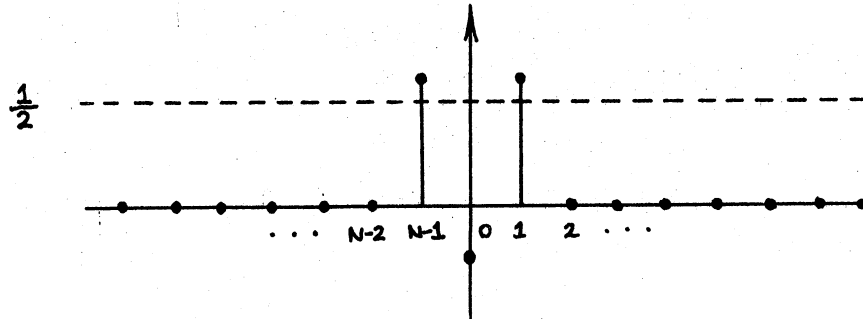
In numerical analysis the decomposition  $\hat{M}'$  is often used to determine the stability of the finite difference scheme to which the LCA  $M'$  corresponds (see Duff and Naylor [1966] or Gary [1972]). To see how this is done, note that the components of  $\hat{M}'$  behave very simply: if  $\delta_s$  is the transition function of component  $s$ , we have for all  $q \in F(Z_N, \mathbb{C})$  that

$$\delta_s^t(q(s)) = [H'(s)]^t(q(s)) \quad \text{for } t = 0, 1, 2, \dots \quad (5.8-2)$$

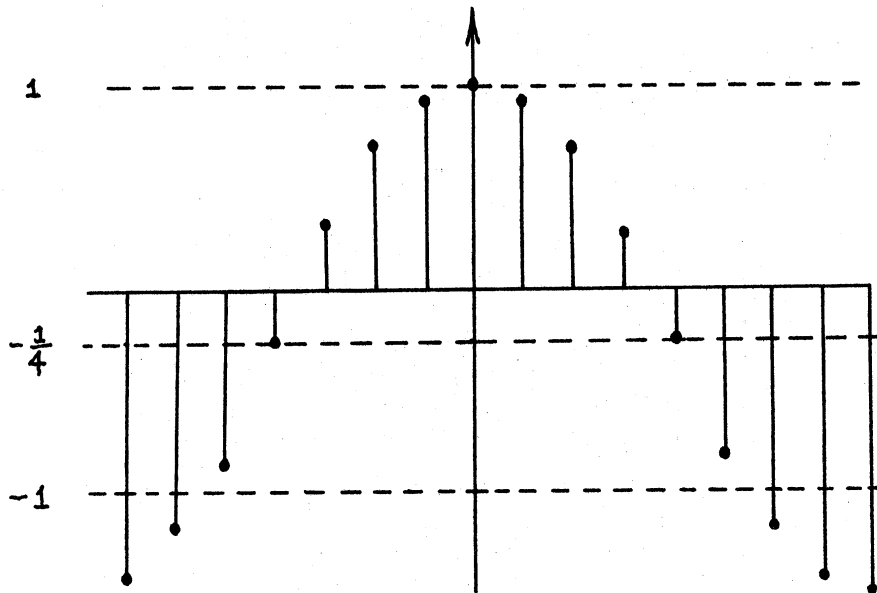
Thus, for  $M'$  to be stable it is necessary and sufficient that

$$|H'(s)| \leq 1 \quad \text{for all } s \in Z_N. \quad (5.8-3)$$

If this is not true for an  $\alpha \in Z_N$ , then the state of component  $\alpha$  in the



a) Real part of  $H'$ .



b) Real part of  $\hat{H}'$ .

Figure 5.4: The functions  $H'$  and  $\hat{H}'$  for  $N = 16$  and  $\frac{k\Delta t}{\Delta x^2} = \frac{5}{8}$ .

parallel decomposition  $\hat{M}'$  will grow without bound, and we say that  $\hat{M}'$  is unstable. When  $\hat{H}'(\alpha) < -1$ , not only does the magnitude of a state grow, but its sign changes at every time step. Using expression (5.8-1), condition (5.8-3) can be written

$$\left| 1 - \frac{2k\Delta t}{\Delta x^2} + \frac{2k\Delta t}{\Delta x^2} \cos(2\pi s/N) \right| \leq 1 \quad \text{for all } s \in Z_N.$$

•  $k \geq 0$  this reduces to the condition

$$\frac{k\Delta t}{\Delta x^2} < \frac{1}{2} \quad (5.8-4)$$

since  $|\cos(2\pi s/N)| \leq 1$  for all  $s \in Z_N$ .

The functions graphed in Fig. 5.3 show the situation for  $\frac{k\Delta t}{\Delta x^2} = \frac{1}{2}$  and one can see that the values of  $\hat{H}'$  are all between +1 and -1. Fig. 5.4, on the other hand, shows what can happen when (5.8-4) is not satisfied. There are five values of  $\hat{H}'$  that are less than -1. The corresponding five components of  $\hat{M}'$  are therefore unstable, and hence we say that the entire machine  $\hat{M}'$  is unstable. Since  $M'$  and  $\hat{M}'$  are isomorphic, this implies the instability of  $M'$ .

It is not the case, however, that all submachines of  $M'$  are unstable. Let  $Y \subseteq S$  be such that  $|\hat{H}'(s)| > 1$  if and only if  $s \in Y$ . If  $M'$  is started in a state  $q$  such that  $\hat{q}(s) = 0$  for  $s \in Y$ , then for all  $t \in Z^+$ ,  $\delta^t(q)$  has the same property, i.e.  $[\delta^t(q)]^{\hat{}}(s) = 0$  for  $s \in Y$ . The set of states with this property is thus the state set of a submachine of  $M'$ , and this submachine is clearly stable. However, the submachine  $\text{Re}M'$  of  $M'$  with state set consisting of real valued configurations is not stable when (5.8-4) fails. Since  $\text{Re}M' \iff P_N M'$ , we can conclude that condition (5.8-4) also guarantees the stability of  $P_N M'$ . It also guarantees the stability of the entire system  $M$ , but we have not developed the

framework necessary to show this.

Example 5.6 Simplified wave equation

The LCA  $M = \langle \mathbb{Z}, \mathbb{R}, H \rangle$  discussed in Ex. 4.4 as a finite difference for a simplified wave equation is not decomposable for the same reasons the LCA for the heat equation isn't: it has an infinite structure group and  $\mathbb{R}$  doesn't contain enough roots of unity. However, the remarks made above for the approximation to the heat equation hold here as well, and we'll consider an LCA  $M' = \langle \mathbb{Z}_N, \mathbb{C}, H' \rangle$  that has the same kind of cells as  $M$  but a cyclic structure group and complex cell state set.  $M'$  is decomposable and Fig. 5.5 shows  $M'$  and its parallel decomposition  $\hat{M}' = [\mathbb{Z}_N, \mathbb{C}, \hat{H}']$  for  $N = 4$ . The impulse response function  $H$  is simply  $(0, 1, 0, 0)$ , and if you refer to the matrix  $U$  of the DFT for  $N = 4$  given in Ex. 5.1, you will see that  $\hat{H} = (1, i, -1, -i)$ , i.e. it is the second row of  $U$ .

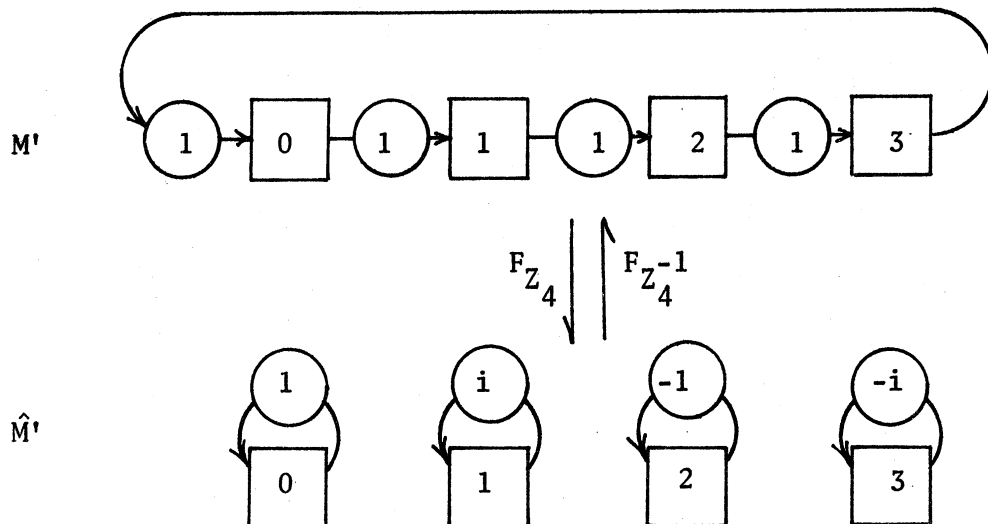
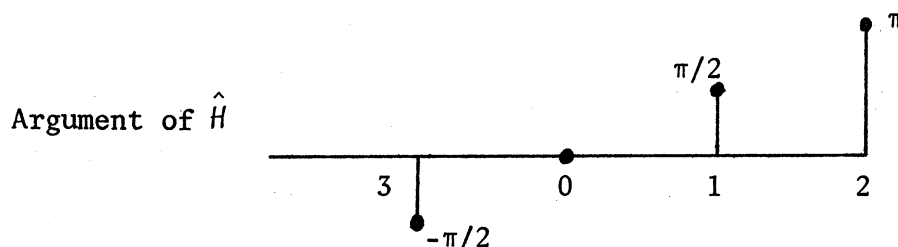


Figure 5.5: The LCA  $M'$  and its parallel decomposition  $\hat{M}'$  for  $N=4$ .



An understanding of the relationship between this  $M'$  and  $\hat{M}'$  will go far toward clarifying the nature of the spatial and the spatial frequency representations of a decomposable LCA. In fact, an intuitive understanding of these systems constitutes an essential step in understanding much of what mathematics has to say about PDEs and wave phenomena. Therefore, despite their simplicity, we will discuss the systems of Fig. 5.5 in some detail.

First, it is obvious what kind of behavior  $M'$  exhibits: it's a circular shift register. Any initial configuration will simply circulate around the loop without changing size or shape. The behavior of  $\hat{M}'$  is simple also since its components are independent. Each component is an oscillator with a different frequency. Since when complex numbers are multiplied their arguments add and their moduli multiply (De Moivre's Theorem), the behavior of component  $\alpha$  of  $\hat{M}'$  can be seen directly from the argument and modulus of  $\hat{H}(\alpha)$ : the argument gives the frequency, and the modulus determines whether the oscillation grows or decays in amplitude. In this case all the moduli are 1 so that the components produce undamped oscillations. A graph of the arguments of the coefficients looks like this:



It turns out that this graph is linear for any number of cells (even an uncountably infinite number). Thus, for  $N = 4$ , the frequencies are such that component 0 simply stores its initial state without change for all  $t \in \mathbb{Z}^+$ . The state of component 2 changes sign at each time step, and components 1 and 3 return to their initial state after every 4<sup>th</sup> time step.

How are the very different behaviors of  $M'$  and  $\hat{M}'$  related? Suppose  $M'$  is started in initial configuration  $\frac{1}{N}\phi_{-\alpha}$ ,  $\alpha \in \mathbb{Z}_4$ , i.e. its initial configuration is a normalized group character. According to Theorem 5.4 and the definition of  $F_{\mathbb{Z}_4}$ , the corresponding initial state of  $\hat{M}'$  is  $F_{\mathbb{Z}_4}(\frac{1}{N}\phi_{-\alpha}) = (0, \dots, 1, \dots, 0)$  where the 1 is in the  $\alpha^{\text{th}}$  position. Then the next state of  $\hat{M}'$  is  $(0, \dots, \hat{H}'(\alpha) \cdot 1, \dots, 0) = (0, \dots, \hat{H}'(\alpha), \dots, 0)$ . Since  $M'$  and  $\hat{M}'$  are isomorphic, this means that the next configuration of  $M'$  is  $F_{\mathbb{Z}_4}^{-1}(0, \dots, \hat{H}'(\alpha), \dots, 0) = \hat{H}'(\alpha) \cdot \frac{1}{N}\phi_{-\alpha} = \frac{1}{N}\hat{H}'(\alpha) \cdot \phi_{-\alpha}$ . In other words, the group characters are eigenvectors of the linear global transition function of  $M'$ . The corresponding eigenvalues are the values of  $\frac{1}{N}\hat{H}'$ . Since any configuration can be represented as a linear combination of the group characters, and since the transition function of  $M'$  is linear, it is possible to find the coefficients in the linear combination (i.e. compute  $F_S$ ) and multiply each by the appropriate value of  $\frac{1}{N}\hat{H}'$ . The resulting values are used as the coefficients in the group character expansion of the next state of  $M'$  (i.e.  $F_S^{-1}$  is computed). The transition function of  $\hat{M}'$  is just the diagonalized representation of the global transition function of  $M'$  (see Ex. 5.1).

But how are the group characters shifted (i.e. the action of  $M'$ )

by means of multiplying them by constants? Suppose  $M'$  is in configuration  $\phi_2 = (1, -1, 1, -1)$ . Then the next configuration is this function shifted (circularly) to the right, i.e.  $(-1, 1, -1, 1)$ . But this is  $-\phi_2$ . Similarly,  $i(1, -i, -1, i) = (i, 1, -i, -1)$ , etc. Fact (5) of Sec. 5.2 expresses this more simply: the characters are eigenfunctions of the shift operators  $L_\alpha$ ,  $\alpha \in S$ . Here, the global transition function of the LCA  $M'$  is  $L_{-1}: F(Z_4, \mathbb{C}) \rightarrow F(Z_4, \mathbb{C})$ . Group characters in  $\mathbb{C}$  (at least for simply generated groups) can be viewed as helices (in the space  $S \times \mathbb{C}$ ) whose "pitch" is given by their frequency. Multiplying by a complex constant can be thought of as a "twist" of the helix and, therefore, results in a translation. The linear graph of the argument of  $\hat{H}'$  reflects the fact that higher pitched helices need to be twisted more to produce the same lateral translation as their lower pitched counterparts. If a value of  $\hat{H}'$  has modulus not equal to 1, the corresponding helix will shrink ( $<1$ ) or grow ( $>1$ ) in diameter as it is twisted (cf. the previous example where, in addition, the argument of each value of  $\hat{H}'$  is zero so that no twists are involved). Although such imagery is not directly applicable in the more abstract cases (e.g. Walsh functions), it captures all the essential intuition needed to understand the relationship between a decomposable LCA and its parallel composition.

Example 5.7 Rashevsky-Turing morphogenic model

Rosen [1970] discusses a simple example of a class of models proposed by Rashevsky in 1940 and by Turing in 1952 as possible mechanisms underlying morphogenic processes. The example provides one reasonable view of how an initially homogeneous collection of cells can give rise to cellular differentiation.

We will present this model in an extremely simple version that retains only the barest essentials necessary to produce the required behavior. The system consists of two cells which are thought of as actual biological cells. The state of each cell is characterized by the concentration of a substance, called a morphogen, which is taken to determine the characteristics of a cell (the model Rosen discusses has two morphogens per cell). Suppose each cell produces the morphogen at a particular rate and, depending on the parameter values chosen, the morphogen either diffuses or is actively transported to the other cell.

Let us assume that a cell can manufacture and store an arbitrary amount of morphogen so that we can use the following simple model:

let  $M = \langle \mathbb{Z}_2, \mathbb{R}, H \rangle$  be the LCA with  $H(0) = a$  and  $H(1) = b$ . This LCA is decomposable and Fig. 5.6 shows  $M$  and its parallel decomposition  $\hat{M}$ .

In this case the DFT has this matrix representation:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} .$$

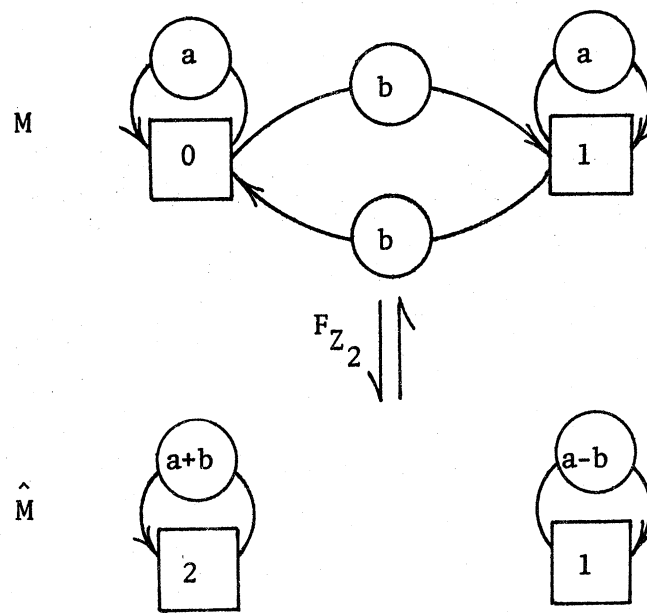


Figure 5.6: A two-celled LCA and its parallel decomposition

Thus, component 0 of  $\hat{M}$  keeps track of the total concentration of the morphogen in  $M$ , and component 1 keeps track of the difference between the concentrations in cell 0 and 1.

Suppose  $a+b = 1$  so that the total concentration remains unchanged, but let  $a-b > 1$ . Then, if the initial state of component 1 of  $\hat{M}$  is non-zero (say it is positive), the amplitude of the spatial frequency component  $\phi_1 = (1, -1)$  will increase without bound as  $M$  operates.

Using a term from the theory of signal processing, we can say that  $\phi_1$  is a frequency that resonates in  $M$ .

What does this spatial resonance mean in terms of the model's morphogenetic interpretation? First, note that any initial configur-

ation of concentrations that is a constant function (i.e. the same value for each cell) is an equilibrium configuration of  $M$  since the Fourier coefficient corresponding to  $\phi_1$  is 0. Thus, in the isomorphic system  $\hat{M}$ , the initial state of component 1 is 0 so that neither component changes state. However, if there is the slightest difference between the concentrations in cell 0 and 1 of  $M$ , this difference will resonate. In other words, a constant configuration is an unstable equilibrium of  $M$ .

Thus, the model shows that it is not logically impossible for an initially homogeneous structure to differentiate provided there is a source of outside disturbances however small. Note, though, that  $M$  remains uniform in the sense of our definition of a cellular automaton since the function  $H$  does not change. The differentiation occurs between the states of the cells. This has no bearing on the argument, however, since presumably biological cells in different states can exhibit different physical characteristics.

Finally, we would like to make the simple observation that in this modeling situation the instability of the system  $M$  is a desirable aspect of its behavior and has not been introduced by careless approximation techniques.

## Chapter 6

### Homomorphisms of Linear Cellular Automata by Spatial Frequency Filtering

#### 6.1 Introduction

The concept of an automaton homomorphism provides one suggestive idealization of what it means for one system to be a model of another. In this chapter we present a number of ways that homomorphisms of linear cellular automata can arise and interpret these homomorphisms in the context of modeling and simulation. The point of view taken is that when a model is formulated as an LCA, two related sets of issues arise. First, the original LCA proposed as a model might be so complex that it does little to increase our understanding of the real system. It might even be too complex to simulate on a computer due to storage and/or time limitations. Under these circumstances it is natural to ask how the original LCA can be simplified and what relationship can be expected to hold between the LCA and its simplified form. The second set of issues involves the justification for using the LCA formalism in specifying a model. While Chapter 8 deals more directly with the latter problems, the results developed here and in Chapter 7 provide insight into various simplification procedures and the kind of difficulties that can arise when they are not used carefully.

Following an intuitive discussion of automaton-homomorphisms, we define several operators that will be used to express relationships between LCAs and their simplifications. Sec. 6.4 is somewhat of a digression. It contains a result we will use later on in discussing homomorphisms, but the presentation was chosen to emphasize some ideas

that are less central to our line of development. Finally, we discuss the basic facts about LCAs and spatial frequency filtering. Although this chapter and the next contain numerous examples, the possibilities for system simplification are not adequately illustrated since it is impossible to draw diagrams for all but the least complex systems.

## 6.2 Homomorphisms

Let  $M = \langle Q, \delta \rangle$  and  $M' = \langle Q', \delta' \rangle$  be sequential machines. In Chapter 2 we said that if there is an onto map  $\xi: Q \rightarrow Q'$  (not necessarily one-one) such that  $\xi\delta(q) = \delta'(\xi(q))$  for all  $q \in Q$ , then  $M'$  is a homomorphic image of  $M$ . The map  $\xi$  is called an automaton-homomorphism and induces an equivalence relation on  $Q$  such that  $q_1 \equiv q_2$  if and only if  $\xi(q_1) = \xi(q_2)$ . Since  $\xi$  is a homomorphism it is true that if  $q_1 \equiv q_2$  then  $\delta(q_1) \equiv \delta(q_2)$ , i.e. equivalence classes are mapped to equivalence classes by the transition function  $\delta$ . Recall our convention of writing  $M \xrightarrow{\xi} M'$  to show that  $\xi$  is a homomorphism from  $M$  to  $M'$ .

The notion of an automaton homomorphism, which we have presented in its simplest form, has been very suggestive both formally and intuitively in studies of the modeling and simulation process. The reader is referred to Ashby [1963], Burks [1973] and Zeigler [1970,1974] for more complete discussions. Here we would like only to bring out several intuitive aspects of the idea that are particularly well illustrated by the homomorphisms of LCAs that we discuss below.

Thinking of the sequential machine  $M$  as representing a "real" system, the homomorphic image  $M'$  can be viewed as a model of  $M$  that captures certain aspects of  $M$ 's behavior. Operating with map  $\xi$  causes certain distinctions which are made in the state set  $Q$  of  $M$  to be lost so that



$\xi(Q)$  is a view of  $Q$  that is "less distinct " in some sense. Each equivalence class produced by the equivalence relation discribed above consists of states of  $M$  that look the same when viewed through  $\xi$ , i.e. that are blurred to the same image. That  $\xi$  is a homomorphism means that the level of resolution of the image of any state of  $M$  is sufficiently detailed so that the next state of  $M$  can be predicted from it to the same level of resolution. In other words,  $\xi$  provides a less distinct view of  $M$ , but one that is not misleading as far as the behavior of  $M$  is concerned.

It turns out that for LCAs this metaphorical language can be taken quite literally. For a decomposable LCA  $M$ , since all the components in its decomposition  $\hat{M}$  operate independently of each other, the view of  $M$  resulting from "filtering out" any subset of these components is a homomorphic image of  $M$ . This homomorphic image is literally a low-resolution image of  $M$  in the spatial, visual sense. It's as if the LCA's behavior were being viewed through a fog. That the resulting image is a homomorphic image means that each blurry view of a configuration contains enough information so that the blurry view of the next configuration can be predicted exactly. An observer, compelled to watch only through the fog, could, therefore, formulate a valid deterministic model for what he was able to see.

Although these observations correspond to fairly elementary mathematical concepts, we feel that the connection provided to the concept of automaton-homomorphism, and hence to theoretical investigations of the modeling process, warrants the careful development that we provide in this chapter.

### 6.3 Restrictions, Projections, and Embeddings

In this chapter and the one that follows, we will be discussing various homomorphisms and isomorphisms between LCAs, between LCAs and parallel compositions, and between different parallel compositions. According to Def. 4.1 and the remarks at the end of Sec. 5.7, the state sets of these kinds of systems can be viewed as sets of functions. Thus, a map from one state set to another is an operator on a set of functions, i.e. it takes one function and produces another. An example that we have already seen is the Fourier transform as it is used in Theorem 5.4. We now define three other operators on functions, each much simpler than the Fourier transform, that will be used to relate the state sets of various machines.

Let  $S$  and  $I$  be sets such that there is an injection  $\rho$  of  $I$  into  $S$ , i.e.  $\rho: I \rightarrow S$  is one-one. Let  $V$  be a set with a distinguished element denoted by  $0$ .

#### Definition 6.1

The restriction w.r.t.  $\rho(I)$  is the operator  $\text{Res}_{\rho(I)}: F(S,V) \rightarrow F(I,V)$  given by  $\text{Res}_{\rho(I)} f = f\rho$  for  $f \in F(S,V)$ .

#### Definition 6.2

The projection w.r.t.  $\rho(I)$  is the operator  $\text{Proj}_{\rho(I)}: F(S,V) \rightarrow F(S,V)$  given by

$$(\text{Proj}_{\rho(I)} f)(s) = \begin{cases} f(s) & \text{for } s \in \rho(I) \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } s \in S.$$

#### Definition 6.3

The embedding w.r.t.  $\rho(I)$  is the operator  $\text{Emb}_{\rho(I)}: F(I,V) \rightarrow F(S,V)$

given by

$$(\text{Emb}_{\rho(I)} f)(s) = \begin{cases} f(\rho^{-1}(s)) & \text{for } s \in \rho(I) \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } s \in S.$$

Suppose  $S = \{0, 1, \dots, 7\}$  and  $I = \{x, y, z\}$ . Fig. 6.1 illustrates the action of the above three operators for the map  $\rho: I \rightarrow S$  shown at the top of the figure.

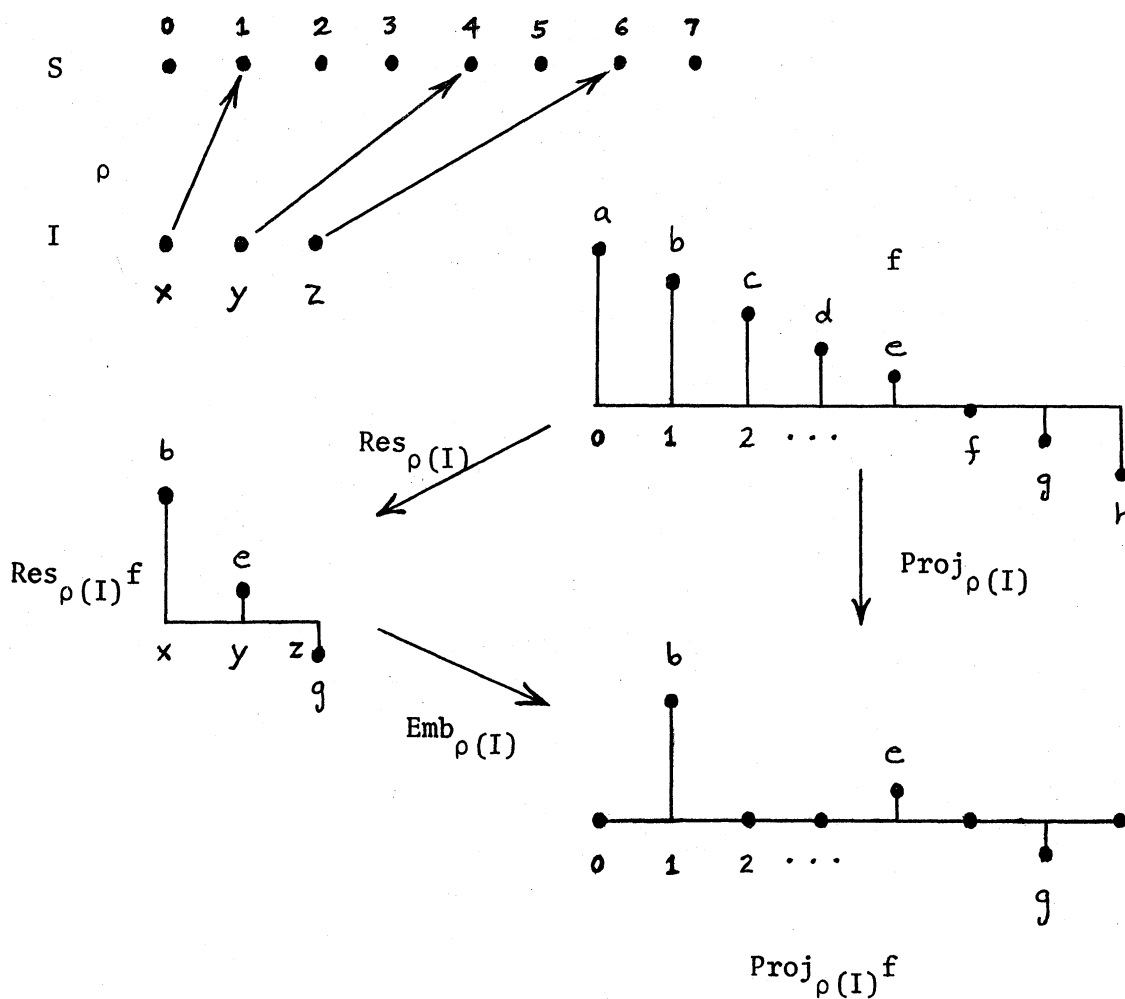


Figure 6.1: The action of the operators  $\text{Res}_{\rho(I)}$ ,  $\text{Proj}_{\rho(I)}$ , and  $\text{Emb}_{\rho(I)}$ .

It should not be difficult to see that for any injection  $\rho$  the following properties hold:

- 1)  $\text{Proj}_{\rho(I)} \text{Proj}_{\rho(I)} = \text{Proj}_{\rho(I)}$ ,
- 2)  $\text{Res}_{\rho(I)} \text{Proj}_{\rho(I)} = \text{Res}_{\rho(I)}$ , and
- 3)  $\text{Emb}_{\rho(I)} \text{Res}_{\rho(I)} = \text{Proj}_{\rho(I)}$  .

These operators are perhaps more familiar when  $\rho$  is an inclusion map (see Sec. 2.2). Suppose  $W$  is any subset of  $S$  and  $\rho:W \rightarrow S$  is the inclusion map. Then we have that  $\text{Res}_{\rho(W)} f = f|_W$  in the notation of Sec. 3.2; the projection operator is given by

$$(\text{Proj}_{\rho(W)} f)(s) = \begin{cases} f(s) & \text{for } s \in W \\ 0 & \text{otherwise} \end{cases} ;$$

and the embedding operator is given for  $f \in F(W,V)$  by

$$(\text{Emb}_{\rho(W)} f)(s) = \begin{cases} f(s) & \text{for } s \in W \\ 0 & \text{otherwise} \end{cases} .$$

When  $\rho:I \rightarrow S$  is not an inclusion map, the situation is exactly the same except the subset of  $S$  is  $\rho(I)$ , and we can label its elements with elements of  $I$ .

We will let the context indicate what set  $V$  denotes in particular instances: it will be either a field  $K$ , the vector space  $K^n$ , the vector space  $\text{Mat}_K(n,n)$ , an integral domain  $R$ , or the module  $R^n$ . In each case the distinguished element 0 will be the usual zero element of the space, i.e. the zero of  $K$ , the zero vector in  $K^n$ , the zero matrix in  $\text{Mat}_K(n,n)$ , etc. The reader can assume that unless otherwise indicated the same set  $S$ , set  $I$ , and injection  $\rho$  will underlie all usages of the symbols  $\text{Res}_{\rho(I)}$ ,  $\text{Proj}_{\rho(I)}$ , and  $\text{Emb}_{\rho(I)}$ . We will write  $\text{Res}_W$ ,  $\text{Proj}_W$ , and

$\text{Emb}_W$  in case  $W \subseteq S$  and  $\rho$  is the inclusion map. In many cases, such as in the proofs of theorems, it will be possible to completely drop the subscripts without causing confusion. When it is clear what  $I$  and  $\rho$  are, we will simply write  $\text{Res}$ ,  $\text{Proj}$ , and  $\text{Emb}$ .

We now establish some basic facts about the operator  $\text{Proj}$  and decomposable LCAs. Let  $M = \langle S, K^n, H \rangle$  be a decomposable LCA with the decomposition  $\hat{M} = [S, K^n, \hat{H}]$ . Let  $\hat{\delta}: F(S, K^n) \rightarrow F(S, K^n)$  be the transition function of  $\hat{M}$ , i.e.  $(\hat{\delta}q)(s) = \hat{H}(s) \cdot q(s)$ . Let  $\rho$  be an injection of a set  $I$  into  $S$ . Then from the fact that  $\hat{M}$  is a parallel composition where each component  $M_s$  is linear we have the following:

Lemma 6.1

$$\hat{\delta}(\text{Proj}_{\rho(I)}(F(S, K^n))) \subseteq \text{Proj}_{\rho(I)}(F(S, K^n)).$$

Proof:

Let  $q \in \text{Proj}_{\rho(I)}(F(S, K^n))$  and  $s \in S - \rho(I)$ . Then  $(\hat{\delta}q)(s) = \hat{H}(s) \cdot q(s) = \hat{H}(s) \cdot (0) = 0$ . Thus  $\hat{\delta}q \in \text{Proj}_{\rho(I)}(F(S, K^n))$ .

Q.E.D.

In other words, if the parallel composition  $\hat{M}$  is in a state at time  $t$  such that the components in the set  $S - \rho(I)$  are in state zero, then those components will be in state zero at time  $t+1$ . Thus, if  $M$  is started in such a state, those components in  $S - \rho(I)$  will stay in state zero as long as the machine operates. It therefore makes sense to define a submachine of  $\hat{M}$ :

Definition 6.4

$\hat{M}_{\rho(I)}$  is the submachine of  $\hat{M}$  with state set  $\text{Proj}_{\rho(I)}(F(S, K^n))$ .

From Theorem 5.4 we know that  $M \xrightarrow{F_S} \hat{M}$  so that there is a submachine of  $M$  isomorphic to  $\hat{M}_{\rho(I)}$ . In fact, this submachine can be defined as follows:

Definition 6.5

$M_{\rho(I)}$  is the submachine of  $M$  with state set  $F_S^{-1} \text{Proj}_{\rho(I)}(F(S, K^n))$ .

We have, then, that  $M_{\rho(I)} \xrightarrow{F_S} \hat{M}_{\rho(I)}$ . For  $I \subseteq S$  and  $\rho$  the inclusion map, we will write  $\hat{M}_I$  and  $M_I$ .

Intuitively, a state of  $M_{\rho(I)}$  is one whose Fourier transform is zero on the set  $S - \rho(I)$ . If  $\rho(I)$  is a set of "low frequencies" (c.f. the discussion in Sec. 5.2), then it is reasonable to say that the state set of  $M_{\rho(I)}$  consists of only "smooth" functions of  $S$ .

Finally, since the operator  $F_S^{-1} \text{Proj}_{\rho(I)} F_S$  will appear so frequently, we provide an abbreviation for it:

Definition 6.6

$\xi_{\rho(I)}$  is the operator given by  $F_S^{-1} \text{Proj}_{\rho(I)} F_S : F(S, V) \rightarrow F(S, V)$ .

Following our usual practice, the set  $V$  will be specified by the context so that different usages of  $\xi_{\rho(I)}$  may refer to different operators. For  $I \subseteq S$  and  $\rho$  the inclusion map, we will write  $\xi_I$ . Intuitively, this operation can be interpreted as a "filter": it filters out those frequencies in the set  $S - \rho(I)$ . We return to this idea in Sec. 6.5.

## 6.4 Distributed Realizations of Sets of LSMs

In Sec. 5.7 we pointed out that the decomposition of an LCA  $M$  could be considered a realization  $M$  through the assignment  $F_S$ . Since  $F_S$  is a bijection on the set  $F(S, K^n)$ , it is equally possible to consider the LCA a realization of its decomposition through the assignment  $F_S^{-1}$ . Moreover, under appropriate conditions, any parallel composition  $\prod_{i \in I} M_i$  of LSMs is a part of the decomposition of some LCA. We can therefore consider that LCA to be a realization of  $\prod_{i \in I} M_i$ . The next theorem precisely states this observation using the Emb operator.

Let  $I$  be any finite set, let  $A: I \rightarrow \text{Mat}_K(n, n)$ , and assume that  $|S| = N$ .

Theorem 6.1

Let  $\{M_i = \langle K^n, A(i) \rangle \mid i \in I\}$  be a finite set of autonomous LSMs and let  $\rho: I \rightarrow S$  be any injection of  $I$  into a finite abelian group  $S$ . If  $K$  contains a principle  $m^{\text{th}}$  root of unity where  $m$  is the exponent of  $S$ , then the LCA  $M = \langle S, K^n, H \rangle$  where  $H = F_S^{-1} \text{Emb}_{\rho(I)} A$  is a realization of the parallel composition  $\prod_{i \in I} M_i$  through the assignment  $F_S^{-1} \text{Emb}_{\rho(I)}$ .

Proof:

We need to show that  $\prod_{i \in I} M_i$  is isomorphic to a submachine of  $M$ . This submachine is  $M_{\rho(I)}$  (Def. 6.5) and the isomorphism is  $F_S^{-1} \text{Emb}_{\rho(I)}$ .

To see this, note that  $M$  is decomposable with decomposition  $\hat{M} = [S, K^n, \text{Emb}_{\rho(I)} A]$  having transition function  $\hat{\delta}$  given for  $q \in F(S, K^n)$  by

$$(\hat{\delta}q)(s) = (\text{Emb}_{\rho(I)} A)(s) \cdot q(s), \quad s \in S.$$

The transition function  $\delta$  of  $\prod_{i \in I} M_i$ , on the other hand, is given as follows (if we identify the state set of  $\prod_{i \in I} M_i$  with  $F(I, K^n)$ ): for  $f \in F(I, K^n)$

$$(\delta f)(i) = A(i) \cdot q(i), \quad i \in I.$$

Noticing that  $\text{Emb}: F(I, K^n) \rightarrow \text{Proj}(F(S, K^n))$  is a bijection, we have that

$$\prod_{i \in I} M_i \xrightleftharpoons{\text{Emb}} \hat{M}_{\rho(I)}$$

as the following argument shows:

i) for  $s \in \rho(I)$

$$\begin{aligned} (\text{Emb} \delta f)(s) &= \delta f_{\rho^{-1}(s)} = A(\rho^{-1}(s)) \cdot f(\rho^{-1}(s)) \\ &= (\text{Emb} A)(s) \cdot (\text{Emb} f)(s) \\ &= (\hat{\delta \text{Emb} f})(s), \quad \text{and} \end{aligned}$$

ii) for  $s \in S - \rho(I)$

$$(\text{Emb} \delta f)(s) = 0 = (\hat{\delta \text{Emb} f})(s).$$

Therefore we have

$$\prod_{i \in I} M_i \xrightleftharpoons{\text{Emb}} \hat{M}_{(I)} \xrightleftharpoons{F_S^{-1}} M_{(I)}$$

Q.E.D.

Note that the elements of any set of LSMs  $\{M_i = \langle K^{n_i}, A_i \rangle \mid A_i \in \text{Mat}_K(n_i, n_i), i \in I\}$  can be given the same state space with dimension  $n = \max_{i \in I} n_i$  by augmenting the deficient  $A_i$ 's with zeros, i.e. by introducing dummy components. Since each augmented  $M_i$  realizes the original  $M_i$ , we can apply Theorem 6.1 and conclude that a parallel composition of any set of LSMs over  $K$  can be realized non-trivially by an LCA provided  $K$  contains the necessary principle root of unity. (By non-trivial, we mean realizations other than an LCA with one large cell and the trivial



structure group).

Since the field  $\mathbb{C}$  contains a principle  $m^{\text{th}}$  root of unity for any positive integers  $m$ , we have the following:

Corollary 6.1

The parallel composition of any finite set of finite dimensional LSMs over  $\mathbb{C}$  can be non-trivially realized by an LCA.

If the LSMs all have transition functions given by matrices containing only real numbers, it can be shown that the realizing LCA has weighting functions that are real and even. This follows from the fact that real valued functions are Fourier transforms of even functions (see, for example, Cooley, Lewis, and Welch [1969]).

We have explicitly stated these facts using the notion of realization because certain properties of the assignment  $\text{Emb}_{\rho(I)} F_S^{-1}$  are suggestive of interesting applications. In particular, for a function  $f$ , it is true that every value of its transform  $\hat{f}$  depends on every value of  $f$ ; and conversely, every value of the inverse transform of  $\hat{f}$  depends on every value of  $\hat{f}$ . This follows from the fact that the support of each group character  $\phi_s$  is all of  $S$ . It can therefore be said that the process represented by a single LSM in  $\prod_{i \in I} M_i$  is distributed by the assignment  $F_S^{-1} \text{Emb}_{\rho(I)}$  throughout the entire structure of the cellular automaton  $M$ . Of course, other linear operators have similar properties. All that is required is that the rows of the matrix representations contain few zeros. With the Fourier transform, however, a heterogeneous collection of LSMs can be realized by a uniform structure. We will call the realization given by Theorem 6.1 a distributed realization of  $\prod_{i \in I} M_i$ .

It would not have been misleading to call these realizations holographic realizations. In fact, the connection to holography can be made fairly precise. Suppose a holographic image is to be made of an illuminated object. One can think of each point on the object's surface as an electro-magnetic oscillator having an amplitude depending on its brightness and a frequency depending on its color. The relative phases of the oscillations depend on the relative "depths" of the corresponding points on the object's surface. Each oscillation can be modeled by a simple LSM over  $\mathbb{C}$  with a one-dimensional state space. *Then the object can be thought of as the parallel composition of LSMs.*

The holographic image corresponds to the distributed realization of this collection of oscillators in the following way. A holographic image is recorded on a photographic plate over a short interval of time. Think of taking a time-exposure of the distributed realization as it is operating. The result will be a superposition of all the configurations assumed by the cellular automaton in that time interval. In general, this superposition will not contain the desired information about the object. However, if all of the independent oscillators making up the object have approximately the same frequency, the distributed realization has a small enough neighborhood so that the time-exposure captures essential information pertaining to the object's three-dimensional appearance. This condition corresponds to the requirement of a coherent light source (e.g. a laser) for making holographic images.

We return now to the major theme of this chapter and use Theorem 6.1 to prove the existence of certain kinds of homomorphisms of LCAs. For more information on the holographic process, we refer the reader

to Stroke [1969].

### 6.5 Homomorphisms of LCAs via Spatial Frequency Filtering

In Sec. 6.2 we said that since all the components in the decomposition of an LCA  $M$  operate independently, when any subset of them are "filtered out", the remaining components form a homomorphic image of the complete set. This smaller set of components can then be realized by another LCA  $M'$  such that a submachine of  $M'$  is a homomorphic image of  $M$ .

The simplest example of this procedure is shown in Fig. 6.2.

The LCA  $M = \langle Z_2, \mathbb{R}, H \rangle$  is decomposable into the parallel composition  $\hat{M}$  by the DFT w.r.t.  $Z_2$  whose matrix appears in Ex. 5.1. The component  $M_1$  is eliminated from  $\hat{M}$  to produce what we've called  $\prod_{s \in W} M_s$  ( $W \subseteq S$  is the set of components retained). The map  $\text{Res}_W$  shown in the figure is an automaton-homomorphism. The LCA  $M'$  is a distributed realization of  $\prod_{s \in W} M_s$ . Since we know from Theorem 6.1 that the assignment  $F_{Z_2}^{-1} \text{Emb}_W$  is an isomorphism from  $\prod_{s \in W} M_s$  to the submachine  $M'_W$  of  $M'$ , the composition of  $F_{Z_2}^{-1} \text{Emb}_W$ ,  $\text{Res}_W$ , and  $F_{Z_2}$  (which turns out to be the map  $\xi_W$  given by Def. 6.6), is an automaton-homomorphism. Since  $S-W = \{1\}$ , we say that the map  $\xi_W$  filters the group character  $\phi_{-1} = \phi_1 = (1, -1)$  from the LCA  $M$  (see Sec. 5.2).

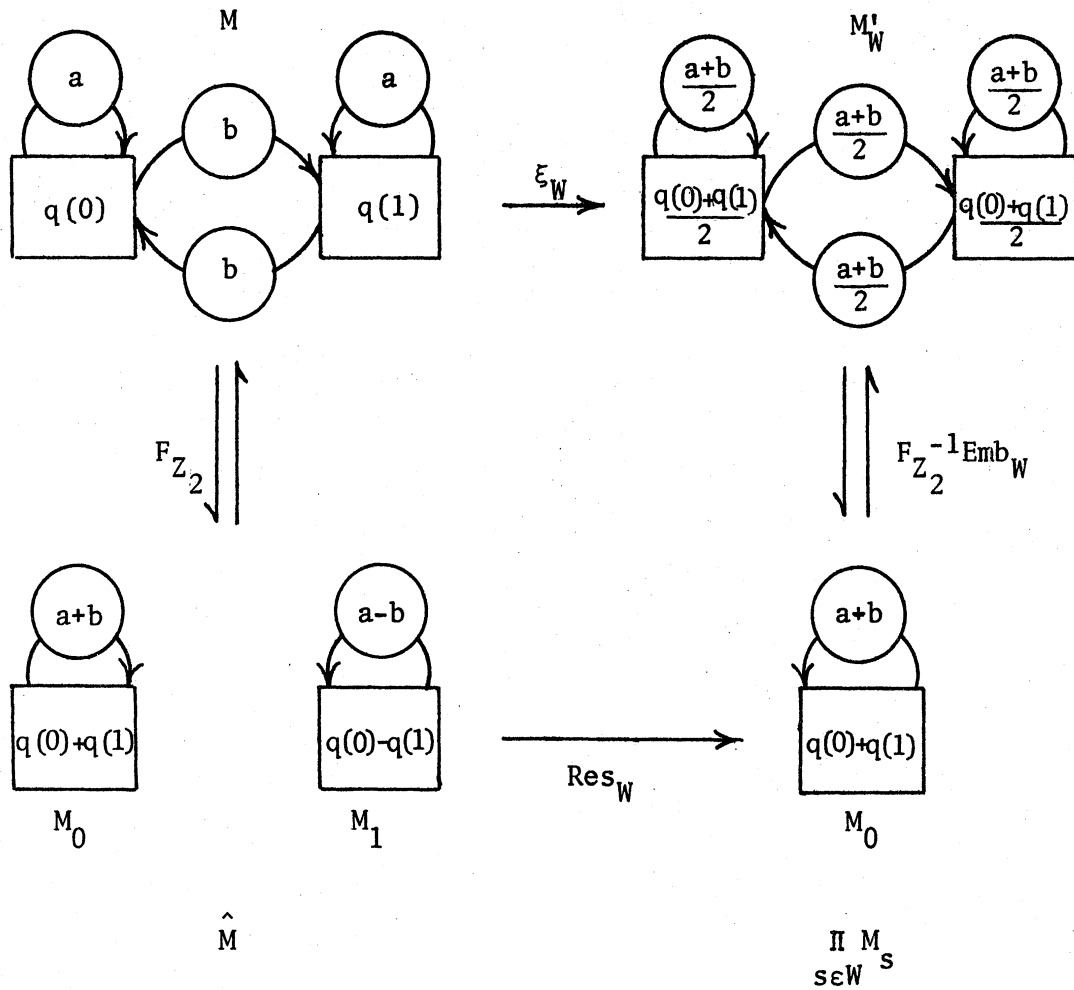


Figure 6.2:  $M'_W$  is a homomorphic image produced by filtering the character  $\phi_1 = (1, -1)$  from the LCA  $M = \langle \mathbb{Z}_2, \mathbb{R}, H \rangle$ .

Formally, we have the following theorem:

Theorem 6.2

Let  $M = \langle S, K^n, H \rangle$  be a decomposable LCA and  $W$  any subset of  $S$ . Then the submachine  $M'_W$  of the LCA  $M' = \langle S, K^n, H' \rangle$  with  $H' = \xi_W H$  is a homomorphic image of  $M$ .

Proof:

The set  $\{M_s = \langle K^n, \hat{H}(s) \rangle \mid s \in W\}$  satisfies the hypotheses of Theorem 6.1 for  $\rho$  the inclusion map of  $W$  into  $S$ . Then the LCA  $M'$  is the distributed realization of  $\prod_{s \in W} M_s$  given by Theorem 6.1 since

$$\begin{aligned} H' &= \xi_W H = [F_S^{-1} \text{Proj}_W F_S] H \\ &= F_S^{-1} \text{Proj}_W \hat{H} \\ &= F_S^{-1} \text{Emb}_W(\text{Res}_W \hat{H}) . \end{aligned}$$

In other words, the function called  $A$  in Theorem 6.1 is  $\text{Res}_W \hat{H}$ . It's easy to check that the map  $\text{Res}_W$  is a homomorphism from  $\hat{M} = \prod_{s \in S} M_s$  to  $\prod_{s \in W} M_s$ . Then composing the maps shown in the diagram, we have that

$$F_S^{-1} \text{Emb}_W \text{Res}_W F_S = F_S^{-1} \text{Proj}_W F_S = \xi_W$$

is the required homomorphism.

$$\begin{array}{ccc} M & \xrightarrow{\xi_W} & M'_W \\ \begin{array}{c} \uparrow \\ F_S \\ \downarrow \end{array} & & \begin{array}{c} \downarrow \\ F_S^{-1} \text{Emb}_W \\ \uparrow \end{array} \\ \hat{M} = \prod_{s \in S} M_s & \xrightarrow{\text{Res}_W} & \prod_{s \in W} M_s \end{array}$$

Q.E.D.

We have maintained the distinction between the LCA  $M'$  and the submachine  $M'_W$  whose state set consists of all "smooth" configurations of  $M'$ . However, the difference between  $M'$  and  $M'_W$  is very slight. It is easiest to see this by looking at  $\hat{M}'$  and  $\hat{M}'_W$ , the parallel decompositions of  $M'$  and  $M'_W$  respectively.  $\hat{M}'$  and  $\hat{M}'_W$  consist of exactly the same components  $M_s = \langle K^n, \hat{H}'(s) \rangle$ ,  $s \in S$ , where  $\hat{H}' = F_S \xi_W H = \text{Proj}_W \hat{H}$ , but

any state  $f$  of the submachine  $\hat{M}'_W$  must be such that  $f(s) = 0$  for  $s \in S-W$ .

Now suppose  $\hat{M}'$  is started in an arbitrary initial state  $q \in F(S, K^n)$ . Then the next state, say  $q'$ , is given by

$$q'(s) = (\text{Proj}_W \hat{H})(s) \cdot q(s), \quad s \in S.$$

Thus, for  $s \in S-W$ ,  $q'(s) = 0 \cdot q(s) = 0$  so that  $q'$  is a state of the submachine  $\hat{M}'_W$ . The only difference, therefore, between  $\hat{M}'$  and  $\hat{M}'_W$  is that the initial state of  $\hat{M}'$  is sometimes not a state of  $\hat{M}'_W$ . Since these machines are isomorphic to  $M'$  and  $M'_W$  respectively, we have that the submachine  $M'_W$  is always entered in the first transition of the LCA  $M'$ . Technically, however, we need to distinguish between  $M'$  and  $M'_W$  since our definition of an automaton-homomorphism (Sec. 2.5) requires the map  $\xi_W$  to be onto.

Theorem 6.2 is an automata theoretic restatement of the classical principle which, if the structure group  $S$  were allowed to be  $\mathbb{R}$ , could be succinctly stated as the Principle of Rayleigh. We quote from Duff and Naylor ([1966] p. 87):

*"Principle of Rayleigh: A continuous system can be approximated by a finite system, so that the eigenvalues and eigenvectors of the finite system will approximate to a finite number of the eigenvalues and eigenfunctions of the continuous system."*

Here we see that projection onto a subspace spanned by eigenvectors is an example of the concept of automaton homomorphism.

The symbol  $W$  was chosen to suggest the term "window" used in the signal processing literature (e.g. Gold and Rader [1969]). The operator  $\xi_W = F_S^{-1} \text{Proj}_W F_S$  represents an ideal band-pass filter of signals which here we interpret as spatial signals, i.e. they are cellular auto-

maton configurations. The frequency band is the set  $W$ , and we say that windowing in the frequency domain is filtering in the spatial domain.

The filtering operation can be represented more concretely through the use of characteristic functions. Let  $V$  be a vector space with zero vector denoted by  $0$  and let  $\hat{\chi}_W \in F(S, V)$  be given by

$$\hat{\chi}_W(s) = \begin{cases} 1 & \text{for } s \in W \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } s \in S.$$

The function  $\hat{\chi}_W$  is the characteristic function of  $W$ . Again, in using this symbol, we leave it to the context to indicate what space  $V$  is.

It is clear that  $\text{Proj}_W f = \hat{\chi}_W \cdot f$  for all  $f \in F(S, V)$ . Letting  $\chi_W = F_S^{-1} \hat{\chi}_W$ , we have for all  $f \in F(S, V)$  that

$$(F_S^{-1} \text{Proj}_W F_S)(f) = (F_S^{-1} \text{Proj}_W)(\hat{f}) = F_S^{-1}(\hat{\chi}_W \cdot f) = \chi_W *_S f$$

by the convolution theorem. Thus, letting  $V = \text{Mat}_K(n, n)$ , the function  $H'$  of  $M'$  is given by

$$H' = \chi_W *_S H \quad (6.4-1)$$

and (with  $V = K^n$ ) the homomorphism  $\xi_W$  is given by

$$\xi_W q = \chi_W *_S q \text{ for all } q \in F(S, K^n). \quad (6.4-2)$$

For the temporal interpretation of the group  $S$ , ideal filters of the kind we have described are not physically realizable since  $\chi_W$  given in (6.4-2) is a non-causal operator (Gold and Rader [1969]) for  $W \not\subseteq S$ . This is true since the support of  $\chi_W$  is not "one-sided" for any  $W \not\subseteq S$ . For our spatial interpretation, however, this fact creates no difficulty.

Fig. 6.3 a) shows the functions  $\chi_W$  and  $\hat{\chi}_W$  associated with the homomorphism  $\xi_W$  of Fig. 6.2. The "window"  $W$  is the set  $\{0\}$ . Fig. 6.3 b)

shows the action of  $\xi_W$  on a particular function of  $Z_2$ . Each value of  $\xi_W q$  is the average of the values of  $q$ . This is always the case for  $W = \{e\}$  where  $e$  is the identity element of the underlying group.

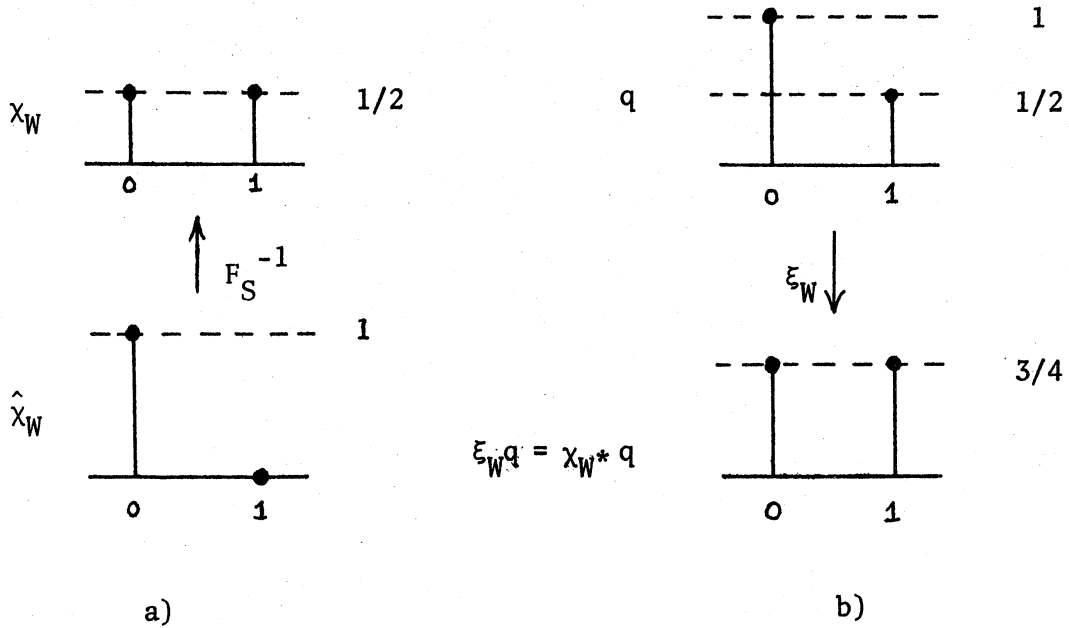
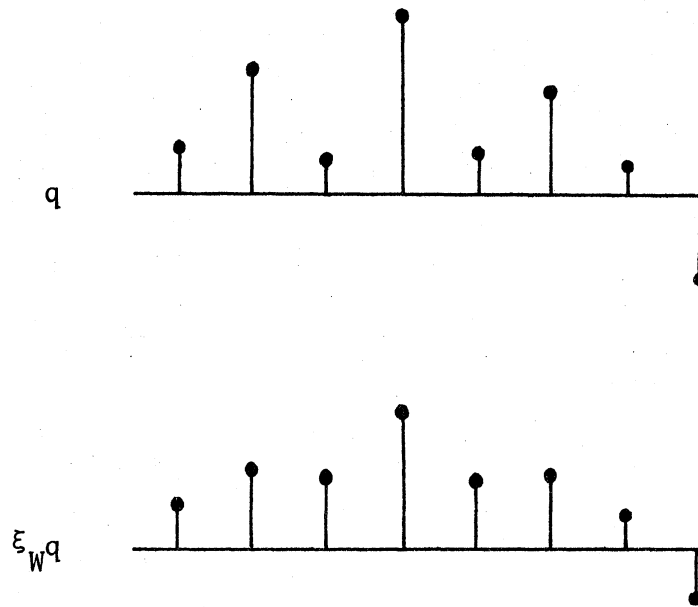


Figure 6.3: Filtering on  $Z_2$  with window  $W = \{0\}$ :

- a) the characteristic function  $\hat{x}_W$  and its inverse transform  $x_W$ ,
- b) the action of  $\xi_W$  on a  $q \in F(Z_2, \mathbb{R})$ .

Section 6.2 intuitively described an automaton homomorphism as a map that produces a low resolution version of a given automaton. When the set  $W$  underlying the homomorphism  $\xi_W = F_S^{-1} \text{Proj}_W^F$  consists of the lower multiples of the generators of  $S$  and their inverses, the map  $\xi_W$  might be called a "low-pass" filter and can be thought of as a smoothing operator. Figure 6.4 illustrates the action of such a map on a particular function  $q \in F(Z_8, \mathbb{C})$  for  $W = \{6, 7, 0, 1, 2\}$ .





$$W = \{6, 7, 0, 1, 2\}$$

Figure 6.4: Low-pass filtering of  $q \in F(\mathbb{Z}_8, \mathbb{C})$  .

We refer the reader to picture processing literature (e.g. Andrews [1970]) for evidence that the homomorphic images of LCAs produced by low-pass spatial-frequency filtering are indeed low-resolution or blurry images.

Finally, we would like to point out an interesting fact about the situation illustrated in Fig. 6.2. If we let  $\delta$  denote the transition of  $M$  and  $\delta_0$  denote the transition function of  $M_0$ , then the crucial fact underlying the existence of the homomorphism  $\xi_W$  is represented by the following commutative diagram:

$$\begin{array}{ccc}
 (q(0), q(1)) & \xrightarrow{\delta} & ([aq(0)+bq(1)], [bq(0)+aq(1)]) \\
 \downarrow \text{Res}_W^F Z_2 & & \downarrow \text{Res}_W^F Z_2 \\
 (q(0)+q(1)) & \xrightarrow{\delta_0} & (aq(0)+bq(1)+bq(0)+aq(1)) \\
 & & \parallel \\
 & & (a+b)(q(0)+q(1))
 \end{array}$$

The equality on the right follows from the distributive property required of rings. In fact, the commutativity of this diagram is equivalent to the distributive property of the underlying commutative ring.

This is always the case, in fact, for  $W = \{e\}$ , where  $e$  is the identity of  $S$ , since the character  $\phi_e$  for any group  $S$  is identically equal to 1. We see, though, that any  $W \subseteq S$  will produce a commutative diagram. This suggests a generalization of distributivity that perhaps has interesting applications. We have not yet explored the mathematical literature to see if this observation has been incorporated into some algebraic framework.

## 6.6 Complexity of Homomorphic images

The homomorphism  $\xi_W$ ,  $W \subseteq S$ , defined in the previous section is a map between LCAs that have the same structure group so that  $M$  and  $M'$  have the same number of cells. Thus, despite the lower resolution of the LCA  $M'$ , it is not clear how  $M'$  is less complex than  $M$ , something we would desire of a homomorphic image for simulation purposes. In

fact,  $M'$  might be considered more complex than  $M$  since in general the neighborhoods of cells in  $M'$  are larger than those of cells in  $M$ . This is true since filtering the function  $H$  always tends to increase its support.

Intuitively,  $M'$  is less complex than  $M$  in the sense that the cells of  $M'$  process less information than cells of  $M$ : they are sensitive to a narrower spatial frequency band. We will not pursue the precise information theoretic characterization of this observation but will ask, instead, how this kind of complexity reduction can be exploited for simulation purposes. Stated another way, how can this "information theoretic" complexity measure be interpreted in terms of more standard complexity measures associated with the number of components and density of interconnections? We will examine one answer to this question in the next chapter by interpreting aspects of sampling theory in the language of LCAs.

## Chapter 7

### Linear Cellular automata and Sampling Theory

#### 7.1 Introduction

A natural technique for modeling a "real" homogeneous medium by a cellular automaton is to view each cell of the automaton as a representation of a small area of the system to be modeled. For example, a cellular automaton specified to model a section of fairly homogeneous neural tissue might be formulated so that each "cell" of the automaton is intended to model a small block of neural tissue containing thousands of neurons. In other words, the cellular automaton is an aggregated model of neural tissue. Each state variable of the model (i.e. each cell of the cellular automaton) represents a set of components of the real system (i.e. a set of neurons). Examples of this kind of aggregation are Mortimer's model of the mammalian cerebellar cortex [1970] and Foy's model of cardiac muscle tissue [1974]. Intuitively, it is felt that the behavior of the aggregated model ought to be essentially like that of the more complex natural system except for a lack of fine spatial detail.

While these kinds of models tend to be much more complicated than the linear cellular automata considered here, it is nevertheless important to understand some of the difficulties that can arise when LCAs are aggregated. The reason for this is that even in the comparatively simple linear case, aggregated models tend to produce misleading behavior *for reasons that seem very counter-intuitive*. It is therefore not sufficient to appeal to intuitive explanations for a justification of the aggregation process.

To examine these problems using the framework developed so far, we will take the point of view that the underlying real system is some LCA with structure group  $S$ . We will then be interested in describing homomorphic images of this LCA that are also LCAs but having structure groups smaller than  $S$ . Using Zeigler's terminology [1974], we will postulate an LCA  $M$  as a base model and examine lumped models which are LCAs with fewer cells than  $M$  and which homomorphically preserve aspects of  $M$ 's behavior.

Theorem 6.2, proved in the previous chapter, establishes a homomorphic relation between an LCA  $M$  and an LCA  $M'$ , but  $M$  and  $M'$  have the same structure group  $S$  and hence have the same number of cells. It's just as easy, however, to obtain a similar result in which  $M'$  has fewer cells than  $M$ . In the next section we develop this idea more carefully and then motivate the application of sampling theory to questions arising from using LCAs to model natural systems.

## 7.2 Homomorphisms between LCAs and having Different Structure Groups

Referring back to Fig. 6.2 in the previous chapter, the LCA  $M'$  in the upper right of the figure is a particular distributed realization of  $\prod_{s \in W} M_s$ , namely the one with structure group  $S$ . Any distributed realization, however, will give rise to a homomorphism analogous to  $\xi_W$  of Fig. 6.2. In particular,  $M'$  can have structure group  $S'$  where  $|W| \leq S' < S$  since  $W$  can then be injected into  $S'$  instead of  $S$ . Using this observation we can prove the following theorem:

Theorem 7.1

Let  $M = \langle S, K^n, H \rangle$  be a decomposable LCA,  $W$  any subset of  $S$ , and let  $M' = \langle S', K^n, H' \rangle$  be a decomposable LCA such that  $|W| \leq |S'|$ . Then for any injection  $\rho$  of  $W$  into the underlying set of  $S'$ , if

$H' = F_{S'}^{-1} \text{Emb}_{\rho(W)} \text{Res}_W F_S H$ , then the submachine  $M'_{\rho(W)}$  of  $M'$  is a homomorphic image of  $M$ .

Proof:

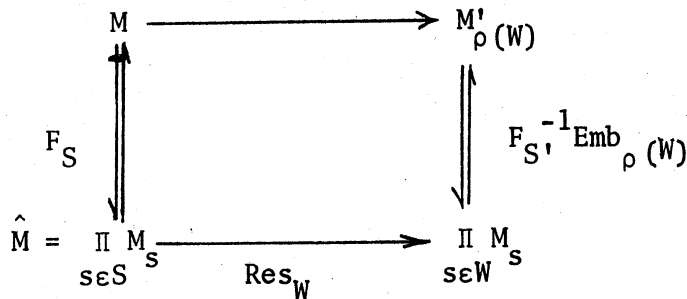
This is proved in basically the same way that we proved Theorem 6.1.

First,  $M'_{\rho(W)}$  is the submachine of  $M'$  with state set  $F_{S'}^{-1} \text{Proj}_{\rho(W)}(F(S', K^n))$ , i.e. the group  $S'$  is used in Def. 6.5 instead of  $S$ . Also,  $\text{Res}_W : F(S, K^n) \rightarrow F(W, K^n)$  since, by our notation,

$\text{Res}_W = \text{Res}_{i(W)}$  where  $i$  is the inclusion map; but

$\text{Emb}_{\rho(W)} : F(W, K^n) \rightarrow F(S', K^n)$ .

Then consider the following diagram:



All we need to show is that  $M'$  with  $H' = F_{S'}^{-1} \text{Emb}_{\rho(W)} \text{Res}_W F_S H$  is a distributed realization of  $\coprod_{s \in W} M_s$ . Applying Theorem 6.1 with  $A = \text{Res}_W F_S H$  and  $S'$  taking the place of  $S$  (and  $M'$  taking the place of  $M$ !), it can be seen that  $M'$  is indeed a distributed realization of

$\coprod_{s \in W} M_s$ . Thus, composing the maps in the diagram, the required homomorphism is

$$F_{S'}^{-1} \text{Emb}_{\rho(W)} \text{Res}_W F_S. \quad (7.2-1)$$

Q.E.D.

For  $S = S'$  and  $\rho$  the inclusion map of  $W$  into  $S$ , this homomorphism reduces to  $\xi_W = F_S^{-1} \text{Proj}_W F_S$  given by Theorem 6.2.\*

There are several difficulties with the homomorphism  $F_{S'}^{-1} \text{Emb}_{\rho(W)} \text{Res}_W F_S$ . First, it is obviously complicated. Since it is a linear operator, it can be represented by a matrix, but the matrix does not in general help clarify matters. In Sec. 6.5 we showed that the map  $\xi_W$  has a concrete representation in terms of the convolution operation (see (6.4-1) and (6.4-2)). Here, however, the matrix representation is not so nicely structured.

For example, let  $S = Z_3$ ,  $S' = Z_2$ ,  $W = \{0,1\}$ , and let  $\rho: W \rightarrow Z_2$  be the identity map. Then if  $\mathbb{C}$  is the underlying field (and  $n = 1$ ), the operator (7.2-1) maps  $\mathbb{C}^3$  into  $\mathbb{C}^2$  and has the matrix representation:

$$\begin{bmatrix} 1 & \frac{1+\omega}{2} & \frac{1+\bar{\omega}}{2} \\ 0 & \frac{1-\omega}{2} & \frac{1-\bar{\omega}}{2} \end{bmatrix} \quad (7.2-2)$$

where  $\omega$  is a principle 3<sup>rd</sup> root of unity in  $\mathbb{C}$ , i.e.  $\omega = e^{2\pi i/3}$ . In

---

\* It's interesting to note here that if  $S = S'$ ,  $W = S$ , and  $\rho$  is a permutation of the elements of  $S$ , then the map  $\xi_{\rho(S)} = F_S^{-1} \text{Proj}_{\rho(S)} F_S$  is an isomorphism of the LCAs  $M$  and  $M'$  each having structure group  $S$ . In fact, if  $n=1$ , all such isomorphisms are associated with permutations so that there are exactly  $N!$  of them where  $N = |S|$ . See Nicholson [1971]. We will not pursue the modeling theoretic interpretation of these maps.

the general case, the most intuitively satisfying way to think about these kinds of operators may be to go back to expression (7.2-1) i.e. to view them as appropriate compositions of restrictions, embeddings, and DFTs.

Another difficulty with the homomorphism given by Theorem 7.1 is that in order for it to exist, both of the LCAs  $M$  and  $M'$  must be decomposable over the same field  $K$ . This means that the group  $S'$  must be chosen so that the DFT exists over  $K$  for  $S$  and  $S'$ . When  $K = \mathbb{C}$  there is no problem, but for finite fields care must be taken in choosing the group  $S'$ .

Neither of these difficulties exist for the homomorphism  $\xi_W$  given by Theorem 6.2. It can be represented by a convolution operation, and since  $S = S'$ ,  $M'$  is decomposable whenever  $M$  is. As indicated in Sec. 7.1, however, we would like to consider cases in which  $M'$  has fewer cells than  $M$ . It turns out that when  $S'$  is a certain kind of subgroup of  $S$ , the situation is different, and all of these objections are met.\* In order to establish these results, it is necessary to develop aspects of sampling theory which we do in the next section. In Sec. 7.4 we will return to the discussion of automaton-homomorphisms.

### 7.3 Generalized Sampling Theory

As we have indicated, several results about LCAs are essentially cellular automatic theoretic interpretations of well known facts about filtering and sampling signals which, in our case, are spatial

---

\*It seems likely that if  $S'$  is any subgroup of  $S$ , the same results can be proved. However, within the framework developed so far, it's easiest to provide  $S'$  with an explicit representation. Whether or not all subgroups can be so represented is a purely group theoretic question that we do not pursue here.



signals (i.e. configurations of a cellular automaton). However, the group theoretic framework of our discussion requires us to develop the sampling results in similar group theoretic terms. This sort of presentation is not new but is 1) more abstract than the engineering literature (e.g. Cooley, Lewis, and Welch [1969]), and 2) less abstract than the mathematical literature (e.g. Hewitt and Ross [1963]).

Our development owes the most to algebraic developments of the Fast Fourier Transform algorithm, in particular to Cairns [1971] and Cooley, Lewis, and Welch [1969], but provides the generalization from complex numbers to integral domains containing suitable roots of unity.

Let  $G$  be an arbitrary finite abelian group with  $|G| = N$  and canonical decomposition as in (5.2-1), i.e.

$$G \approx \mathbb{Z}_{n_1} \oplus \dots \oplus \mathbb{Z}_{n_d} \quad \text{where} \quad \prod_{j=1}^d n_j = N \quad \text{and} \quad n_i | n_{i+1}. \quad (7.3-1)$$

Let  $D$  be the finite abelian group,  $|D| = M$ , with canonical decomposition

$$D \approx \mathbb{Z}_{m_1} \oplus \dots \oplus \mathbb{Z}_{m_d} \quad \text{where} \quad \prod_{j=1}^d m_j = M, \quad m_i | m_{i+1} \quad \text{and} \quad m_i | n_i. \quad (7.3-2)$$

Suppose  $n_i = m_i k_i$  for  $i=1, \dots, d$ . Define two maps  $\rho: D \rightarrow G$  and  $\eta: G \rightarrow D$  as follows:

$$\rho((a_1, \dots, a_d)) = (k_1 a_1, \dots, k_d a_d) \quad \text{for all } (a_1, \dots, a_d) \in D \quad (7.3-3)$$

where  $k_i a_i = a_i + \dots + a_i$  ( $k_i$  times) and

$$\eta((s_1, \dots, s_d)) = (s_1 \bmod m_1, \dots, s_d \bmod m_d) \in D \quad \text{for all } (s_1, \dots, s_r) \in G. \quad (7.3-4)$$

The following facts can be proven about  $\rho$  and  $\eta$ :

1)  $\rho$  is an injection of  $D$  into  $G$  which is a group-homomorphism.

Hence  $\rho(D)$  is a subgroup of  $G$  and  $M|N$ .

2)  $\eta$  is a group-homomorphism of  $G$  onto  $D$ .

3) The kernel  $K$  of  $\eta$  is (isomorphic to)  $\mathbb{Z}_{k_1} \oplus \dots \oplus \mathbb{Z}_{k_d}$ .

Let  $k$  be such that  $N = kM$ , and let  $L$  be a module. In addition to the operators  $\text{Res}_{\rho(D)}: F(G, L) \rightarrow F(D, L)$ ,  $\text{Proj}_{\rho(D)}: F(G, L) \rightarrow F(G, L)$ ,

and  $\text{Emb}_{\rho(D)} : F(D,L) \rightarrow F(G,L)$  given by Defs. 6.1, 6.2, and 6.3 respectively, we define two more operators:

Definition 7.1

The periodic extension w.r.t.  $\eta$  is the operator  $\text{PE}_{\eta} : F(D,L) \rightarrow F(G,L)$  given by  $\text{PE}_{\eta} f = f\eta$  for all  $f \in F(D,L)$ .

Definition 7.2

The aggregation w.r.t.  $\eta$  is the operator  $\text{Agg}_{\eta} : F(G,L) \rightarrow F(D,L)$  given by  $(\text{Agg}_{\eta} f)(a) = \frac{1}{k} \sum_{\eta(r)=a} f(r)$  for all  $f \in F(G,L)$  and  $a \in D$ .

All five of these operators are illustrated in Fig. 7.1 for  $G = Z_8$ ,  $L = \mathbb{R}$ , and  $\eta$  and  $\rho$  as shown. Note that the operator  $\text{Agg}_{\eta}$  averages over cosets of  $K$  and that  $\text{Agg}_{\eta} \text{PE}_{\eta}$  is the identity on  $F(D,L)$ . We will let the context indicate what module  $L$  denotes in particular instances : it will be an integral domain  $R$ , the module  $R^n$ , or the module of matrices  $\text{Mat}_R(n,n)$ . Note that we need the structure of a module so that Def. 7.2 makes sense. When no confusion is likely to occur, we will drop the subscripts from the names of these operators.

The operator  $\text{Emb}_{\rho(D)}$  is the natural generalization of the operator given by Cooley, Lewis, and Welch [1969] for cyclic groups which they called "Stretch<sub>k</sub>". Similarly, their operator "Sample<sub>k</sub>" corresponds to our  $\text{Res}_{\rho(D)}$ . However, they identify the functions  $\text{Res}_{\rho(D)} f$  and  $\text{PE}_{\rho(D)} \text{Res}_{\rho(D)} f$  by saying that  $\text{Sample}_k f$  is regarded as extended to all the integers in a periodic way (hence our name "periodic extension"). Intuitively,  $\text{Res}_{\rho(D)}$  takes equally spaced samples of a function where  $k$ , the size of the kernel of  $\eta$ , is the sampling rate. For the example shown in Fig. 7.1,  $k = 2$  and  $\text{Res}_{\rho(D)} f$  consists of every second value of  $f$ .

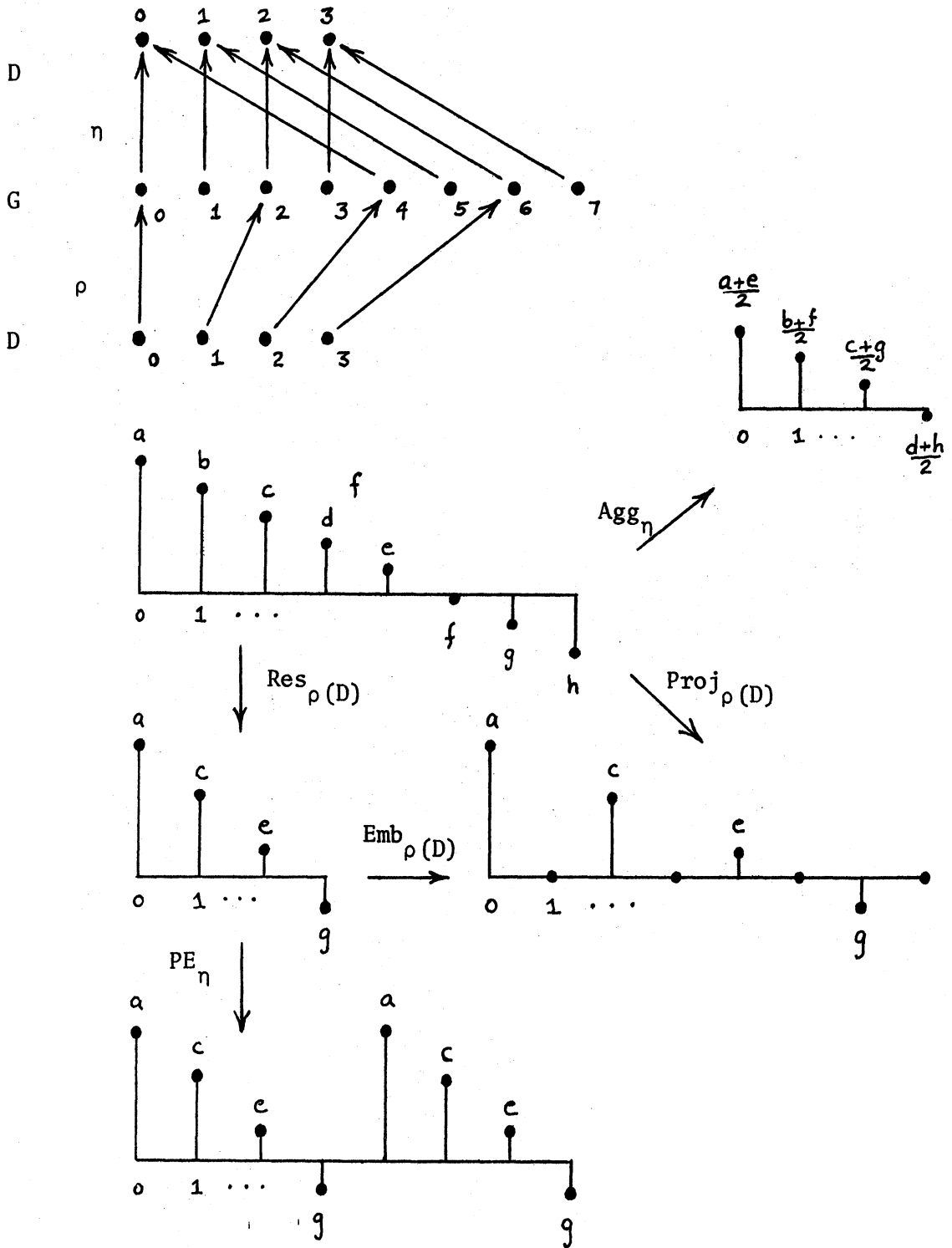


Figure 7.1: Action of the operators  $\text{Res}_\rho(D)$ ,  $\text{Proj}_\rho(D)$ ,  $\text{Emb}_\rho(D)$ ,  $\text{PE}_\eta$ , and  $\text{Agg}_\eta$  on an  $\text{feF}(Z_g, \mathbb{R})$ .

The results of this section express relationships between the above five operators and the Fourier transforms  $F_G$  and  $F_D$ . First it is necessary to show that the conditions given in Theorem 5.1 which insure the existence of  $F_G$  also insure the existence of  $F_D$ . According to the canonical decompositions of  $G$  and  $D$  given by (7.3-1) and (7.3-2), the exponents of  $G$  and  $D$  are  $n_d$  and  $m_d$  respectively, and we know that  $n_d = k_d m_d$ . Consequently, if  $R$  is an integral domain containing a principle  $n_d^{\text{th}}$  root of unity  $\omega$ , then it contains a principle  $m_d^{\text{th}}$  root of unity, namely  $\omega' = \omega^{n_d/m_d} = \omega^{k_d}$ . Similarly, since  $N = kM$ , we know that if  $R$  contains  $N^{-1}$  then it contains  $M^{-1} = N^{-1} + \dots + N^{-1}$  ( $k$  times).

Therefore, let  $R$  be an integral domain containing both a principle  $n_d^{\text{th}}$  root of unity  $\omega$  and  $N^{-1}$ . Following Def. 5.1 we can define the set of characters of  $D$  in  $R$ :

$$\begin{aligned} \Phi(D,R) &= \{\theta_a \mid a \in D\} \subseteq F(D,R) \quad \text{where} \\ \theta_{(a_1, \dots, a_d)}((b_1, \dots, b_d)) &= \prod_{j=1}^d \omega^{a_j b_j n_d / m_j} \end{aligned} \quad (7.3-5)$$

for all  $(a_1, \dots, a_d), (b_1, \dots, b_d) \in D$ .

The facts 1-7 of Sec. 5.2 are true of  $\Phi(D,R)$  so that we have the following lemma:

#### Lemma 7.1

For  $G$  and  $D$  with canonical decompositions given by (7.3-1) and (7.3-2) respectively, if  $R$  is an integral domain containing  $N^{-1}$  and a principle  $n_d^{\text{th}}$  root of unity, then both  $F_G$  and  $F_D$  exist for functions with values in  $R$ .

We will always denote the characters of  $G$  by  $\phi_s$ ,  $s \in G$ , and those of  $D$  by  $\theta_a$ ,  $a \in D$ . The next result gives a pertinent relation between  $\phi(G, R)$  and  $\phi(D, R)$ .

Lemma 7.2

For all  $a \in D$ ,  $\text{PE}_\eta \theta_a = \phi_\rho(a)$ .

Proof:

We need to show that  $\theta_a \eta = \phi_\rho(a)$  for  $a \in D$ . Directly from (5.2-2) we have for  $(a_1, \dots, a_d) \in D$  and  $(r_1, \dots, r_d) \in G$  that

$$\begin{aligned} \phi_\rho(a_1, \dots, a_d)(r_1, \dots, r_d) &= \phi(k_1 a_1, \dots, k_d a_d)(r_1, \dots, r_d) & (7.3-6) \\ &= \prod_{j=1}^d \omega_j^{k_j a_j r_j^{n_d/n_j}} \\ &= \prod_{j=1}^d \omega_j^{k_j a_j r_j^{n_d/m_j k_j}} \quad (\text{since } m_j k_j = n_j) \\ &= \prod_{j=1}^d \omega_j^{a_j r_j^{n_d/m_j}} . \end{aligned}$$

For each  $j$ , we have that  $r_j = r_j \bmod m_j + \ell_j m_j$  where  $0 \leq \ell_j < k_j$ .

Thus (7.3-6) equals the following:

$$\begin{aligned} &\prod_{j=1}^d \omega_j^{a_j (r_j \bmod m_j + \ell_j m_j)^{n_d/m_j}} \\ &= \prod_{j=1}^d \left[ \omega_j^{a_j (r_j \bmod m_j)^{n_d/m_j}} \omega_j^{a_j \ell_j m_j^{n_d/m_j}} \right] . \end{aligned}$$

However, since  $\omega$  is a principle  $n_d^{\text{th}}$  root of unity, we have that

$$\omega_j^{a_j \ell_j m_j^{n_d/m_j}} = \omega_j^{a_j \ell_j n_d} = 1 .$$

Thus (7.3-6) equals

$$\prod_{j=1}^d \omega^{a_j (r_j \bmod m_j) n_d / m_j}$$

$$= \theta_{(a_1, \dots, a_d)}(r_1 \bmod m_1, \dots, r_d \bmod m_d) \quad \text{by (7.3-5)}$$

$$= [\theta_{(a_1, \dots, a_d)}]^\eta(r_1, \dots, r_d) \quad .$$

Q.E.D.

Thus, we know that for each  $s$  in the subgroup  $\rho(D)$ , the character  $\phi_s$  of  $G$  is the periodic extension of a character of  $D$ . This fact can be applied directly to the formulas for  $F_G$  and  $F_D$  (see (5.2-3)) to obtain a relationship between the DFT w.r.t.  $D$  of a function  $f \in F(D, R)$  and the DFT w.r.t.  $G$  of the function  $\text{Emb}_{\rho(D)} f \in F(G, R)$ . A similar relationship holds for the inverse DFTs.

### Theorem 7.2

$$1) \quad \text{PE}_{\eta D} F_D = F_G \text{Emb}_{\rho(D)}$$

$$2) \quad \text{PE}_{\eta D}^{-1} = k F_G^{-1} \text{Emb}_{\rho(D)}$$

Proof:

To prove part 1 let  $f \in F(D, R)$  and  $\alpha \in G$ . Applying Def. 7.1 to the formula for the DFT w.r.t.  $D$  (see 5.2-3) we have

$$(\text{PE}_{\eta D} f)(\alpha) = \sum_{b \in D} f(b) \theta_{\eta(\alpha)}(b) \quad (7.3-7)$$

From fact 4) stated in Sec. 5.2 we know that

$$\theta_{\eta(\alpha)}(b) = \theta_b(\eta(\alpha)) = (\theta_b \eta)(\alpha) \quad \text{for any } \alpha \in G \text{ and } b \in D.$$

Thus, from Def. 7.1 and Lemma 7.2, we have

$$\theta_{\eta(\alpha)}(b) = (\text{PE}_{\eta D} \theta_b)(\alpha) = \phi_{\rho(b)}(\alpha)$$

so that (7.3-7) becomes

$$(\text{PE}_{\eta} F_D f)(\alpha) = \sum_{b \in D} f(b) \phi_{\rho(b)}(\alpha) .$$

Since  $\rho$  is an injection, we can view this as a sum over  $\rho(D)$  by making the change of variable  $b \rightarrow \beta = \rho(b)$ . Thus  $b = \rho^{-1}(\beta)$  and we have

$$(\text{PE}_{\eta} F_D f)(\alpha) = \sum_{\beta \in \rho(D)} (f \rho^{-1})(\beta) \phi_{\beta}(\alpha),$$

and by fact 4) Sec. 5.2,

$$= \sum_{\beta \in \rho(D)} (f \rho^{-1})(\beta) \phi_{\alpha}(\beta) .$$

By the definition of the Emb operator (Def. 6.3) this is the same as

$$\sum_{\beta \in G} (\text{Emb}_{\rho(D)} f)(\beta) \phi_{\alpha}(\beta) .$$

But this is the DFT of  $\text{Emb}_{\rho(D)} f$  evaluated at  $\alpha$  so that

$$(\text{PE}_{\eta} F_D f)(\alpha) = (F_G \text{Emb}_{\rho(D)} f)(\alpha) .$$

This proves part 1. Part 2 is proved analogously but with the observations that  $\rho(-b) = \beta$  implies  $\rho(b) = -\beta$  since  $\rho$  is a group homomorphism .

Q.E.D.

### Corollary 7.1

$$k \text{Emb}_{\rho(D)} F_D = F_G \text{PE}_{\eta} .$$

Proof:

From Theorem 7.2 part 2 we have



out, even in the abstract case we are considering here, that this is the kind of smoothness that allows a rigorous connection to be made between a function and its sampled version.

Within our framework these results take the form of a relationship between the DFT w.r.t.  $G$  of a function and the DFT w.r.t.  $D$  of its sampled version. The operator  $\text{Res}_{\rho(D)}$  can be thought of as performing the sampling operation where the samples are taken at equally spaced elements of  $G$ . Thus, we will establish a relationship between the operators  $F_G$  and  $F_D \text{Res}_{\rho(D)}$ .

First, we need a relationship between the DFT of a function and the DFT of its projection, i.e. between the operators  $F_G$  and  $F_G \text{Proj}_{\rho(D)}$ . The appropriate relationship is a consequence of the following lemma:

Lemma 7.3

$$\text{Agg}_{\eta} \phi_s \equiv 0 \quad \text{for } s \in G - \rho(D).$$

Proof:

This is difficult to prove directly so we will take the following approach: first we will show that for any  $s \in G$ ,  $\text{Agg}_{\eta} \phi_s$  is a homomorphism from  $D$  into the multiplicative monoid of  $R$ . Since  $\phi(D, R)$ , the set of characters of  $D$  in  $R$ , contains all such homomorphisms that are not identically zero (fact 1 Sec. 5.2), we will then have that  $\text{Agg}_{\eta} \phi_s$  is either a character of  $D$  or identically zero. But all the characters of  $D$  are accounted for by Lemma 7.2. We will then conclude that  $\text{Agg}_{\eta} \phi_s \equiv 0$  for  $s \in G - \rho(D)$ .

To show for  $s \in G$  that  $\text{Agg}_{\eta} \phi_s$  is a homomorphism from  $D$  into  $R$  we need to show that

$$(\text{Agg}_{\eta} \phi_s)(a+b) = (\text{Agg}_{\eta} \phi_s)(a) \cdot (\text{Agg}_{\eta} \phi_s)(b) \quad \text{for all } a, b \in D. \quad (7.3-8)$$

Using Def. 7.2 of the Agg operator, the right hand side of (7.3-8)

is

$$\begin{aligned} & \left[ \frac{1}{k} \sum_{\eta(x)=a} \phi_s(x) \right] \cdot \left[ \frac{1}{k} \sum_{\eta(y)=b} \phi_s(y) \right] \\ &= \frac{1}{k^2} \sum_{\eta(x)=a} \sum_{\eta(y)=b} \phi_s(x) \phi_s(y) \\ &= \frac{1}{k^2} \sum_{\eta(x)=a} \sum_{\eta(y)=b} \phi_s(x+y) \quad . \quad (7.3-9) \end{aligned}$$

If we make the change of variable  $x+y \rightarrow r$ , then  $y \rightarrow r-x$  and the inside summation is taken over all  $r$  such that  $\eta(r-x)=b$  or, since  $\eta$  is a group-homomorphism, all  $r$  such that  $\eta(r) = \eta(x)+b$ . Thus (7.3-9)

becomes

$$\frac{1}{k^2} \sum_{\eta(x)=a} \left[ \sum_{\eta(r)=a+b} \phi_s(r) \right] \quad .$$

but since  $\eta(x) = a$  inside the brackets, we have

$$\frac{1}{k^2} \sum_{\eta(x)=a} \left[ \sum_{\eta(r)=a+b} \phi_s(r) \right] \quad .$$

Now  $\{x | \eta(x) = a\}$  is the coset  $x+K$  of  $G$  (where  $x$  is such that  $\eta(x) = a$ ) and therefore has cardinality  $k = |K|$ . Therefore (7.3-9) becomes

$$\frac{1}{k^2} \cdot k \sum_{\eta(r)=a+b} \phi_s(r) = \frac{1}{k} \sum_{\eta(r)=a+b} \phi_s(r) = (\text{Agg}_\eta \phi_s)(a+b)$$

which is the desired result.

Therefore we know that for any  $s \in G$ ,  $\text{Agg}_\eta \phi_s$  is a homomorphism from  $D$  into  $R$  and hence is either a character of  $D$  or identically 0. According to Lemma 7.2, however, for all  $a \in D$

$$\text{PE}_\eta \theta_a = \phi_\rho(a) \cdot$$

Aggregating both sides of this equality we have

$$\text{Agg}_\eta \text{PE}_\eta \theta_a = \text{Agg}_\eta \phi_\rho(a) \cdot$$

Since  $\text{Agg}_\eta \text{PE}_\eta$  is the identity on  $F(D, R)$  this means that for all  $a \in D$

$$\theta_a = \text{Agg}_\eta \phi_\rho(a) \cdot$$

These are all the characters of  $D$  so for  $s \in G - \rho(D)$  it must be true that

$$\text{Agg}_\eta \phi_\rho(s) \equiv 0 \cdot$$

Q.E.D.

Applying this result directly to the formula for  $F_G$ , we prove that for any  $f \in F(G, R)$ , the functions  $F_G f$  and  $F_G \text{Proj}_{\rho(D)} f$  are in the same partition of the equivalence relation induced by  $\text{Agg}_\eta$ , i.e. when you aggregate them they look the same. In other words  $\text{Agg}_\eta F_G = \text{Agg}_\eta F_G \text{Proj}_{\rho(D)}$ .

Lemma 7.4

- 1)  $\text{Agg}_{\eta} F_G = \text{Agg}_{\eta} F_G \text{Proj}_{\rho(D)}$
- 2)  $\text{Agg}_{\eta} F_G^{-1} = \text{Agg}_{\eta} F_G^{-1} \text{Proj}_{\rho(D)}$

Proof:

Clearly the operator  $\text{Agg}_{\eta}$  is a module homomorphism (or a linear operator for  $R$  a field). It suffices, therefore, to prove that  $\text{Agg}(F_G - F_G \text{Proj})$  maps every function in  $F(G, R)$  to the zero function of  $F(D, R)$ .

Let  $f \in F(G, R)$  and let  $\alpha \in G$ . We have from (5.2-3) and Def. 6.2 that

$$\begin{aligned}
 [(F_G - F_G \text{Proj})f](\alpha) &= \sum_{\beta \in G} f(\beta) \phi_{\alpha}(\beta) - \sum_{\beta \in \rho(D)} f(\beta) \phi_{\alpha}(\beta) \\
 &= \sum_{\beta \in G - \rho(D)} f(\beta) \phi_{\alpha}(\beta) \\
 &= \sum_{\beta \in G - \rho(D)} f(\beta) \phi_{\beta}(\alpha) \quad \text{since } \phi_{\alpha}(\beta) = \phi_{\beta}(\alpha).
 \end{aligned}$$

Therefore, since  $\text{Agg}$  is a module homomorphism, we have that

$$\begin{aligned}
 \text{Agg}[(F_G - F_G \text{Proj})f] &= \text{Agg} \left( \sum_{\beta \in G - \rho(D)} f(\beta) \phi_{\beta} \right) \\
 &= \sum_{\beta \in G - \rho(D)} f(\beta) \text{Agg} \phi_{\beta}.
 \end{aligned}$$

Now from Lemma 7.3 we know that  $\text{Agg} \phi_{\beta} \equiv 0$  for  $\beta \in G - \rho(D)$ . Therefore

$$[\text{Agg}(F_G - F_G \text{Proj})]f \equiv 0$$

and part 1 is proved. Part 2 is proved analogously with the observation that if  $\beta \in G - \rho(D)$  then  $-\beta \in G - \rho(D)$ , since  $\rho(D)$  is a subgroup of  $G$ .

Q.E.D.

We are now in a position to relate the operators  $F_G$  and  $F_D \text{Res}_\rho(D)$

Theorem 7.3

- 1)  $F_D \text{Res}_\rho(D) = \text{Agg}_\eta F_G$
- 2)  $F_D^{-1} \text{Res}_\rho(D) = k \text{Agg}_\eta F_G^{-1}$

Proof:

To prove part 1 we start with Theorem 7.2 part 1, i.e. we have that

$$\text{PE } F_D = F_G \text{Emb.}$$

Aggregating both sides we arrive at

$$\text{AggPE } F_D = \text{Agg} F_G \text{Emb.}$$

Thus, since AggPE is the identity on  $F(D,R)$  we get

$$F_D = \text{Agg} F_G \text{Emb.}$$

Hence, the equality holds also for restricted functions or

$$F_D \text{Res} = \text{Agg} F_G \text{EmbRes.}$$

Since EmbRes = Proj we get

$$F_D \text{Res} = \text{Agg} F_G \text{Proj} .$$

Finally, applying Lemma 7.4 part 1, we get

$$F_D \text{Res} = \text{Agg} F_G$$

and part 1 is proved. Part 2 follows similarly from Theorem 7.2 part 2 and Lemma 7.4 part 2.

Q.E.D.

Corollary 7.2

$$k F_D \text{Agg}_\eta = \text{Res}_\rho(D) F_G .$$

Proof:

Immediate from Theorem 7.3 part 2 by applying  $F_D$  and  $F_G$  as in the proof of Corollary 7.1.

Q.E.D.

Recalling that operating with  $\text{Res}_{\rho(D)}$  can be thought of as sampling at rate  $k$ , we can think of Theorem 7.3 as saying that:

- 1) *Sampling in space (time) corresponds to aggregating in the spatial (temporal) frequency domain.*
- 2) *Aggregating in space (time) corresponds to sampling in the spatial (temporal) frequency domain (Corollary 7.2).*

In most applications emphasis is placed on statement 1. (This is especially true when the group  $G$  is interpreted as time since sampling is a causal operation but aggregation is not.) The usual statement of this result takes the form of the well-known "sampling theorem" and can be roughly stated as follows (e.g. Schwarz and Friedland [1965]) for functions of  $\mathbb{R}$ :

If a signal has a Fourier transform which vanishes for frequencies above a certain value, then it is completely determined by its values at a denumerable set of equally spaced sampling points. The higher the cut-off frequency, the smaller the sampling intervals.

The emphasis is thus placed on the ability to exactly reconstruct a function from its values at sample points, or, in our terms, to exactly

recover a function  $f \in F(G, R)$  from  $\text{Res}(f) \in F(D, R)$ .

Since both  $F_G$  and  $F_D$  are invertible, it suffices to recover  $\hat{f} = F_G f$  from  $F_D \text{Res}(f)$ , and we can see from Theorem 7.3 part 1 that this involves "disaggregating"  $\text{Agg} F_G f$ . Of course this can't be accomplished in general since the map  $\text{Agg}$  loses information, i.e. is not one-one. If, however, we know a priori that the function  $f \in F(G, R)$  has a non-zero value for exactly one element of every coset of  $K$  and we know which element of the coset it is, then no information is lost by the map  $\text{Agg}$ . It is customary to say that no aliasing occurs. In particular, for  $G$  with the decomposition in (7.2-1), this is true if the support of  $f$  is the set  $W = \{(s_1, \dots, s_d) \mid 0 \leq s_i < m_i\}$ . Noting that  $W$  is the underlying set of  $D$ , it can be seen that

$$\hat{f}(s_1, \dots, s_d) = k(\text{Agg} f)(s_1, \dots, s_d) \text{ for } (s_1, \dots, s_d) \in W.$$

Then given  $\text{Agg}(f)$ , the function  $\hat{f}$  can be obtained by filling in the missing values with zero and dividing by  $k$ . The function  $f$  can be obtained by taking the inverse DFT w.r.t.  $G$ .

A function like  $f$  is called band-limited since its transform has an appropriately limited support. The recovery of a function with no higher frequency components (a smooth function) is very naturally interpreted as an interpolation procedure. However, these smooth functions are not the only kind of band limited functions that can be recovered: they are a special case of functions whose transforms have one non-zero value on each coset of  $K$ . The reconstruction process in these other cases, though, does not resemble interpolation.

Finally, we should point out that there is nothing profoundly unique about band-limited functions that makes their recovery possible. Any function  $f \in F(G, R)$  can be recovered from  $\text{Res}(f)$  provided that we have,

in addition to  $\text{Res}(f)$ , enough other information about  $f$ , namely that we know 1) that  $f$  belongs to a subset of  $F(G,R)$  on which the operator  $\text{Res}$  is invertible and 2) we know how to compute the inverse operator for that subset. The set of band-limited functions is one such subset and is important in applications because membership in it is often a natural consequence of underlying assumptions. For material related to these issues we refer this reader to deBoor's discussion of the method of projections in numerical analysis [1966] and to Zeigler's discussion of justifying conditions for the postulational approach to modeling [1974].

Our application of Theorems 7.2 and 7.3 to linear cellular automata will, however, involve band limited functions since the process of band-limiting (filtering) produces homomorphic images (Theorem 6.2). Thus, it will not be necessary to require a priori knowledge of a function's spectrum since the object is not to reconstruct a function exactly but to recover only enough information for predictive purposes. Whether this situation underlies the importance of band-limited functions in other applications is an interesting question the discussion of which may require careful thought about the process of building models.

Before returning to the theory of LCAs, it's necessary to point out that all the results in this section have been proved only for operators on the spaces  $F(G,R)$  and  $F(D,R)$  where  $R$  is an integral domain containing a principle  $n_d^{\text{th}}$  root of unity and the multiplicative inverse of  $|G|$ . However, using the extensions of  $F_G$  and  $F_D$  to vector and matrix valued functions that we gave in Sec. 5.4, it is not difficult to see that it is possible to extend all of our results to operators



on the spaces  $F(G, \mathbb{R}^n)$ ,  $F(G, \text{Mat}_{\mathbb{R}}(n, n))$ ,  $F(D, \mathbb{R}^n)$ , and  $F(D, \text{Mat}_{\mathbb{R}}(n, n))$ . The theorems proved above are simply applied to all of the component functions.

#### 7.4 Homomorphisms of LCAs via Filtering and Sampling

Let  $S$  and  $D$  be finite abelian groups with canonical decompositions as in (7.3-1) and (7.3-2) respectively. Let the maps  $\rho: D \rightarrow S$  and  $\eta: S \rightarrow D$  be defined according to formulas (7.3-3) and (7.3-4) and let  $K$  be the kernel of  $\eta$ . Assume also that  $|S| = N$ ,  $|D| = M$  and  $|K| = N/M = k$ .

For every LCA  $M$  with structure group  $S$  we define an LCA with structure group  $D$  whose impulse response functions (and hence weighting functions) are simply related to those of  $M$ .

##### Definition 7.3

For any LCA  $M = \langle S, K^n, H \rangle$ ,  $M_R$  is the LCA given by  $\langle D, K^n, k\text{Res}_{\rho(D)} H \rangle$ .

The impulse response functions of  $M_R$  are  $k$  times the restrictions of those of  $M$ . An immediate consequence of Lemma 7.1 is the following:

Lemma 7.5

If  $M = \langle S, K^n, H \rangle$  is a decomposable LCA, then  $M_R$  is decomposable also.

Proof:

The LCA  $M_R$  is decomposable if it satisfies the hypotheses of Theorem 5.4.

We have that  $M$  satisfies them so that  $M_R$  does also by Lemma 7.1.

Q.E.D.

Even though the structures of  $M$  and  $M_R$  are simply related, it is unfortunately not true that a simple relationship between their behaviors can always be found. For certain LCAs  $M$ , however, a simple relationship does exist. We need the following definition:

Definition 7.4

A subset  $W$  of  $S$  is compatible with  $\eta$  whenever  $|W \cap (\eta^{-1}(a))| = 1$  for all  $a \in D$ .

In other words,  $W$  is compatible with  $\eta$  whenever it contains exactly one element of each coset of  $K$ . It's clear that if this is true, then  $|W| = |D|$  and we can write  $W = \{w_a | a \in D\}$  where  $\eta(w_a) = a$ .

Suppose  $M = \langle S, K^n, H \rangle$  is a decomposable LCA for which  $\hat{H}(s) = 0$  for  $s \in S - W$ . One can think of  $M$  as having "smooth" impulse response functions like those of the LCA  $M'$  given in Theorem 6.2. Its cells are insensitive to a portion of the spatial frequency spectrum. Now consider the LCA  $M_R = \langle D, K^n, k\text{Res}_{\rho(D)} H \rangle$ . From Theorem 7.3 part 1 we know that the decomposition of  $\hat{M}_R$  is given by  $M_R = [D, K^n, k\text{Agg}_{\eta G} H]$ .

If  $W$  is compatible with  $\eta$ , then it is not hard to see that each non-zero component of the parallel composition  $\hat{M} = [S, K^n, \hat{H}]$  is isomorphic to a component of  $\hat{M}_R$  and vice-versa. In fact, component  $w_a$  of  $\hat{M}$  is the same as component  $a \in D$  of  $\hat{M}_R$ . However, since the first state of  $\hat{M}$  need not be in  $\text{Proj}_W(F(S, K^n))$ , there is only a homomorphism from  $\hat{M}$  to  $\hat{M}_R$ . This homomorphism becomes  $\text{Res}_{\rho(D)}$  when transformed back into the spatial domain. We then have this commutative diagram:

$$\begin{array}{ccc}
 M & \xrightarrow{\text{Res}_{\rho(D)}} & M_R \\
 \updownarrow F_S & & \updownarrow F_D^{-1} \\
 \hat{M} = [S, K^n, \hat{H}] & \longrightarrow & \hat{M}_R = [D, K^n, k\text{Agg}_{\eta} \hat{H}].
 \end{array}$$

More formally, we can prove the following theorem. Although the proof essentially follows the verbal argument above, we give it in a direct form in order to explicitly show that the homomorphism is the map  $\text{Res}_{\rho(D)}$ .

Theorem 7.4

Let  $W$  be any subset of  $S$  that is compatible with  $\eta$ . Suppose  $M$  is a decomposable LCA given by  $\langle S, K^n, H \rangle$  where  $\hat{H}(s) = 0$  for  $s \in S - W$ . Then  $\text{Res}_{\rho(D)}$  is a homomorphism from  $M$  to  $M_R$ .

Proof:

Let  $\delta$  and  $\delta_R$  be the transition functions of  $M$  and  $M_R$  respectively and let  $q \in F(S, K^n)$ . We need to show that

$$\text{Res}_{\rho(D)}(\delta q) = \delta_R(\text{Res}_{\rho(D)} q).$$

Or, expanding, and dropping unnecessary subscripts,

$$\text{Res}(H *_S q) = (k\text{Res}H) *_D (\text{Res}(q)).$$

By Lemma 7.5,  $M_R$  is decomposable so that it suffices to show (noting that  $F_D$  is linear) that:

$$F_D \text{Res}(H *_S q) = (kF_D \text{Res}H) \cdot (F_D \text{Res}(q)).$$

By Theorem 7.3 part 1 (for vector valued functions) this becomes:

$$\text{Agg}F_S(H *_S q) = (k\text{Agg}\hat{H}) \cdot (\text{Agg}(\hat{q})).$$

Using the convolution theorem we can write:

$$\text{Agg}(\hat{H} \cdot \hat{q}) = (k\text{Agg}\hat{H}) \cdot (\text{Agg}(\hat{q})).$$

Both sides of this equation are functions of  $D$  so for any  $a \in D$  we need

$$(\text{Agg}(\hat{H} \cdot \hat{q}))(a) = (k\text{Agg}\hat{H})(a) \cdot (\text{Agg}(\hat{q}))(a).$$

Using Definition 7.2, this is the same as:

$$\frac{1}{k} \sum_{\eta(r)=a} (\hat{H} \cdot \hat{q})(r) = \left[ k \cdot \frac{1}{k} \sum_{\eta(r)=a} \hat{H}(r) \right] \cdot \left[ \frac{1}{k} \sum_{\eta(r)=a} \hat{q}(r) \right].$$

But since  $W$  is compatible with  $\eta$ ,  $\hat{H}(r) \neq 0$  only in case  $r = w_a$ . We are thus left with the equality:

$$\begin{aligned} \frac{1}{k} (\hat{H} \cdot \hat{q})(w_a) &= \hat{H}(w_a) \cdot \frac{1}{k} \hat{q}(w_a) \\ &= \frac{1}{k} (\hat{H} \cdot \hat{q})(w_a). \end{aligned}$$

Since  $\text{Res}_{\rho(D)}$  is clearly onto  $F(D, K^n)$  we are done.

Q.E.D.

Since the only thing preventing  $\text{Res}_{\rho(D)}$  from being an isomorphism is that  $M$  can start in a larger number of states than  $M_R$  can, we can

prove the following:

Corollary 7.3

Let  $M$  be as in Theorem 7.4. The map  $\text{Res}_{\rho(D)}$  is an isomorphism from the submachine  $M_W$  of  $M$  (see Def. 6.5) to  $M_R$ .

Proof:

It's necessary only to show that  $\text{Res}_{\rho(D)}$  is a bijection when restricted to  $F_S^{-1}\text{Proj}_W(F(S, K^n))$ , the state set of  $M_W$ .

First, to see that  $\text{Res}_{\rho(D)}$  (so restricted) is onto  $F(D, K^n)$  let  $f$  be any function in  $F(D, K^n)$ . Then the function  $F_S^{-1}q$  where

$$q(s) = \begin{cases} k(F_D f)(a) & s = w_a \\ 0 & \text{otherwise} \end{cases}$$

is in  $F_S^{-1}\text{Proj}_W(F(S, K^n))$  and  $\text{Agg}(q) = F_D f$ . From Theorem 7.3 part 1 it can be seen that  $F_D \text{Res} F_S^{-1} = \text{Agg}$ . Thus  $F_D \text{Res} F_S^{-1} q = F_D f$  and since  $F_D$  is one-one,  $\text{Res} F_S^{-1} q = f$ . Therefore  $\text{Res}_{\rho(D)}$  is onto even when restricted to functions in  $F_S^{-1}\text{Proj}_W(F(S, K^n))$ .

Now, to see that  $\text{Res}_{\rho(D)}$  is one-one, suppose  $q, q' \in F_S^{-1}\text{Proj}_W(F(S, K^n))$  and that

$$\text{Res}_{\rho(D)} q = \text{Res}_{\rho(D)} q'.$$

Then

$$F_D \text{Res}(q) = F_D \text{Res}(q')$$

or, by Theorem 7.3 part 1:

$$\text{Agg} F_S q = \text{Agg} F_S q'.$$

Then for all  $a \in D$  we have that

$$(\text{Agg}_{F_S} q)(a) = (\text{Agg}_{F_S} q')(a)$$

or

$$\frac{1}{k} \sum_{\eta(r)=a} \hat{q}(r) = \frac{1}{k} \sum_{\eta(r)=a} \hat{q}'(r) .$$

But, since  $W$  is compatible with  $\eta$ , we have

$$\frac{1}{k} \hat{q}(w_a) = \frac{1}{k} \hat{q}'(w_a) \quad \text{or} \quad \hat{q}(w_a) = \hat{q}'(w_a) \quad \text{for all } w_a \in W.$$

And since  $q, q' \in F_S^{-1} \text{Proj}_W(F(S, K^n))$  we have that  $\hat{q}(s) = \hat{q}'(s) = 0$  for  $s \in S-W$ . Therefore  $\hat{q} = \hat{q}'$  and hence  $q = q'$ .

Q.E.D.

Again, as for the submachine of Theorem 6.2, the submachine  $M_W$  is always entered by the initial transition of  $M$  so that we really need not be concerned with distinction between  $M$  and  $M_W$ . Intuitively (bearing in mind this caveat) these results say that if it is known a priori that an LCA  $M$  has "smooth" weighting functions, then it is the same as an LCA in which the weighting functions are sampled versions of those of  $M$ . This machine will occupy a smaller "space" and will keep exact track of the sampled versions of the states of  $M$ . No information is lost since the cells of  $M$  are insensitive to a part of the spatial frequency spectrum. In the LCA  $M_R$  this insensitivity is manifested as a smaller number of cells.

As indicated by our remarks at the conclusion of Sec. 7.3, however, it is possible to eliminate the requirement of a priori knowledge about the cells of  $M$ . The price paid at this lower level of knowledge is that  $M_R$  will be a homomorphism of  $M$  and will process less information than  $M$ : it will keep exact track of smoothed and sampled states of  $M$ .

To see how this is done, notice that for an LCA to satisfy the hypothesis of Theorem 7.4, it must have impulse response functions each with spectrum having support (contained in) the set  $W$ . The LCA  $M'$  of Theorem 6.2 satisfies this requirement since

$$H' = \xi_W H = [F_S^{-1} \text{Proj}_W F_S] H .$$

Taking  $F_S$  of both sides, we get

$$\hat{H}' = \text{Proj}_W \hat{H}$$

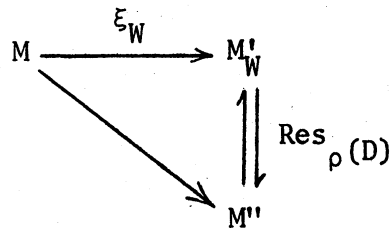
so that  $\hat{H}'(s) = 0$  for  $s \in S - W$ . Moreover,  $M'$  is a homomorphic image of the LCA  $M$  of Theorem 6.2. Using observation we can prove the following result:

Theorem 7.5

Let  $M = \langle S, K^n, H \rangle$  be a decomposable LCA and let  $W$  be any subset of  $S$  compatible with  $\eta$ . Then  $\text{Res}_{\rho(D)} \xi_W$  is a homomorphism from  $M$  to the LCA  $M'' = \langle D, K^n, k\text{Res}_{\rho(D)} \xi_W H \rangle$ .

Proof:

Let  $M'$  be the LCA  $\langle S, K^n, \xi_W H \rangle$ . Theorem 6.2 says that the submachine  $M'_W$  of  $M'$  is a homomorphic image of  $M$ , and from its proof we have that the homomorphism is  $\xi_W : F(S, K^n) \rightarrow F(S, K_n)$ . Since  $\xi_W = F_S^{-1} \text{Proj}_W F_S$ , its clear that  $(F_S \xi_W H)(s) = 0$  for  $s \in S - W$ . Thus, Corollary 7.3 applies to  $M'$ , and since  $M'' = M'_R$ , we have the following commutative diagram.

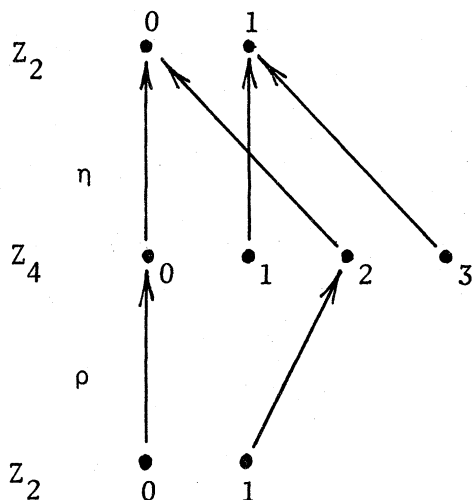


Q.E.D.

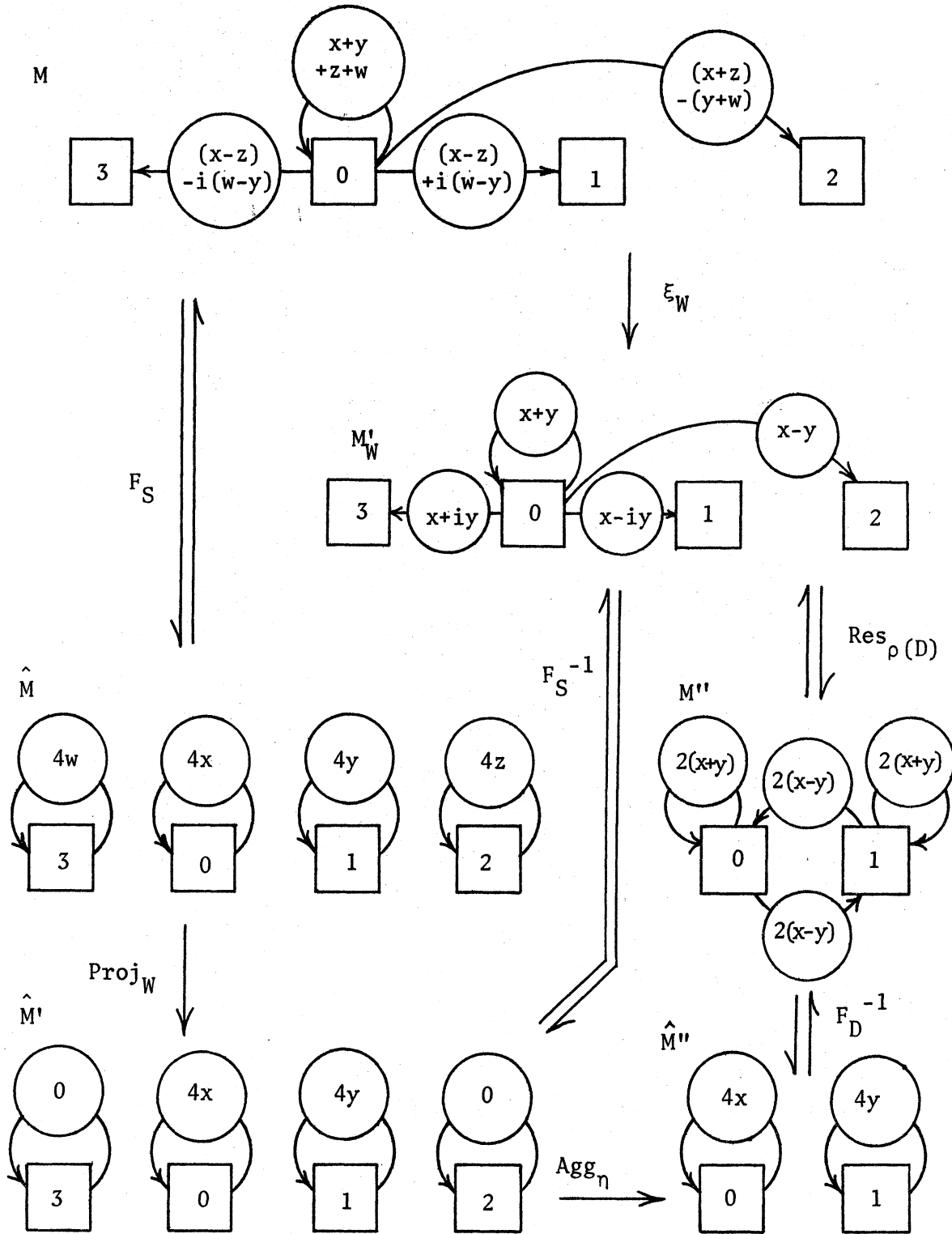
We have succeeded in relating LCAs which have different structure groups. The map  $\text{Res}_{\rho(D)} \xi_W$  is used two ways in Theorem 7.5: first, interpreted as operating on  $n \times n$  matrix valued functions (and multiplied by  $k$ ) it relates the structures of  $M$  and  $M'$ , and, second, operating on vector valued functions, it relates the states of  $M$  and  $M'$  as indicated in the proof. Intuitively, operating with this map first filters a function and then samples it. It is therefore not necessary to know a priori that a state of  $M$  is appropriately band-limited. The homomorphic image  $M'$  will keep track of the states of  $M$  exactly, whether they are band-limited or not, but only to a certain level of resolution. Ex. 7.2 illustrates Theorems 7.4 and 7.5 and their proofs.

### Example 7.2

Consider the LCA  $M = \langle Z_4, \mathbb{C}, H \rangle$  where  $H$  is shown in Fig. 7.2. The group  $D$  is  $Z_2$  and  $W = \{0, 1\}$ . For the sake of clarity, interconnections in  $M$  and  $M'$  are shown only for cell 0. However, it should be kept in mind that these networks resemble the one shown in Fig. 5.2. The maps  $\eta$  and  $\rho$  are as follows:







**Figure 7.2:** Illustration of Theorems 7.4 and 7.5 for  $S = Z_4$ ,  $D = Z_2$ , and  $W = \{0,1\}$ . The maps shown are between state sets.

Since  $\eta$  maps each element in  $W = \{0,1\}$  to a different element of  $Z_2$ ,  $W$  is compatible with  $\eta$ . The maps shown in Fig. 7.2 are those on state sets. Note that the structures of  $\hat{M}$  and  $\hat{M}'$  are related by  $2\text{Agg}_\eta$  and that the structures of  $M$  and  $M'$  are related by  $2\text{Res}_{\rho(D)}$ .

In Sec. 6.5 we showed that the map  $\xi_W$  can be represented more concretely as convolution by  $\chi_W = F_S^{-1} \hat{\chi}_W$  where  $\hat{\chi}_W$  is the characteristic function of the window  $W$ . Thus the map  $\text{Res}_{\rho(D)} \xi_W$  can be given as follows:

$$\begin{aligned} (\text{Res}_{\rho(D)} \xi_W q)(a) &= [\text{Res}_{\rho(D)} (\chi_W *_{\mathcal{S}} q)](a) \\ &= \sum_{r \in \mathcal{S}} \chi_W(r) q(\rho(a) - r) \\ &= \sum_{r \in \mathcal{S}} \chi_W(\rho(a) - r) q(r) . \end{aligned}$$

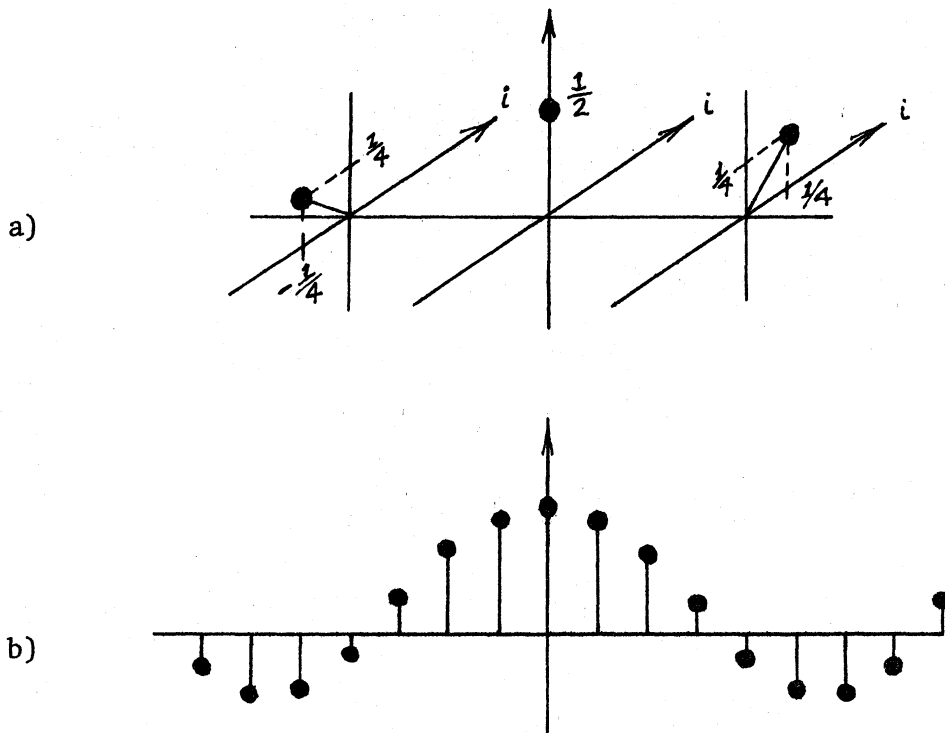
For the special case illustrated in Fig. 7.2,  $\text{Res}_{\rho(D)} \xi_W$  can be taken to map  $\mathbb{C}^4$  into  $\mathbb{C}^2$  and has matrix representation given by:

$$\begin{bmatrix} \frac{1}{2} & \frac{1+i}{4} & 0 & \frac{1-i}{4} \\ 0 & \frac{1-i}{4} & \frac{1}{2} & \frac{1+i}{4} \end{bmatrix} \quad (7.4-1)$$

The second row is the same as the first but moved (cyclically) two spaces to the right. Compare this matrix with matrix (7.2-2).

This kind of operator can be interpreted as follows: to compute  $\text{Res}_{\rho(D)} \xi_W q$  at a point  $a \in D$ , center the function  $\chi_W$  over  $\rho(a) \in \mathcal{S}$ , point-wise multiply, and then sum over  $r \in \mathcal{S}$ . One can think of this operator as an instrument with which one views states of  $M$ . A probe (e.g. a kind of microelectrode) is inserted into a configuration at a point in  $\rho(D)$ ,

i.e. at one of an equally spaced number of sites. Instead of producing a reading that is the state of the cell at that site, though, it produces a value that is a weighted sum over an area around the site. In other words, the probe has a non-local receptive field. The weighting function is determined by the window  $W$ . Fig. 7.3 a) shows the weighting function of the "probe" associated with the map  $\text{Res}_{\rho(D)} \xi_W$  used in Ex. 7.2. It has a complex-valued weighting function which with appropriate shifts, constitutes the rows of matrix (7.4-1). Fig. 7.3 b) shows a more familiar weighting function associated with low-pass filtering. It has real values since the associated characteristic function of the window  $W$  is an even function. For large groups it is possible for  $\hat{\chi}_W$  to be even and  $W$  to be compatible with  $\eta$ .



**Figure 7.3:** Weighting functions associated with filtering and sampling "probes".

We have shown, therefore, that the two objections raised in connection with expression (7.2-1) can be met by a map that relates LCAs with different structure groups. The first objection, i.e. the intuitive complexity of map (7.2-1), is met since the map  $\text{Res}_{\rho(D)} \xi_W$  has an interpretation as "smoothing and sampling" that might provide a plausible model for many kinds of real data gathering techniques. The second objection was that care must be taken in the case of finite fields so that the smaller LCA would be decomposable. This difficulty is avoided for LCAs related by  $\text{Res}_{\rho(D)} \xi_W$  since when  $F_S$  exists, so does  $F_D$ .

There are, however, several facts about the map  $\text{Res}_{\rho(D)} \xi_W$  that create other problems. First, the function  $\chi_W$  will generally have very large support and will always have non-zero values arbitrarily far away. This is because they are inverse transforms of functions with small support. This means that the receptive fields of the corresponding "probes" are arbitrarily wide. For most reasonable cases (e.g. low-pass filtering) this is mitigated somewhat by the fact that these functions have values very close to zero at outlying points (c.f. Fig. 7.3 b)) and might be approximated by more locally supported functions. But the second and more serious problem is that *for the sampling rate required to produce a homomorphism, the receptive fields associated with neighboring sites generally overlap very non-trivially* (an exception is treated in the next section). It is therefore misleading to think of a cell of the homomorphic image  $M'$  as a model of a section or block of cells in a partition of the structure group of the structure group of the original cellular automaton, e.g., in Yamada

and Amoroso's terms [1971], a kernel block (which is not the same as a coset).\*

In fact, as we remarked in Sec. 7.1, care must be taken when cellular automata are simplified by regarding a block of neighboring cells as a single, more complex cell. Such aggregations can indeed give rise to homomorphic images (see Yamada and Amoroso [1971]), but care must be taken to prevent the lumped model from producing artifactual behavior due to aliasing of spatial frequencies. One kind of aggregation that seems natural is to take the state of a block at any time  $t$  to represent the sum of the states at time  $t$  of all the cells that make up that block. In our framework, this kind of aggregation can be viewed as convolving in space with the characteristic function, say  $\lambda$ , of a block, and then sampling at a slow enough rate to remove overlap of blocks in space. This kind of "filtering" and sampling does not cleanly remove high spatial frequencies since the DFT of  $\lambda$  is not a window. In fact, it has very broad support so that the sampling rate required to eliminate overlap of the blocks in space is not fast enough to eliminate aliasing of spatial frequencies (at least for nontrivial blocks, i.e. blocks which are less than the whole space but bigger than single cells. An exception, however, is discussed in the next section).

---

\* Filtering theory as developed for signal processing is concerned primarily with approximating ideal filters with ones that involve weighting functions that are more local, in particular, with ones that can be causally realized (in time). A variety of window shapes (i.e. not characteristic functions) have been investigated which have helpful properties. However, while approximate homomorphisms can be produced using these windows which require less global aggregation, disjoint aggregation always involves undersampling. See Gold and Rader [1969].

These observations are well known among people who are concerned with any form of signal processing. They are related to what is known as the "uncertainty principle" (see Bracewell [1965]) for Fourier transforms. You can either remove overlap in the spatial domain (i.e. produce a blocking of cells) or in the spatial frequency domain (i.e. eliminate aliasing), *but not in both domains at the same time*. To actually filter out the fine spatial detail requires non-local and overlapping "aggregation". Therefore, in order to justify simplifying a system by means of local disjoint aggregates, one generally has to appeal to assumptions other than a combination of linearity and uniformity. An exception to these remarks is discussed next. We describe how a true homomorphic image of an LCA can be produced by certain kinds of structural blockings and aggregations.

### 7.5 Homomorphisms of LCAs via Spatial Aggregation

The operator given by Def. 7.2 was called an aggregation operator because it produces a set of sums over disjoint subsets of  $G$ , namely, the cosets of  $K$ . Applying Corollary 7.2 to an LCA shows that homomorphisms can be produced by aggregation in space. Let  $S, D, \eta$  and  $\rho$  be as defined in the previous section. The kernel of  $\eta$  is  $K$  and  $|K| = k$ .

For every LCA  $M$  with structure group  $S$  we define an LCA with structure group  $D$  that is an aggregated version of  $M$ :

#### Definition 7.5

For any LCA  $M = \langle S, K^n, H \rangle$ ,  $M_A$  is the LCA given by  $\langle D, K^n, k\text{Agg}_\eta H \rangle$ .

An immediate consequence of Lemma 7.1 is the following:

Lemma 7.6

If  $M = \langle S, K^n, H \rangle$  is a decomposable LCA, then  $M_A$  is also decomposable.

We saw in the previous section that under certain conditions on  $M$  the restricted (i.e. sampled) version of  $M$  is a homomorphic image. The situation is more pleasant for the aggregated version of  $M$ : if  $M$  is decomposable then  $M_A$  is a homomorphic image of  $M$ . Intuitively this theorem is true for the following reasons: first, Corollary 7.2 says that aggregating in space corresponds to sampling in the spatial frequency domain; but since the frequency domain representation of an LCA is a parallel composition, selecting any subset of components produces a homomorphic image (cf. Theorem 6.2). Moving back to the spatial domain, this selection appears as aggregation. These ideas are perhaps less well known than those underlying Theorems 7.4 and 7.5 since most interpretations treat  $S$  as a singly generated representation of time. This sort of aggregation in time would always be non-causal (or trivial). In fact, each aggregate would contain information from times arbitrarily far in the future.

Theorem 7.6

If  $M$  is a decomposable LCA, then  $\text{Agg}_\eta$  is a homomorphism from  $M$  onto  $M_A$ .

Proof:

This proof is analogous to that of Theorem 7.4. Let  $\delta$  and  $\delta_A$  be the transition functions of  $M$  and  $M_A$  respectively, and let  $q \in F(S, K^n)$ .

We need to show that

$$\text{Agg}(\delta q) = \delta_A(\text{Agg}(q)).$$

Or, expanding, that

$$\text{Agg}(H *_S q) = (k\text{Agg}H) *_D (\text{Agg}(q)).$$

By Lemma 7.6,  $M_A$  is decomposable so it suffices to show (noting that  $F_D$  is linear) that:

$$F_D \text{Agg}(H *_S q) = (kF_D \text{Agg}H) \cdot (F_D \text{Agg}(q)).$$

By Corollary 7.2 (for vector valued functions) this becomes:

$$\frac{1}{k} \text{Res}_{F_S}(H *_S q) = (\text{Res}\hat{H}) \cdot \left(\frac{1}{k} \text{Res}(\hat{q})\right).$$

Using the convolution theorem we can write:

$$\frac{1}{k} \text{Res}(\hat{H} \cdot \hat{q}) = (\text{Res}\hat{H}) \cdot \left(\frac{1}{k} \text{Res}(\hat{q})\right).$$

From the definition of Res, this is the same as

$$\frac{1}{k} (\hat{H} \cdot \hat{q}) = (\hat{H}\rho) \cdot \left(\frac{1}{k} \hat{q}\rho\right).$$

Both sides of this equation are functions of  $S$  so for any  $s \in S$  we need

$$\frac{1}{k} (\hat{H} \cdot \hat{q})(\rho(s)) = \hat{H}(\rho(s)) \cdot \frac{1}{k} \hat{q}(\rho(s)).$$

But this is just what is meant by pointwise multiplication of functions.

Since Agg is clearly onto  $F(D, K^n)$  we are done.

Q.E.D.

We give two examples of this result. The first shows what happens for a simple cyclic group, and the second shows an application which allows the dimension of the underlying space to be reduced.



Example 7.3

Let  $M = \langle Z_8, \mathcal{C}, H \rangle$ ,  $D = Z_4$ , and the group-homomorphism  $\eta$  be given by:

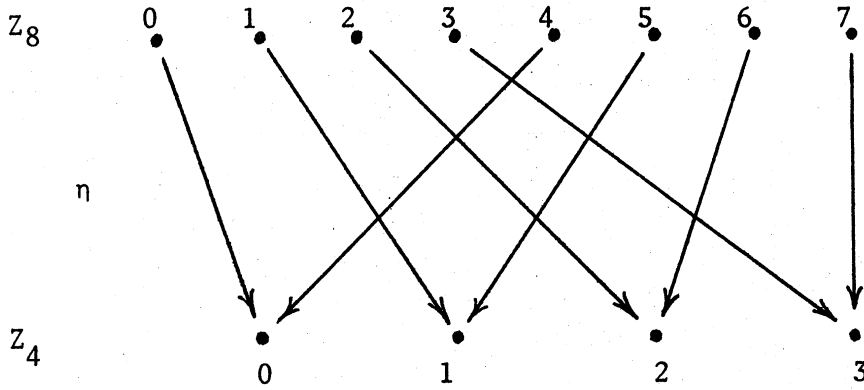
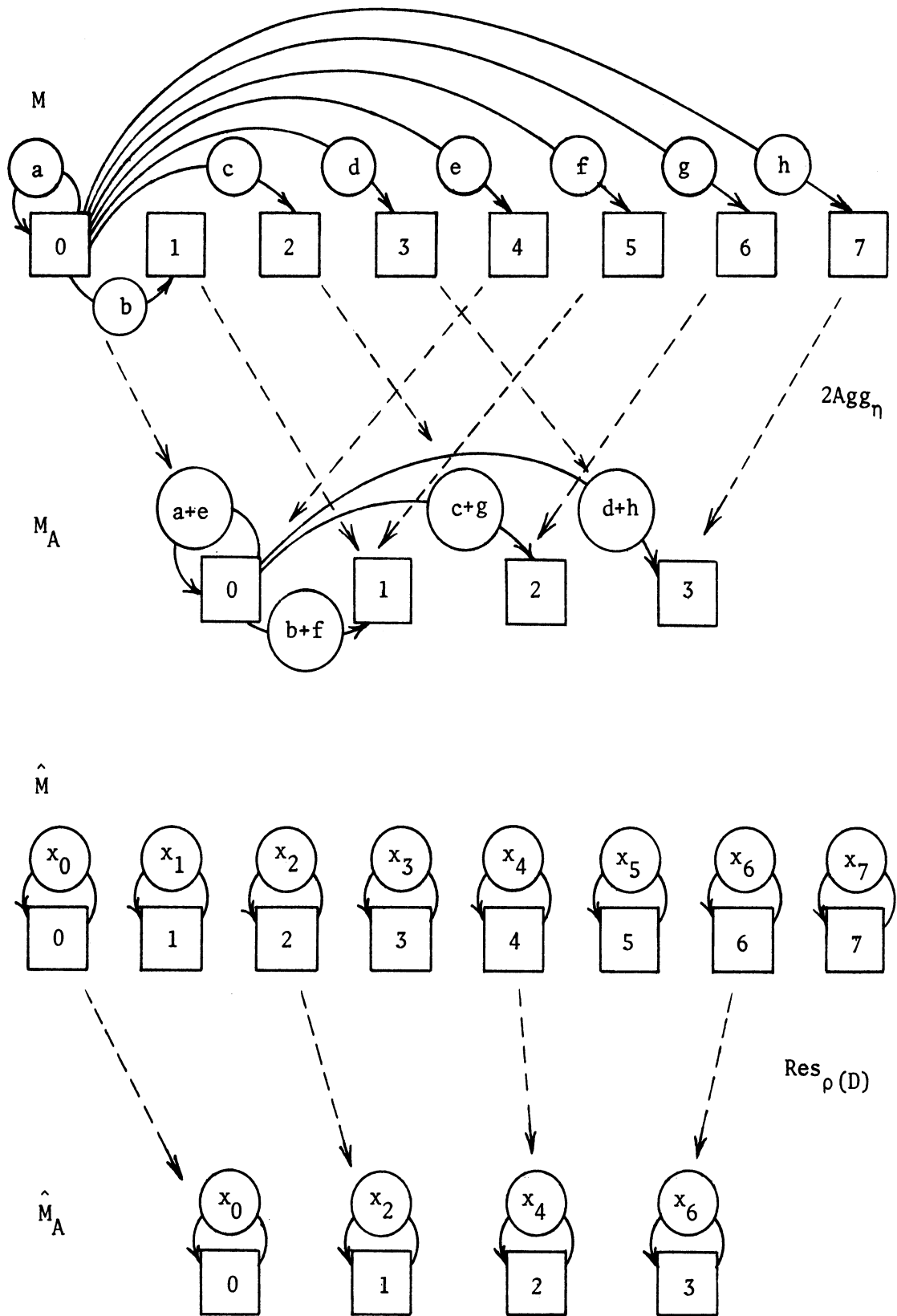


Fig. 7.4 shows  $M$ ,  $M_A$ , and their decompositions. The functions  $H$  and  $2\text{Agg}_\eta H$  are shown for cell 0 only for the sake of clarity.



**Figure 7.4:** A homomorphism from spatial aggregation:  $M_A$  is a homomorphic image of  $M$ .

For the simple cyclic case shown in Fig. 7.4, it can be seen that each cell of  $M_A$  keeps track of the sum of the states of evenly spaced cells in  $M$ , i.e. those cells in a coset of  $K$ . Thus, while the blocks in the aggregation are disjoint, they do not consist of contiguous cells. However, this kind of aggregation has a very useful interpretation for simulation purposes. Suppose one is interested in experimenting with an LCA that has a very large cyclic structure group; one that is so large, in fact, that it can be considered infinite, i.e. non-cyclic. Suppose also that storage and/or time constraints prevent the simulation of any LCA with structure group bigger than  $Z_N$ . If the cells of the original LCA have neighborhoods small enough (e.g. in Fig. 7.4  $d = g = f = 0$ ), then the LCA having the same kind of cells as the larger LSM but with structure group  $Z_N$  is a natural candidate for simulation.

According to Theorem 7.6, this smaller LCA  $M_A$  is a homomorphic image of the large one,  $M$ , via the map  $\text{Agg}_\eta$ . What inferences can be made about the behavior of  $M$  by experimenting with  $M_A$ ? First, it is intuitively clear that for initial configurations with small enough support, the behavioral trajectories of  $M_A$  and  $M$  will be essentially identical until the configurations begin to wrap around  $Z_N$ . So much is also true for non-linear cellular automata. In the case of LCAs, when the configurations wrap around they superimpose (overlap and add). When this happens, the homomorphism  $\text{Agg}_\eta$  begins to act non-trivially by actually losing information, i.e. by producing aliasing in the spatial domain.

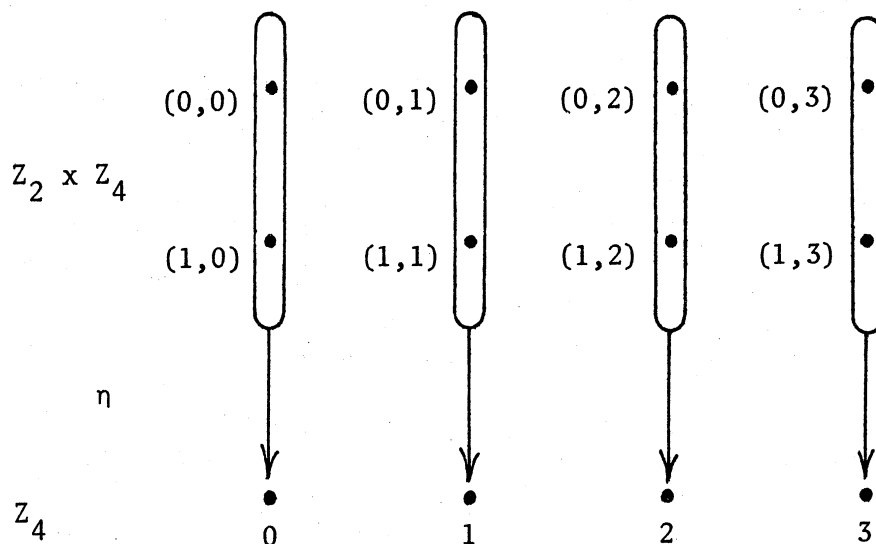
In the spatial frequency domain, however, the situation appears differently. The decomposition of  $M_A$  is a sampled version of the

decomposition of  $M$  as in Fig. 7.4. The larger  $N$  is, the closer these samples are taken. Thus, if a very large LCA is to be specified and/or observed in the spatial frequency domain, it is not misleading to specify and/or observe a frequency sampled LCA.

The significance of these facts is that for simulation purposes our theoretical restriction to LCA with finite structure groups is no real limitation. Keeping in mind the action of the homomorphism  $\text{Agg}_\eta$  in both the spatial and spatial frequency domains, one finds that the infinite case, while perhaps more aesthetically satisfying, does not promise an increase in the understanding of linear cellular automata. The next example shows how Theorem 7.6 can be used to reduce the dimension of an LCA.

#### Example 7.4

Let  $M = \langle \mathbb{Z}_2 \times \mathbb{Z}_4, \mathbb{C}, H \rangle$ ,  $D = \mathbb{Z}_4$ , and  $\eta$  given by



The map  $\eta$  projects the elements of  $\mathbb{Z}_2 \times \mathbb{Z}_4$  onto the second coordinate and is clearly a group-homomorphism. Fig. 7.5 shows  $M$ ,  $M_A$ , and their decompositions.

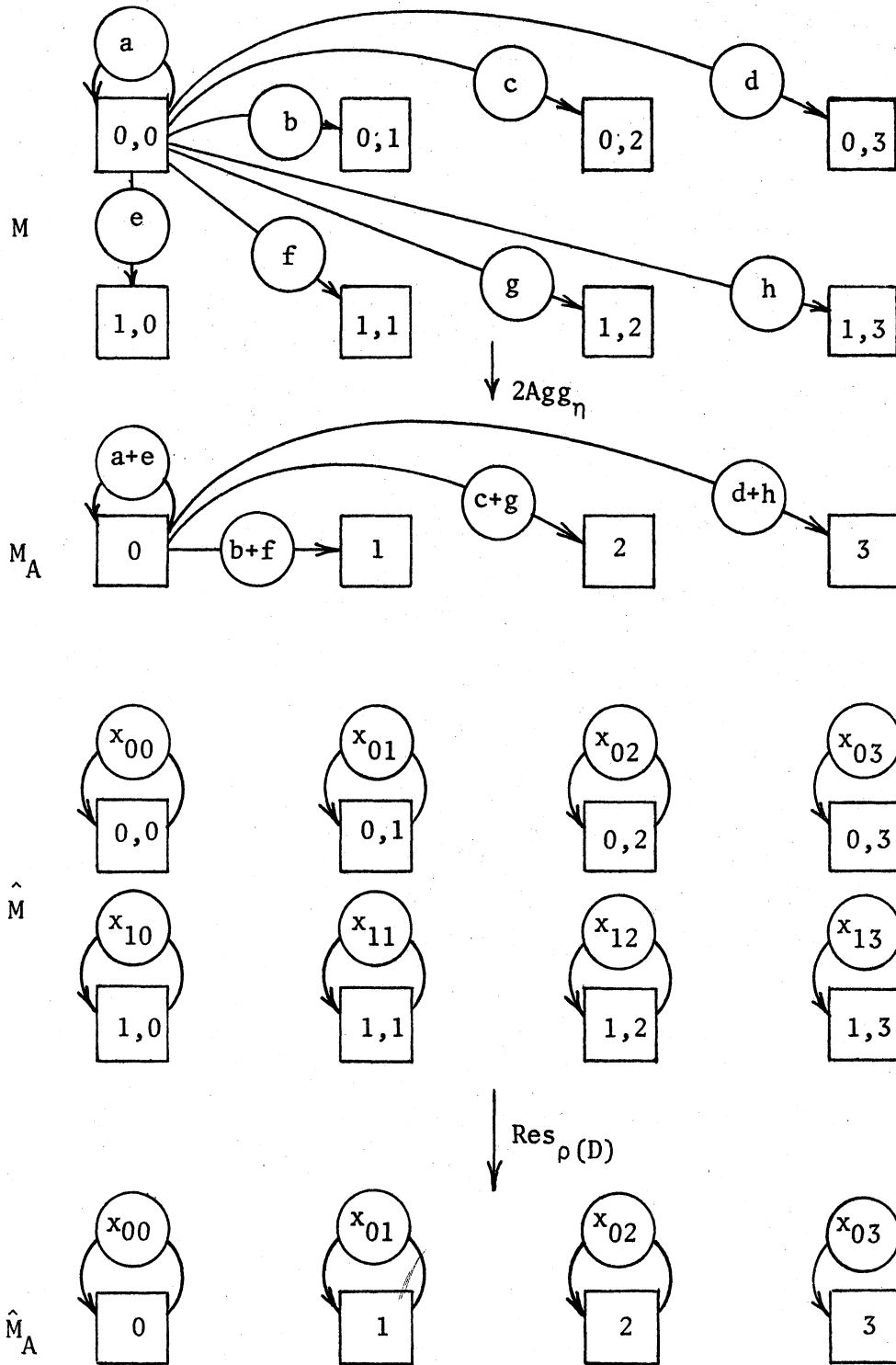


Figure 7.5: Reducing the dimension of an LCA by spatial aggregation.

Compare this figure with Fig. 7.4:  $M$  of Fig. 7.4 and  $M$  of Fig. 7.5 have the same number of cells but have different structures since  $Z_8$  and  $Z_2 \times Z_4$  are not isomorphic and groups (no single generator exists for  $Z_2 \times Z_4$ ). Although the diagrams for these two LCAs appear to be the same except for a rearrangement of the cells, when the connections for all of the cells are filled in, the difference in structure becomes clear. For example, in Fig. 7.4 cell 1 influences cell 0 through the coefficient  $h$ , but in Fig. 7.5 cell (0,1) influences cell (0,0) through the coefficient  $d$ . Both of these LCAs, however, have the same aggregated version  $M_A$ . For  $M$  of Fig. 7.5, though, the cells in a block are adjoining in the structure specified by  $Z_2 \times Z_4$ .

In addition to being an automaton-homomorphism from  $M$  to  $M_A$ , the map  $\text{Agg}_\eta : F(S, K^n) \rightarrow F(D, K^n)$  given in Theorem 7.6 is an example of what Zeigler calls a strong structure preserving morphism. See Zeigler [1971, 1974]. This concept consists of three parts. First, there is a coordinate mapping that maps the coordinate set of  $M$  onto the coordinate set of  $M_A$ . In our case, these coordinate sets are the structure groups  $S$  and  $D$  respectively, and  $\eta$  is the coordinate mapping. It induces a partition of  $S$  consisting of the cosets of  $K$  (the kernel of  $\eta$ ). Secondly, there is a family of coordinate state set mappings. In our case, each of these maps simply averages the states of cells in a coset of  $K$  to produce the state of a cell in  $M_A$ . These maps are required to produce an automaton-homomorphism when acting together. Theorem 7.6 says that this is true since  $\text{Agg}_\eta$  is the result of averaging over each coset of  $K$ . Thus, in Zeigler's terminology,  $\text{Agg}_\eta$  is a weak structure preserving morphism. If, in addition to these properties, a third requirement is met, then  $\text{Agg}_\eta$  is a strong structure preserving morphism. This requirement is that the coordinate mapping (here  $\eta$ ) is a digraph morphism. This means that there is a connection between two cells of  $M_A$  only if there are cells of  $M$  in the corresponding blocks that are connected. In other words, the aggregation process does not introduce misleading dependencies. This is true in our case since the connection coefficients in  $M_A$  are sums of the appropriate connection coefficients in  $M$ . Dependences can disappear (sum to 0) but cannot arise ( $0+0 = 0$ ).

In the next section we complete our application of sampling theory to LCAs by pointing out a connection between a well known fact about

DFTs and an idea that developed in an automata theoretic context.

## 7.6 Laminated LCAs

Yamada and Amoroso [1974] introduced the terms laminations and laminal subautomata for describing the structure of certain cellular automata. For the case of LCAs some potentially interesting connections can be made between these concepts and some well known facts from harmonic analysis. We will describe these concepts as they specialize to the case of finite abelian groups noting that nothing fundamental is lost by abandoning their module theoretic framework.

We pointed out in Chapter 3 that the group graph underlying a cellular automaton is completely connected if and only if the neighborhood template  $X$  contains a set of generators for the structure group  $S$ . If this is not the case, then the cellular automaton is said to be laminated. In general,  $X$  will generate a subgroup of  $S$ . Suppose this subgroup is  $\rho(D)$  for some group  $D$  according to the definitions of Sec. 7.3.

Consider an LCA  $M = \langle S, K^n, H \rangle$  where the support of  $H$  is  $\bar{X}$  where  $X$  generates  $\rho(D)$ . This means that  $X$  is the neighborhood template of  $M$ . Then it should be clear that  $\text{Proj}_{\rho(D)} H = H$ , and that cells in different cosets of  $\rho(D)$  in  $S$  cannot communicate with each other in any number of time steps. There are  $|S|/|D| = k$  distinct cosets of  $\rho(D)$ , and therefore  $k$  different laminal subautomata of  $M$  can be defined. We summarize these ideas using our notation by the following definition. It is essentially the same as that of Yamada and Amoroso except that it allows the possibility of a laminal subautomaton being itself laminated. Let  $S, D \neq S, \rho, \eta$  and  $k$  be as defined in Sec. 7.3 and denote



cosets of  $\rho(D)$  by  $\Lambda_0, \dots, \Lambda_{k-1}$ .

Definition 7.6

An LCA  $M = \langle S, K^n, H \rangle$  is said to be laminated w.r.t.  $D$  if  $\text{Proj}_{\rho(D)} H = H$ . For  $i = 0, \dots, k-1$ ,  $M(\Lambda_i)$  is the subautomaton of  $M$  with state set  $\text{Proj}_{\Lambda_i}(F(S, K^n))$ . These subautomata are called laminal subautomata of  $M$ .

Since cells associated with different cosets cannot influence each other,  $M$  is isomorphic to the parallel composition of the  $M(\Lambda_i)$ . The isomorphism is given by

$$q \mapsto (\text{Proj}_{\Lambda_0} q, \text{Proj}_{\Lambda_1} q, \dots, \text{Proj}_{\Lambda_{k-1}} q) \text{ for } q \in F(S, K^n).$$

Yamada and Amoroso go further and point out that each laminal subautomaton is isomorphic to a nonlaminal automaton. Using our notation this result means that for each  $i = 0, \dots, k-1$ ,  $M(\Lambda_i)$  is isomorphic to the LCA  $M_L = \langle D, K^n, \text{Res}_{\rho(D)} H \rangle$ . If  $M_L$  happens to be laminated also, further restrictions will finally result in a nonlaminal automaton, i.e. one whose neighborhood contains a generating set for its structure group. It's interesting to note that the isomorphism from  $M(\Lambda_i)$  to  $M_L$  can be interpreted simply as a change of spatial scale. It can also be seen that  $\text{Res}_{\rho(D)}$  is an automaton-homomorphism from  $M$  to  $M_L$  (this follows from the fact that  $\text{Res}_{\rho(D)} \text{Proj}_{\rho(D)} = \text{Res}_{\rho(D)}$ ).

All of the facts described above can be formulated and proved for arbitrary cellular automata. What additionally happens in the linear case? It turns out that a laminated LCA has a specific kind of decomposition.

Theorem 7.7

Suppose  $M = \langle S, K^n, H \rangle$  is a decomposable LCA that is laminated w.r.t.  $D$ . Then  $\hat{M} = [S, K^n, \hat{H}]$  where  $\hat{H} = PE_{\eta} F_D Res_{\rho(D)} H$ .

Proof:

From Theorem 7.2 part 1 we have that

$$F_G Emb_{\rho(D)} = PE_{\eta} F_D.$$

Then

$$F_G Emb_{\rho(D)} Res_{\rho(D)} = PE_{\eta} F_D Res_{\rho(D)},$$

or, since

$$Emb_{\rho(D)} Res_{\rho(D)} = Proj_{\rho(D)}, F_G Proj_{\rho(D)} = PE_{\eta} F_D Res_{\rho(D)}.$$

But since  $M$  is laminated w.r.t.  $D$ , we have

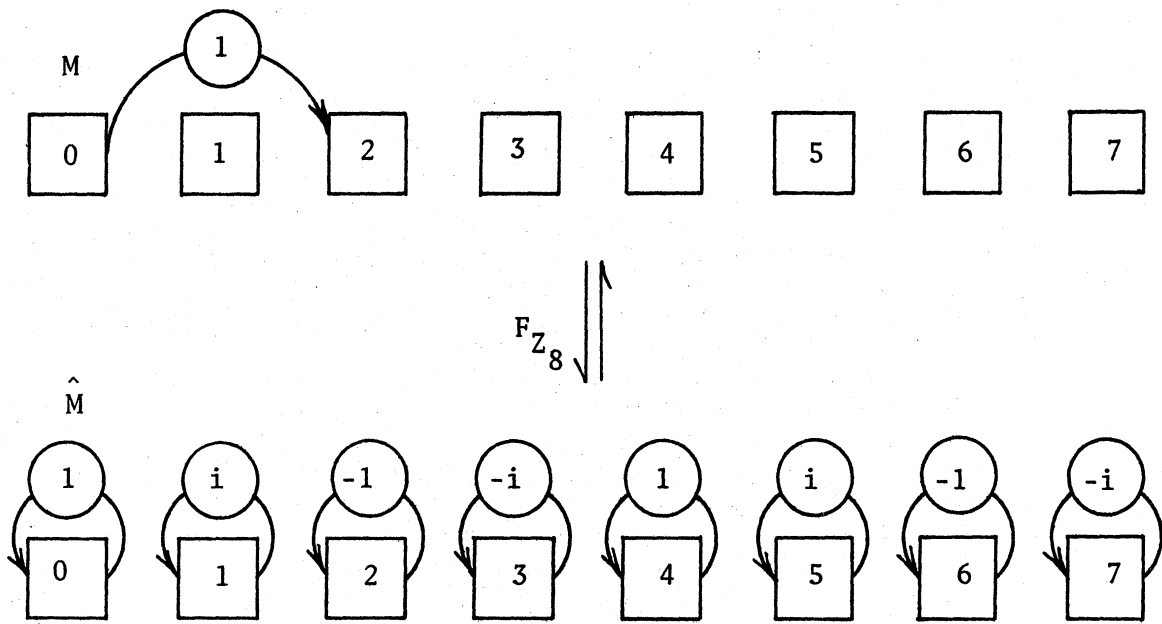
$$F_G Proj_{\rho(D)} H = F_G H = PE_{\eta} F_D Res_{\rho(D)} H.$$

Q.E.D.

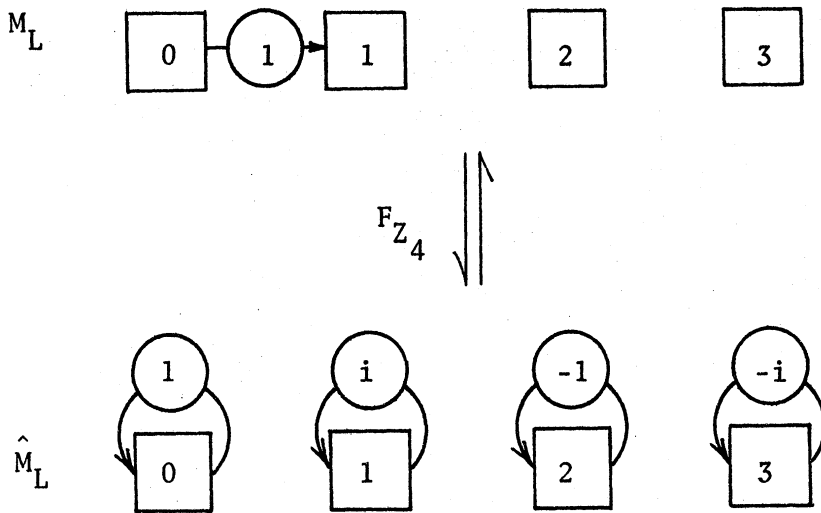
This result says that the decomposition of  $M$  consists of  $k$  identical sets of components where each set is the parallel decomposition of the LCA that we've called  $M_L$ , i.e. the re-scaled version of a laminal subautomaton. Using the terminology of Cooley, Lewis, and Welch [1969],  $M_L$  can be "stretched" to become a laminal subautomaton of  $M$ . Ex. 7.5 illustrates the situation for an LCA like those discussed in Ex. 4.4 and Ex. 5.6.

Example 7.5

Let  $M = \langle Z_8, \mathbb{C}, H \rangle$  be the LCA shown together with its decomposition  $\hat{M}$  in Fig. 7.6 a). Since each cell only depends on the cell two spaces to its left, it is laminated w.r.t.  $D = Z_4$ . The re-scaled LCA  $M_L$  and its decomposition  $\hat{M}_L$  are shown in Fig. 7.6 b). Note that  $\hat{M}$  simply consists of two copies of  $\hat{M}_L$  operating side-by-side.



a)  $M$  is laminated w.r.t.  $Z_4$ .



b) The LCA  $M_L$  and its decomposition.

Figure 7.6: Illustration of Theorem 7.7

Recall that in Ex. 4.4 we discussed an LCA like that shown in Fig. 7.6 b) (but with infinite structure group) as representing a finite difference scheme for the simplified wave equation (4.5-2). Call this LCA  $\tilde{M}$ . We remarked that the approximation produced by this scheme has no truncation error. Of course, this is a trivial observation since the solution of initial value problem (4.5-2) can be explicitly written down, but it provides a point of contact between the terminology of numerical analysis and that of automata theory. If the PDE in (4.5-2) were thought of as specifying a cellular automaton with structure group  $\mathbb{R}$ , then the statement that the LCA  $\tilde{M}$  represents an approximation with no truncation error means that  $\tilde{M}$  is a homomorphic image of the PDE-specified cellular automaton. In fact, for the simplified wave equation,  $\tilde{M}$  is isomorphic to a laminal subautomaton of the continuous cellular automaton. If  $\rho: \mathbb{Z} \rightarrow \mathbb{R}$  is the inclusion map, then this homomorphism is  $\text{Res}_{\rho(D)}$ . In other words,  $\tilde{M}$  is related to the PDE-specified cellular automaton in the same way that  $M_L$  is related to  $M$  in Fig. 7.6. It seems possible that a theory of cellular automata generalized along the lines discussed in Sec. 3.6 would support a rigorous development of this observation.

## Chapter 8

### Linear Cellular Automata as Models of Non-uniform Structures

#### 8.1 Introduction

In constructing models of systems whose components are distributed in space, whether the models are expressed as cellular automata or as partial differential equations, it is often assumed that small spatial non-uniformities exhibited by the "real" system can be ignored without causing the model to produce misleading behavior. This is presupposed by the choice of a cellular automaton model (according to our definition); and, in the case of linear PDEs, leads to an equation with constant coefficients. For the automata theoretic formulation, this simplifying premise is adopted because it simplifies the programming of a simulation model and permits a sometimes necessary reduction in time and/or storage requirements. In the case of PDEs, the assumption of constant coefficients leads to the possibility of applying powerful analysis techniques. In fact, many of these techniques are based on the ability to "separate the variables" of linear PDEs with constant coefficients (and for certain boundary conditions), which is the continuous analog of our parallel decomposition result for LCAs (Theorem 5.4).

Our intuition tells us that if we're interested in large scale spatial behavior, such as the speed and position of wave fronts, then any fine spatial inhomogeneities can be ignored. Formally, this intuitive feeling might be expressed in terms of cellular automata and automaton-homomorphisms: a map which preserves only large-scale aspects of configurations is a homomorphism from certain kinds of spatially inhomogeneous structures to certain cellular automata. In

this chapter we provide a way of looking at spatially inhomogeneous linear structures (i.e. general linear networks) which allows us to specify conditions under which they can be homomorphically modeled by LCAs. As in Chapter 7, the overall point of view follows Zeigler's concepts of base and lumped models (Zeigler [1974]). We will postulate that the "real" system has a certain formal structure (base model) and ask when certain classes of models (lumped models) can be expected to preserve aspects of its behavior. Our results show that for certain kinds of spatial non-uniformities, the intuition often employed in formulating uniform models can be rigorously defended.

## 8.2 Non-uniform Linear Cellular Automata

The title of this section is an inconsistent usage of our terminology, but we use it to emphasize the peculiar point of view to be adopted. By the definition given in Chapter 4, an LCA is a linear sequential machine that has a uniform structure. In this chapter we will study arbitrary linear sequential machines but we will view them within the framework used to study uniform structures. Instead of viewing them as abstract LSMs we will view them as LCAs having certain kinds of non-uniformities. The situation is analogous to standard treatments of time varying linear systems (Chen [1970], Schwarz and Friedland [1965]) in which behavioral trajectories are still regarded as functions of an underlying group (time) even though the system is not uniform with respect to the group (i.e. is not time invariant). In fact, the results we will derive here can automatically be regarded as results for time-varying linear systems with the appropriate considerations of causality and the usual non-cyclic models of time.

Let  $S$  be a finite abelian group and  $K$  a field.

Definition 8.1

A non-uniform linear cellular automaton (NULCA) is an autonomous linear sequential machine  $M = \langle Q, \delta \rangle$  which is associated with a finite abelian group  $S$  in the following way:

- i)  $Q = F(S, K^n)$  for some non-negative integer  $n$ .
- ii) Let  $B_\alpha : S \rightarrow \text{Mat}_K(n, n)$  for each  $\alpha \in S$ .

The transition function  $\delta : Q \rightarrow Q$  of  $M$  is given by:  $\delta(q) =$

where

$$q'(s) = \sum_{r \in S} B_{s-r}(r)q(s-r) \quad (8.1-1)$$

for all  $s \in S$ .

We will denote the NULCA  $M$  by  $\langle S, K^n, \{B\} \rangle$ .

What are these functions  $B_\alpha, \alpha \in S$ , and how do they relate to the function  $H$  of an LCA? To answer these questions suppose  $q \in F(S, K)$  is given by

$$q(s) = \begin{cases} 1 & \text{for } s = \alpha \\ 0 & \text{otherwise} \end{cases}$$

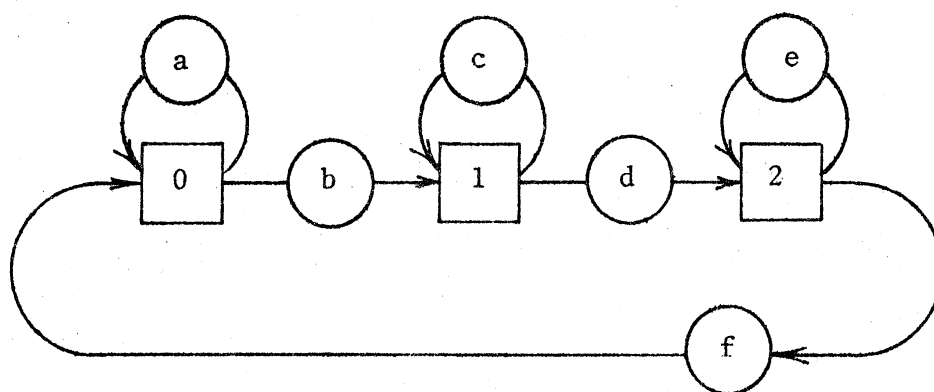
Then according to (8.1-1)  $\delta(q) = q'$  where

$$q'(s) = \sum_{r \in S} B_{s-r}(r)q(s-r) = B_\alpha(s-\alpha) = (L_{-\alpha} B_\alpha)(s) .$$

Thus, the response to an impulse at cell  $\alpha$  is  $L_{-\alpha} B_\alpha$  so that the component functions of  $B_\alpha$  are not the impulse response functions for cell  $\alpha$ . Since  $L_\alpha L_{-\alpha} B_\alpha = B_\alpha$ , they are instead the impulse response functions shifted to the left, i.e., shifted to the origin. Equivalently,

$B_\alpha$  specifies the impulse response of cell  $\alpha$ , but the coordinate system is translated so that cell  $\alpha$  is regarded as located at the origin.

It's important to understand this point so the reader is urged to study the example shown in Fig. 8.1.



	$B_0$	$IR_0$	$B_1$	$IR_1$	$B_2$	$IR_2$
0	a	a	c	0	e	e
1	b	b	d	c	f	0
2	0	0	0	d	0	f

**Figure 8.1:** The functions  $B_\alpha$  and the impulse response functions  $IR_\alpha$ ,  $\alpha \in Z_3$ , for a NULCA

It can be seen, then, that a NULCA is an LCA  $M = \langle S, K^n, H \rangle$  when  $B_\alpha = H$  for all  $\alpha \in S$ , since when this is true equation (8.1-1) is the same as (4.2-3).<sup>\*</sup> Note that the reflection  $\bar{B}_\alpha$  of a function  $B_\alpha$

<sup>\*</sup>This is not a necessary condition for a NULCA  $M'$  to be an LCA since  $M'$  might be uniform with respect to a group other than  $S$  and have cell state set  $K^m$  where  $m \neq n$ . In fact, according to our definition, any LSM is uniform over the trivial group.



(  $\bar{B}_\alpha(s) = B_\alpha(-s)$  ) does not correspond to what we've called the weighting function of cell  $\alpha$ . This can be seen in Fig. 8.1:  $\bar{B}_0 = (a,0,b)$ , but its weighting function is  $(a,0,f)$ .

It is also instructive to compare our definition with the usual representation of linear operators on function spaces. Suppose the underlying group is still  $S$ . Then an arbitrary linear operator  $A$  is represented by a function of  $S \times S$  called the kernel of the operator (which is not related to the kernel of a group-homomorphism). Let  $A: S \times S \rightarrow \text{Mat}_K(n,n)$  denote this kernel. For  $q \in F(S, K^n)$ ,  $Aq = q'$  where  $q'(s) = \sum_{r \in S} A(s,r)q(r)$  for all  $s \in S$ . In this case the

kernel is the same as a matrix where  $A(s,r)$  is the entry in the  $s^{\text{th}}$  row and  $r^{\text{th}}$  column. Usually the term kernel is reserved for situations when the group is  $\mathbb{R}$  and the summation is integration - but it plays the same role as a matrix.

Using this notation, a linear operator on  $F(S, K^n)$  with kernel  $A$  can be viewed as the transition function of a NULCA by defining each  $B_\alpha$ ,  $\alpha \in S$ , as follows:

$$B_\alpha(s-\alpha) = A(s,\alpha) \quad (8.1-2)$$

or

$$B_\alpha(s) = A(s+\alpha,\alpha) \quad \text{for all } s \in S.$$

In fact, it should be clear that any linear operator on a finite dimensional vector space can be viewed as a NULCA. If  $N$  is the dimension of the vector space, take  $S$  to be an abelian group with  $|S| = N$  (e.g.  $Z_N$ ) and derive the functions  $B_\alpha$ ,  $\alpha \in S$ , from the matrix of the operator.

### 8.3 Harmonic Analysis of Non-uniform LCAs

Harmonic analysis techniques are generally reserved for the study of structures with the kind of uniformity we've been discussing here - invariance with respect to a group of shift operators. It is the very frequent assumption of this kind of uniformity, in fact, that makes harmonic analysis techniques so widely used. Other kinds of structural constraints on linear systems make different techniques applicable. For example, systems with spherical symmetry can be analyzed using Bessel and Legendre functions (Duff and Naylor [1966]). Beckmann's book on orthogonal polynomials [1973] describes many other assumptions that give rise to analogous techniques. It therefore seems to be mathematically unusual to apply the Fourier transform to systems that are not uniform in the appropriate way. However, the Fourier transform still can be used to study NULCAs; and, although parallel decompositions are not produced as they are in the uniform case, it is still possible to derive results that have interesting consequences for model building.

In what follows we will restrict our attention to NULCAs which have only one-dimensional "cells", i.e. which have state set  $F(S,K)$ . It seems likely that our results can be extended to the case of n-dimensional cells, but we would like to avoid the notational complexity of such an extension.

Suppose  $S$  is a finite abelian group with  $|S| = N$  and exponent  $m$ . Let  $K$  be any field that contains a principle  $m^{\text{th}}$  root of unity. Then the group characters of  $S$  in  $K$  exist and are given by expression (5.2-2) where  $\omega$  is the root of unity. Then we know (Theorem 5.1) that  $F_S$ , the discrete Fourier transform w.r.t.  $S$ , exists and is invertible. If  $M = \langle S, K, \{B_\alpha\} \rangle$  is a NULCA where  $S$  and  $K$  are as we have

just described, then  $F_S$  is an automaton-isomorphism from  $M$  to another linear sequential machine which we'll call  $\hat{M}$ . Unlike the case of LCAs, however,  $\hat{M}$  is not necessarily a parallel decomposition of  $M$ . If  $A$  is the linear operator with kernel (i.e. matrix) given by (8.1-2), then the transition function of  $\hat{M}$  is the linear operator  $F_S A F_S^{-1}: F(S,K) \rightarrow F(S,K)$ . This operator may not necessarily have a diagonal matrix representation (as it does in the uniform case). We will therefore call the NULCA  $M$  a transformable NULCA, and the linear sequential machine  $\hat{M}$ , the transform of  $M$  since the terms "decomposable" and "parallel decomposition" would be misleading.

Instead of trying to relate the entire structure of a transformable NULCA  $M$  and its transform  $\hat{M}$ , we will only show that certain reasonable constraints on the structure of  $M$  can insure that  $\hat{M}$  is of a special form with useful properties. In the next section, we use this fact to show how non-trivially uniform homomorphic images of NULCAs can be constructed.

For a transformable NULCA with one-dimensional "cells" we now show that if all the functions  $B_\alpha$ ,  $\alpha \in S$ , have the same  $\beta^{\text{th}}$  Fourier coefficient, then the  $\beta^{\text{th}}$  coefficient of any configuration at time  $t+1$  depends only on the  $\beta^{\text{th}}$  coefficient of the configuration at time  $t$ . In order to prove this result, it is convenient to introduce a standard notation for an inner product on the linear space  $F(S,K)$ . For  $f, g \in F(S,K)$ , let

$$\langle f, g \rangle = \sum_{s \in S} f(s)g(s).$$

Using this notation, the orthogonality of the group characters of  $S$  in  $K$  stated in Sec. 5.2 becomes

$$\langle \phi_{-s}, \phi_r \rangle = \begin{cases} N & \text{if } r = s \\ 0 & \text{otherwise.} \end{cases} \quad (8.3-1)$$

The discrete Fourier transform w.r.t.  $S$  (which exists and is invertible) can be written

$$F_S f = g \quad \text{where} \quad g(s) = \langle f, \phi_s \rangle \quad \text{for all } s \in S. \quad (8.3-2)$$

### Theorem 8.1

Let  $M = \langle S, K, \{B_\alpha\} \rangle$  be a transformable NULCA with transition function  $\delta$ . If for any  $\beta \in S$  there is a  $c_\beta \in K$  such that  $\hat{B}_\alpha(\beta) = c_\beta$  for all  $\alpha \in S$  (i.e. the  $\beta^{\text{th}}$  Fourier coefficients of all the  $B_\alpha$ 's are the same), then

$$(\delta(q))^\wedge(\beta) = c_\beta \hat{q}(\beta) \quad \text{for any } q \in F(S, K).$$

**Proof:**

By the definitions of  $F_S$  and  $F_S^{-1}$  and since  $\delta$  is linear, we have for any  $q \in F(S, K)$

$$\begin{aligned} (\delta(q))^\wedge(\beta) &= \langle \delta(q), \phi_\beta \rangle \\ &= \langle \delta\left(\frac{1}{N} \sum_{\alpha \in S} \hat{q}(\alpha) \phi_{-\alpha}\right), \phi_\beta \rangle \\ &= \left\langle \frac{1}{N} \sum_{\alpha \in S} \hat{q}(\alpha) \delta(\phi_{-\alpha}), \phi_\beta \right\rangle \\ &= \frac{1}{N} \sum_{\alpha \in S} [\langle \delta(\phi_{-\alpha}), \phi_\beta \rangle \hat{q}(\alpha)]. \end{aligned} \quad (8.3-3)$$

Now consider only the inner product  $\langle \delta(\phi_{-\alpha}), \phi_\beta \rangle$ .

Let  $\delta^T$  be such that

$$\langle \delta(\phi_{-\alpha}), \phi_\beta \rangle = \langle \phi_{-\alpha}, \delta^T(\phi_\beta) \rangle. \quad (8.3-4)$$

From standard finite dimensional linear algebra we know that  $\delta^T$  exists and  $\delta^T(\phi_\beta)$  is given by

$$(\delta^T(\phi_\beta))(s) = \sum_{r \in S} \beta_s(r-s) \phi_\beta(r) \quad \text{for } s \in S.$$

This is true because from (8.1-1) we have

$$\begin{aligned} (\delta(\phi_\beta))(s) &= \sum_{r \in S} \beta_{s-r}(r) \phi_\beta(s-r) \\ &= \sum_{r \in S} \beta_r(s-r) \phi_\beta(r) \end{aligned}$$

by a change of variable. Then  $\delta^T$ , the transpose of  $\delta$ , is obtained by interchanging  $r$  and  $s$  in  $\beta_r(s-r)$ .

By another change of variable, we get

$$\begin{aligned} (\delta^T(\phi_\beta))(s) &= \sum_{r \in S} \beta_s(r) \phi_\beta(r+s) \\ &= \sum_{r \in S} \beta_s(r) \phi_\beta(r) \cdot \phi_\beta(s) \end{aligned}$$

since  $\phi_\beta$  is a homomorphism from  $S$  into the multiplicative structure of  $K$ . Thus, by (8.3-2), we have

$$\begin{aligned} (\delta^T(\phi_\beta))(s) &= \phi_\beta(s) \langle \beta_s, \phi_\beta \rangle \\ &= \phi_\beta(s) \hat{\beta}_s(\beta). \end{aligned}$$

Returning to (8.3-4) we therefore have that

$$\begin{aligned} \langle \delta(\phi_{-\alpha}), \phi_\beta \rangle &= \langle \phi_{-\alpha}, \delta^T(\phi_\beta) \rangle \\ &= \sum_{s \in S} \phi_{-\alpha}(s) \phi_\beta(s) \hat{\beta}_s(\beta) \end{aligned}$$

$$\begin{aligned}
&= \sum_{s \in S} \phi_{-\alpha}(s) \phi_{\beta}(s) c_{\beta} \\
&= c_{\beta} \langle \phi_{-\alpha}, \phi_{\beta} \rangle .
\end{aligned}$$

Going back to (8.3-3) we have that

$$\begin{aligned}
(\delta(q))^{\wedge}(\beta) &= \frac{1}{N} \sum_{\alpha \in S} [\langle \delta(\phi_{-\alpha}), \phi_{\beta} \rangle \hat{q}(\alpha)] \\
&= \frac{1}{N} \sum_{\alpha \in S} [c_{\beta} \langle \phi_{-\alpha}, \phi_{\beta} \rangle \hat{q}(\alpha)] \\
&= c_{\beta} \frac{1}{N} \sum_{\alpha \in S} \langle \phi_{-\alpha}, \phi_{\beta} \rangle \hat{q}(\alpha) .
\end{aligned}$$

Hence, since the characters are orthogonal in the sense given by expression (8.3-1), this becomes

$$\begin{aligned}
(\delta(q))^{\wedge}(\beta) &= c_{\beta} \frac{1}{N} \cdot N \hat{q}(\beta) \\
&= c_{\beta} \hat{q}(\beta)
\end{aligned}$$

which is the desired result.

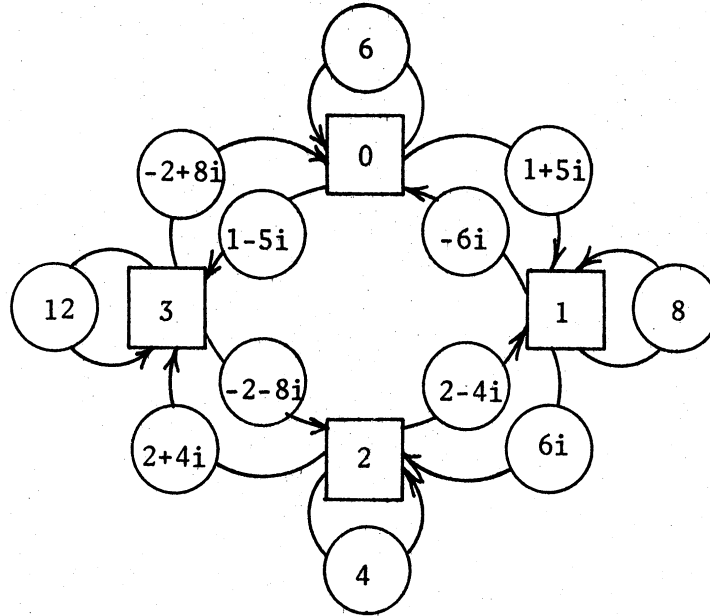
Q.E.D.

We illustrate this theorem by the following example:

### Example 8.1

Consider the NULCA  $M = \langle Z_4, \mathbb{C}, \{B_{\alpha}\} \rangle$  shown in Fig. 8.2 a). Since the underlying field is  $\mathbb{C}$ ,  $M$  is transformable. We've contrived the functions  $B_{\alpha}, \alpha=0, \dots, 3$ , so that they have the same  $\beta^{\text{th}}$  Fourier coefficients for  $\beta=0, 1$ . Therefore, according to Theorem 8.1, the transform of  $M$  is such that components 0 and 1 do not receive input from any components but themselves. This is shown in Fig. 8.2 b). Note that

a)  $M$



	$B_0$	$B_1$	$B_2$	$B_3$	$\hat{B}_0$	$\hat{B}_1$	$\hat{B}_2$	$\hat{B}_3$
0	6	8	4	12	8	8	8	8
1	$1+5i$	$6i$	$2+4i$	$-2+8i$	-4	-4	-4	-4
2	0	0	0	0	4	8	0	16
3	$1-5i$	$-6i$	$2-4i$	$-2-8i$	16	20	12	28

b)  $\hat{M}$

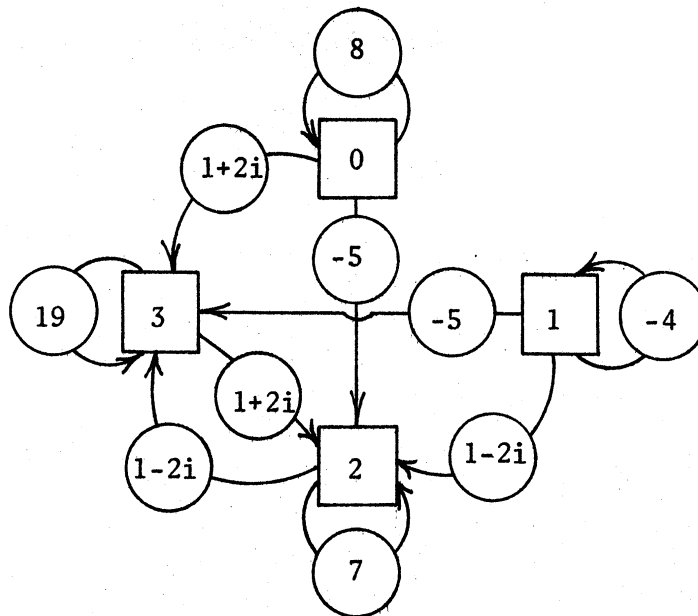


Figure 8.2: A NULCA  $M$  and its transform  $\hat{M}$ .

Theorem 8.1 says that components 0 and 1 are not influenced by other components. It does not imply that they may not influence other components. Note also that component 0 depends on itself via the Fourier coefficient  $\hat{B}(0)$  that is common to all of the  $B_\alpha$ 's and that component 1 depends on itself via  $\hat{B}(1)$ .

#### 8.4 Uniform Homomorphic Images of NULCAs

The automaton-homomorphism  $\xi_W$  of Theorem 6.2, which maps states of an LCA  $M$  onto the "smoother" states of an LCA  $M'$ , exists because the parallel composition  $\hat{M}'$  consists of a subset of the independent components of  $\hat{M}$ . More specifically, however, the map  $\xi_W$  is an automaton-homomorphism because those components of  $\hat{M}$  retained for the composition  $\hat{M}'$  do not depend on components not retained. For a parallel composition, any subset of components forms a homomorphic automaton since its components depend only on themselves. It is clear, then, that a more general situation can be discussed.

In particular, if  $M = \langle S, K, \{B_\alpha\} \rangle$  is a transformable NULCA such that the  $B_\alpha$ 's share some subset of Fourier coefficients, Theorem 8.1 says that each component of the transform  $\hat{M}$  which corresponds to one of these mutual coefficients does not receive input from any other component. It is therefore possible to form a homomorphic image of  $\hat{M}$  by retaining only those components of  $\hat{M}$  corresponding to the mutual Fourier coefficients. Moreover, since these components do not depend on one another, they form a parallel composition that has a distributed realization with structure group  $S$ . The NULCA  $M$ , therefore, has a structurally uniform homomorphic image that can be produced by filtering out the aspects of its behavior which result from its



structural non-uniformity.

Since any NULCA is automatically uniform with respect to the trivial structure group, we will be concerned with homomorphic images that possess, in some sense, maximal uniformity. Let us, therefore, explicitly define the maximal subset of  $S$  on which the spectra of all of the  $B_\alpha$ 's take on the same values.

Definition 8.2

For a transformable NULCA  $M = \langle S, K, \{B_\alpha\} \rangle$ , let  $U$  be the subset of  $S$  such that  $s \in U$  if and only if  $\hat{B}_\alpha(s) = \hat{B}_\beta(s)$  for all  $\alpha, \beta \in S$ .

We have then for all  $\alpha, \beta \in S$  that

$$\hat{B}_\alpha|_U = \hat{B}_\beta|_U,$$

and hence the following definition makes sense:

Definition 8.3

For  $M$  and  $U$  as in Def. 8.2, let  $H_U \in F(S, K)$  be such that

$$\hat{H}_U = \text{Proj}_U \hat{B}_\alpha$$

where  $\alpha \in S$ .

From Theorem 8.1 we can prove the following result:

Theorem 8.2

Let  $M = \langle S, K, \{B_\alpha\} \rangle$  be a transformable NULCA. Then  $\xi_U = F_\xi^-$  is a homomorphism from  $M$  onto a submachine of the LCA  $M^U = \langle S, K \rangle$ .

Proof:

It will suffice to show that  $\text{Proj}_U$  is a homomorphism from  $\hat{M}$ , the transform of  $M$ , to the submachine  $\hat{M}_U^U$  of the parallel decomposition

$\hat{M}^u$  with state set  $\text{Proj}_U(F(S,K))$  (see Def. 6.4). Theorem 8.2 will follow from composing the maps in this diagram:

$$\begin{array}{ccc}
 M & \xrightarrow{\xi_U} & M_U^u \\
 \uparrow F_S & & \uparrow F_S^{-1} \\
 \hat{M} & \xrightarrow{\text{Proj}_U} & \hat{M}_U
 \end{array}$$

Let  $\hat{\delta}$  and  $\hat{\delta}_U$  be the transition functions of  $M$  and  $M^u$  respectively.

Then we need to show that

$$\text{Proj}_U \hat{\delta} = \hat{\delta}_U \text{Proj}_U.$$

More explicitly, we need for all  $q \in F(S,K)$  and  $s \in S$  that

$$(\text{Proj}_U \hat{\delta} q)(s) = (\hat{\delta}_U \text{Proj}_U q)(s)$$

First, assume  $s \in S-U$ . We know from Theorem 5.4 that

$$(\hat{\delta}_U \text{Proj}_U q)(s) = \hat{H}_U(s) \cdot (\text{Proj}_U q)(s) = 0 \cdot 0 = 0.$$

By the definition of  $\text{Proj}_U$  we have also that

$$(\text{Proj}_U \hat{\delta} q)(s) = 0 \quad \text{for } s \in S-U.$$

Therefore suppose  $s \in U$ . This means that  $\hat{B}_\alpha(s) = \hat{H}_U(s)$  for all  $\alpha \in S$ . If  $\delta$  denotes the transition function of  $M$ , since  $M \xrightleftharpoons{F_S} \hat{M}$  we know that

$$F_S \delta = \hat{\delta} F_S$$

or

$$F_S \delta F_S^{-1} = \hat{\delta}.$$

Thus we can apply Theorem 8.1 and get the following: for  $s \in U$

$$\begin{aligned}
(\text{Proj}_U \hat{\delta}q)(s) &= (\hat{\delta}q)(s) \\
&= (F_S \delta F_S^{-1} q)(s) \\
&= (\delta(F_S^{-1} q)) \hat{\delta}(s) \\
&= \hat{H}_U(s) \cdot q(s) .
\end{aligned}$$

But  $\hat{\delta}_U$  is such that for  $s \in U$  we have

$$\begin{aligned}
(\hat{\delta}_U \text{Proj}_U q)(s) &= \hat{H}_U(s) \cdot (\text{Proj}_U q)(s) \\
&= \hat{H}_U(s) \cdot q(s) .
\end{aligned}$$

Since  $\text{Proj}_U$  is clearly onto the set  $\text{Proj}_U(F(S,K))$ , we have that  $\text{Proj}_U$  is an automaton-homomorphism from  $\hat{M}$  to the submachine of  $\hat{M}_U^u$  given in Def. 6.4. We conclude that  $\xi_U = F_S^{-1} \text{Proj}_U F_S$  is a homomorphism from  $M$  to the submachine of  $M_U^u$  of  $M^u$  having state set  $F_S^{-1} \text{Proj}_U(F(S,K))$

Q.E.D.

This theorem is illustrated in Fig. 8.3 where we have represented the LCA  $M^u$  and its decomposition  $\hat{M}^u$  that corresponds to the NULCA of Ex. 8.1 and Fig. 8.2. Since  $M_U$  is uniformly interconnected, we show  $H_U$  for cell 0 only. Compare the parallel composition  $\hat{M}^u$  to the transform  $\hat{M}$  of  $M$  shown in Fig. 8.2 b). One can see that at any time  $t \in \mathbb{Z}^+$  the state of component 0 in  $\hat{M}^u$  will be exactly the same as the state of component 0 in  $\hat{M}$ . Similarly, components 1 will agree for all time  $t \in \mathbb{Z}^+$ .

The map  $\xi_U = F_S^{-1} \text{Proj}_U F_S$  of Theorem 8.2 is exactly the same as the map  $\xi_W$  for  $W = U$  given in Theorem 6.2. Unlike the uniform case, however, not every subset of  $S$  will produce a homomorphism. It should be clear, though, that for any subset  $W$  of  $S$  that is also a subset of  $U$ , the map  $\xi_W$  is a homomorphism from the NULCA  $M$  to an LCA  $M'$ .

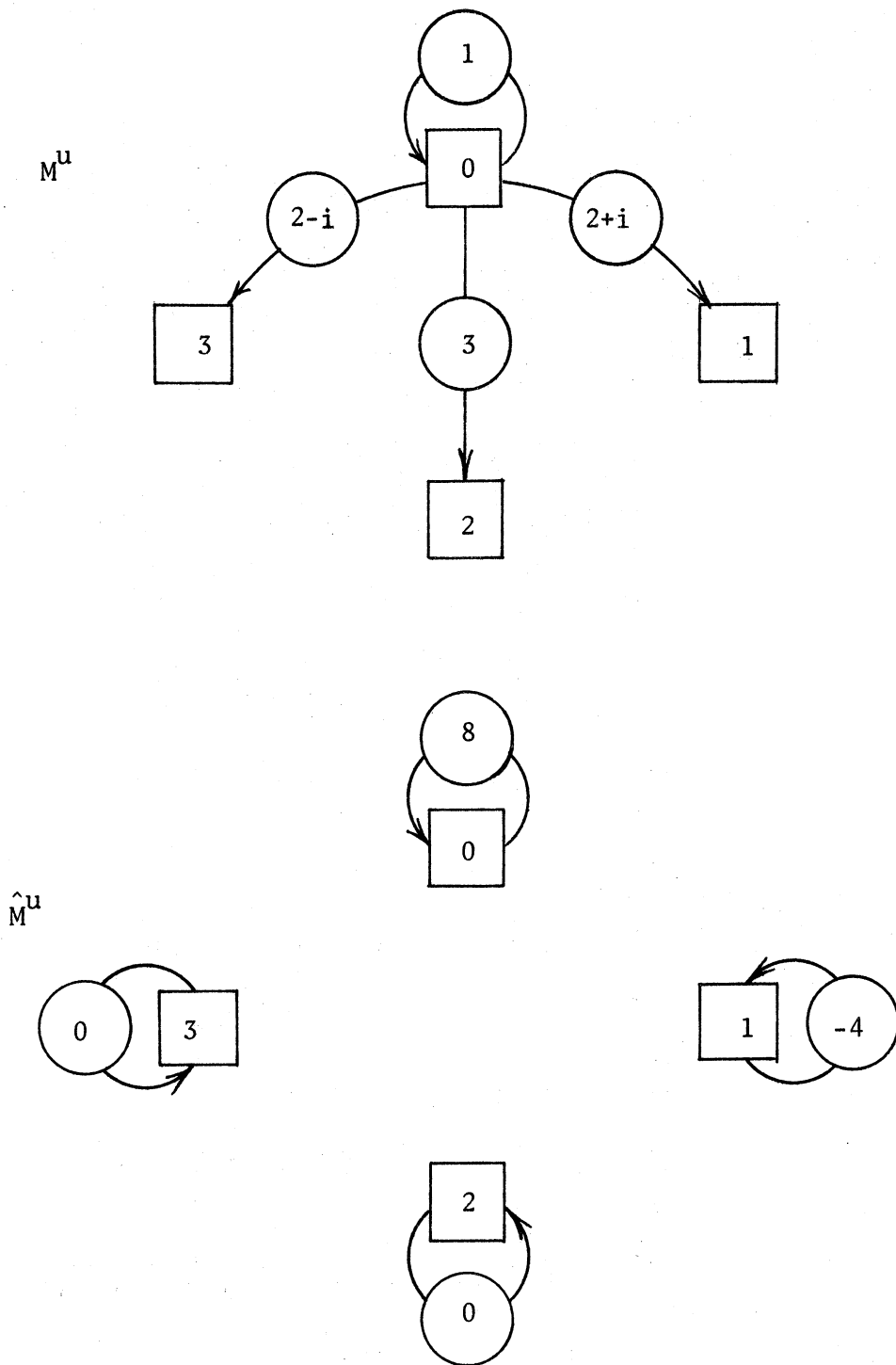


Figure 8.3: The LCA  $M^u$  is a structurally uniform homomorphic image of the NULCA  $M$  shown in Fig. 8.2 a).

We make this explicit in the following corollary:

Corollary 8.1

Let  $M = \langle S, K, \{B_\alpha\} \rangle$  be a transformable NULCA. Then for any  $W \subseteq U$  (where  $U$  is given by Def. 8.2),  $\xi_W$  is a homomorphism from  $M$  to the submachine  $M'_W$  of the LCA  $M' = \langle S, K, H' \rangle$  where  $H' = \xi_W H$ .

Proof:

The LCA  $M^U$  given by Theorem 8.2 satisfies the hypotheses of Theorem 6.2. We therefore have

$$M \xrightarrow{\xi_U} M^U \xrightarrow{\xi_W} M'_W .$$

Thus,  $\xi_W \xi_U$  is the required homomorphism

$$\begin{aligned} \xi_W \xi_U &= F_S^{-1} \text{Proj}_W F_S F_S^{-1} \text{Proj}_U F_S \\ &= F_S^{-1} \text{Proj}_W \text{Proj}_U F_S \\ &= F_S^{-1} \text{Proj}_W F_S \\ &= \xi_W \end{aligned}$$

since  $W \subseteq U$ .

Q.E.D.

This corollary is the generalization to NULCA's of Theorem 6.2. Notice that if the NULCA  $M$  happens also to be an LCA, then  $U = S$  since all the  $B_\alpha$ 's are the same functions, namely they are all equal to the impulse response function  $H$  of the LCA. Therefore, the requirement that  $W \subseteq U$  reduces to the requirement that  $W \subseteq S$ , and Corollary 8.1 reduces to Theorem 6.2. In a similar manner, all the results proved in Chapter 7 for LCAs can be generalized to the case

of NULCAs. We make these generalizations explicit in the next section.

### 8.5 Homomorphisms of NULCAs produced by Filtering-Sampling and Aggregation

Let  $M = \langle S, K, \{\hat{B}_\alpha\} \rangle$  be a transformable NULCA, and let  $M^u = \langle S, K, H_U \rangle$  be the LCA of Theorem 8.2. That is, for  $U$  given by Def. 8.2,  $\hat{H}_U = \text{Proj}_U \hat{B}_\alpha$  for any  $\alpha \in S$ . Let  $S$  and  $D$  be finite abelian groups with canonical decompositions as in (7.3-1) and (7.3-2) respectively. Let the maps  $\rho: D \rightarrow S$  and  $\eta: S \rightarrow D$  be as defined by (7.3-3) and (7.3-4), and let  $K$  be the kernel of  $\eta$  with  $|K| = k$ .

Def. 7.3 can be applied to the LCA  $M^u$  to define the sampled version  $M_R^u$  of  $M^u$ . More specifically:

#### Definition 8.4

For  $M^u = \langle S, K, H_U \rangle$ , the LCA  $M_R^u$  is given by  $\langle D, K, k\text{Res}_{\rho(D)} H_U \rangle$ .

Recalling that a subset  $W$  of  $S$  is compatible with  $\eta$  whenever  $|W \cap (\eta^{-1}(a))| = 1$  (Def. 7.4), we can prove a generalization of Theorem 7.5:

#### Theorem 8.3

Let  $W$  be any subset of  $U$  that is compatible with  $\eta$ . Then  $\text{Res}_{\rho(D)} \xi_W$  is a homomorphism from the NULCA  $M$  onto the LCA  $M_R^u$ .

Proof:

Theorem 8.2 says that

$$M \xrightarrow{\xi_W} M_U^u,$$

and Corollary 7.3 applies to  $M^u$  so that we also have

$$M_U^u \xrightleftharpoons{\text{Res}_{\rho(D)}} M_R^u .$$

Thus, we have the following commutative diagram:

$$\begin{array}{ccc} M & \xrightarrow{\xi_W} & M_U^u \\ & \searrow & \updownarrow \text{Res}_{\rho(D)} \\ & & M_R^u \end{array}$$

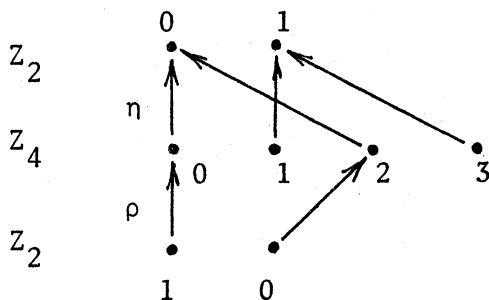
Q.E.D.

The above result is a generalization of Theorem 7.5 since the smoothed and sampled LCA  $M' = \langle S, K, k\text{Res}_{\rho(D)} \xi_W H \rangle$  of that theorem is, for the case considered here, the same as  $M_R^u = \langle D, K, k\text{Res}_{\rho(D)} H_U \rangle$ . The reason for this is that  $\xi_W H_U = H_U$  due to the fact that  $H_U$  is already a smooth function by virtue of its definition:  $\hat{H}_U = \text{Proj}_U \hat{B}_\alpha$ . In other words, the "filtering" part of Theorem 7.5 has already been accomplished in passing from the NULCA  $M$  to the LCA  $M^u$ .

The next example illustrates Theorem 8.3 and its proof:

Example 8.2

We will again consider the NULCA  $M = \langle Z_4, \mathbb{C}, \{B_\alpha\} \rangle$  shown in Fig. 8.2 a). For this system, the set  $U$  is  $\{0,1\}$ . Suppose  $D = Z_2$ , and the maps  $\eta$  and  $\rho$  are as follows:



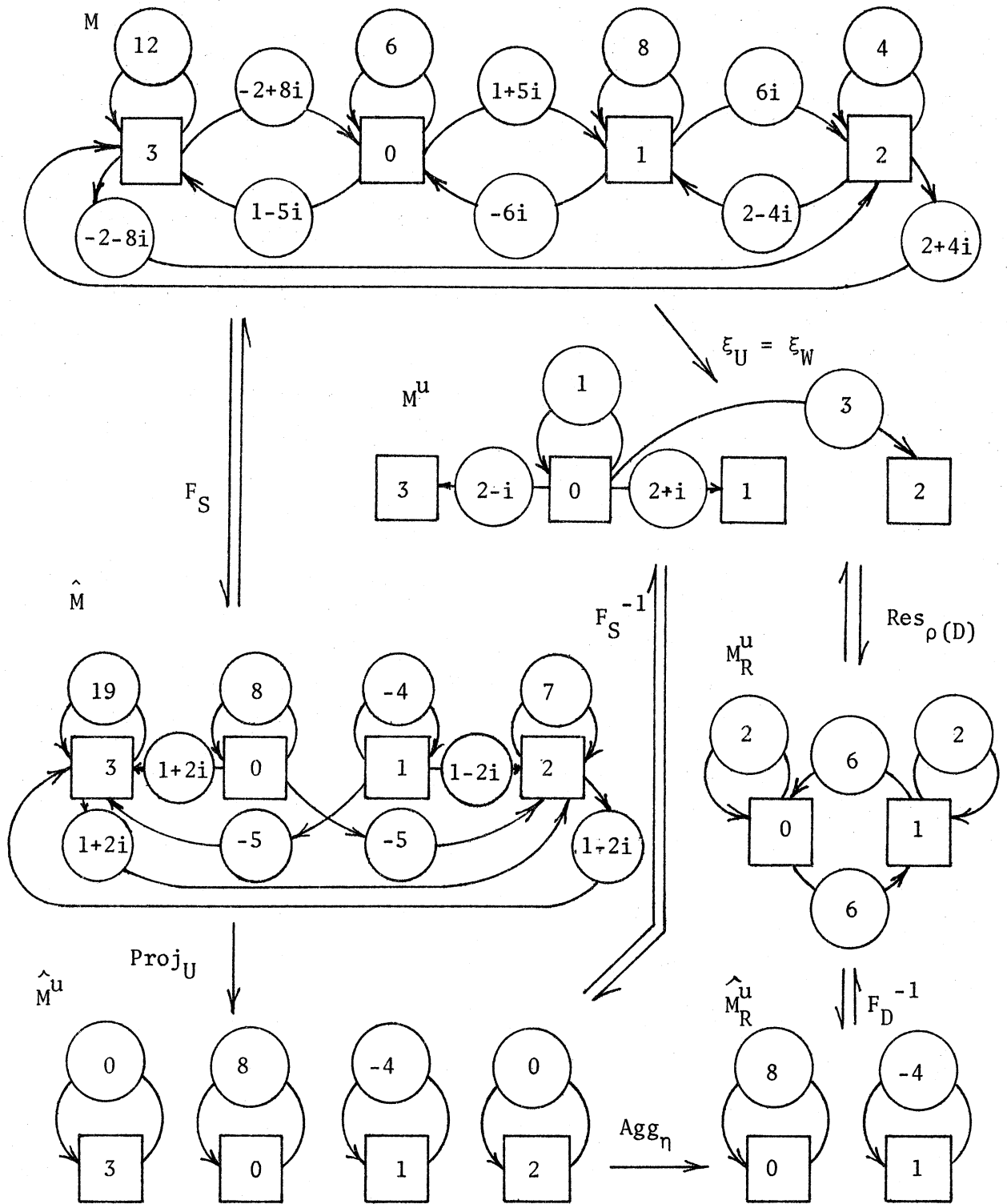


Figure 8.4: Illustration of Theorem 8.3 for  $S = \mathbb{Z}_4$ ,  $D = \mathbb{Z}_2$ , and  $W = U = \{0, 1\}$ . The maps shown are between state sets. Compare this figure with Fig. 7.2.



Therefore  $U$  is compatible with  $\eta$  so that Theorem 8.3 holds for  $W = U = \{0,1\}$ . Fig. 8.4 shows the systems  $M$ ,  $M^u$ ,  $M_R^u$  and their decompositions. The reader should compare Fig. 8.4 with Fig. 7.2 which illustrates the analogous facts for LCAs. The only difference between these figures is that in Fig. 8.4  $M$  is not uniformly interconnected and its transform  $\hat{M}$  is not a parallel composition. The crucial fact provided by Theorem 8.1 is that  $\text{Proj}_U$  in Fig. 8.4 is still an automaton-homomorphism.

We should point out that the homomorphism  $\text{Res}_{\rho(D)} \xi_W$  of Theorem 8.3 is exactly the same map that we discussed in Sec. 7.4. as a homomorphism between two LCAs. Thus, even for the non-uniform case, it can be interpreted as filtering ( $\xi_W$ ) followed by sampling ( $\text{Res}_{\rho(D)}$ ) and can intuitively be thought of as a means of observing  $M$  using "probes" with overlapping receptive fields. For the case of NULCAs, the band of frequencies filtered out by  $\xi_W$  is wide enough so that *any effect on behavior produced by structural non-uniformities are eliminated.* The resulting uniform system will give exact information about the remaining frequencies.

We now turn to the results summarized in Theorem 7.6 which permit homomorphic images of LCAs to be formed by means of spatial aggregation. We can also extend this simplification technique to the case of NULCAs. Consistent with Def. 7.5 we define an aggregated version of  $M^u$ :

Definition 8.5

For  $M^u = \langle S, K, H_U \rangle$ , the LCA  $M_A^u$  is given by  $\langle D, K, k\text{Agg}_{\eta} H_U \rangle$ .

Theorem 7.6 can be applied directly to the LCA  $M^u$  so that we can prove the following:

Theorem 8.4:

Let  $M = \langle S, K, \{B_\alpha\} \rangle$  be a NULCA and let  $W$  be any subset of  $U$ . Then  $\text{Agg}_\eta \xi_W$  is a homomorphism from  $M$  onto a submachine the LCA  $M_A^u$ .

Proof:

Theorem 8.2 says that

$$M \xrightarrow{\xi_W} M_U^u,$$

and Theorem 7.6 applies to  $M^u$  giving

$$M^u \xrightarrow{\text{Agg}_\eta} M_A^u.$$

Then the submachine of  $M_A^u$  with state space  $\text{Agg}_\eta \xi_W(F(S, K))$  is a homomorphic image of  $M_U^u$ . Call this submachine  $\text{Sub}M_A^u$ . Then we have the following commutative diagram:

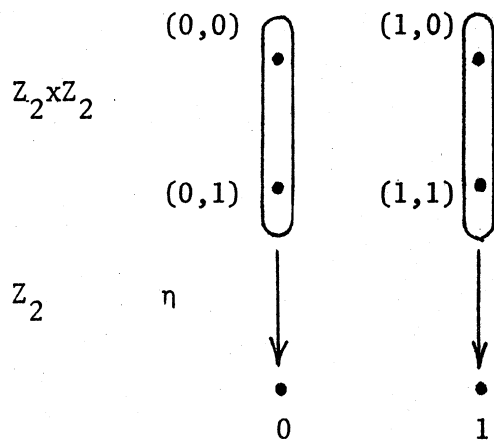
$$\begin{array}{ccc} M & \xrightarrow{\xi_W} & M_U^u \\ & \searrow & \downarrow \text{Agg}_\eta \\ & & \text{Sub}M_A^u \end{array}$$

Q.E.D.

We illustrate this theorem with a non-uniform analog of the situation presented by Ex. 7.4. There we showed that spatial aggregation could be used to reduce the dimension of an LCA's structure group. Here we show that the same technique can be used under less restrictive conditions.

Example 8.3

Consider the NULCA  $M = \langle \mathbb{Z}_2 \times \mathbb{Z}_2, \mathbb{R}, \{B_\alpha\} \rangle$  shown in Fig. 8.5. Since  $\mathbb{R}$  contains the principle 2<sup>nd</sup> root of unity  $-1$ ,  $M$  is transformable. In fact, the appropriate transform is an example of the finite Walsh transform whose matrix is shown in Ex. 5.4. Suppose  $\eta$  is as follows:



Then the kernel  $K$  of  $\eta$  is the subgroup  $\{(0,0), (0,1)\}$  of  $\mathbb{Z}_2 \times \mathbb{Z}_2$ .

We've specified the functions  $B_\alpha$ ,  $\alpha \in \mathbb{Z}_2 \times \mathbb{Z}_2$ , such that the following holds: for any coset  $\Lambda$  of  $K$ , each cell in  $\Lambda$  is such that the sum of its outputs which go to cells also in  $\Lambda$  is a constant  $x \in \mathbb{R}$ . The connections between cells in different cosets are uniform. Thus, the set  $U$  corresponding to this NULCA is  $\{(0,0), (1,0)\}$  which is compatible with the group-homomorphism  $\eta$ . We can therefore apply Theorem 8.4 for  $W = U$ . Fig. 8.5 shows  $M$ , its uniformly interconnected homomorphic image  $M^u$ , and the aggregated version  $M_A^u$ .

We'd like to make explicit several observations about the simplification process illustrated in Fig. 8.5. First, notice that in going from  $M^u$  to  $M_A^u$  we've aggregated over the cosets of  $\eta$ . However, this kind of aggregation can also be viewed as smoothing followed by

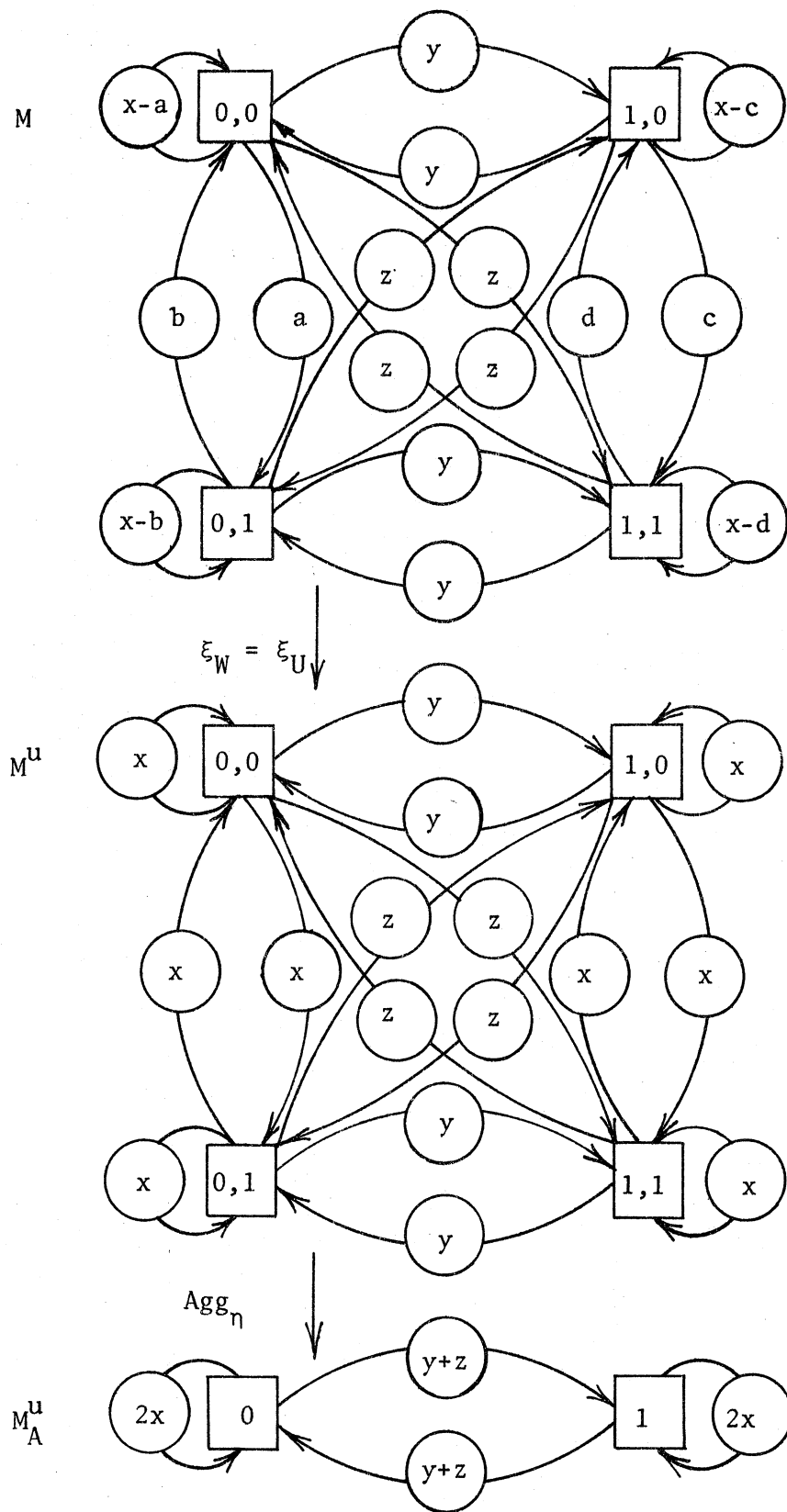


Figure 8.5: The NULCA  $M$  can be aggregated to form the LCA  $M_A^u$ .

sampling. If  $\hat{\chi}_U$  is the characteristic function of  $U$ , then its inverse transform  $\chi_U$  is equal to 0 on all of  $Z_2 \times Z_2$  except for values of 1/2 at  $(0,0)$  and  $(0,1)$ . Convoluting (w.r.t.  $Z_2 \times Z_2$ ) with  $\chi_U$  and then taking one sample from each block induced by  $\eta$  produces the same result as the map  $\text{Agg}_\eta$ . A second observation is that although the simplification process of Fig. 8.5 can easily be justified without any mention of the DFT, the case shown involving equal sums of influences within the blocks is a special case of our more general method. That the response functions sum to the same value within a block is the same as saying that their  $e^{\text{th}}$  Fourier coefficients are the same where  $e$  is the identity of the underlying structure group. As we mentioned earlier, this is the case for any group since the character  $\phi_e$  is always a constant function. Here, however, we've shown the notion of sum in this context can be generalized to a subset of spatial frequency components.

Finally, we would like to point out that the LCA  $M_A^u$  of Fig. 8.5 is of the same form as the LCA discussed in Ex. 5.7 as a model of an aspect of morphogenesis. Following the analysis given for that example, we can conclude that when  $2x - (y+z) > 1$ , any initial inhomogeneity in the state of  $M_A^u$  will resonate, and thus cause the difference between the states of cell 0 and cell 1 to increase without bound. Therefore, we can conclude that the non-uniform system  $M$  behaves in the same way for  $2x - (y+z) > 1$ : the difference between the sums of cell states over the two blocks will increase without bound.

## 8.6 Discussion

The results described in this chapter go part way toward justifying the intuition that small structural non-uniformities do not affect large-scale spatial features of a system's behavior. If we postulate that such a system can be validly modeled by a NULCA  $M = \langle S, K, \{B_\alpha\} \rangle$  (the base model), then the presence of small non-uniformities might reasonably be represented by requiring the  $B_\alpha$ 's to differ only in their higher spatial frequency components. That is, the set  $U$  would consist of only the lower frequencies. Theorem 8.2 would then say that if one is interested in aspects of the system's behavior having relatively large spatial scale (i.e. involving low spatial frequencies), then it is not misleading to assume that the components of the system are identical and uniformly interconnected (the LCA  $M^u$ ).

The difficulty with this analysis, of course, is in postulating the NULCA  $M$  as a base model, and it turns out that this may involve some non-trivial assumptions about the structure of the "real" system. To make this point clear we will briefly discuss the dual situation in which NULCA's are characterized by weighting functions instead of impulse response functions. In Sec. 8.2 we pointed out that the reflection  $\bar{B}_\alpha$  of  $B_\alpha$ ,  $\alpha \in S$ , cannot be interpreted as the weighting function of component  $\alpha$ . Suppose, however, that we give the name  $C_\alpha$  to the weighting function of component  $\alpha$ . It seems reasonable that small spatial inhomogeneities can just as well be modeled by assuming that the  $C_\alpha$ 's are the same except for their higher spatial frequency components. In this case, however, the analog of Theorem 8.2 says that a homomorphic image can be formed by retaining the higher frequen-

cies and that this homomorphic image is not necessarily uniformly structured.\* An interesting question, therefore, is to ask under what conditions the smoothness of weighting functions (the  $B_\alpha$ 's) implies the smoothness of impulse response functions ( $C_\alpha$ 's).\*\* It seems plausible that very reasonable conditions can be given, but we have not pursued this question.

Another aspect of the results of this chapter arises if we ask how an LCA with time-varying impulse response functions might be modeled. If the temporal changes of the impulse response functions are confined to certain spatial frequencies, then it is possible to filter out these frequencies and obtain a uniformly interconnected and time-invariant homomorphic model. To see why this is true, note that at any time  $t$ , the time-varying LCA can be regarded as a NULCA  $M$  for which the set  $U$  is the set of spatial frequencies which do not change in time. At time  $t+1$ , the system is a different NULCA, but the set  $U$  is the same. It is not hard to convince oneself that the single LCA  $M^U$  is a homomorphic image of each of these different NULCAs. Therefore,  $M^U$  is a homomorphic image of the original time-varying system. This is true even when the spatial impulse response function of a cell varies with time independently of changes in the impulse response functions

---

\*More specifically, if in Fig. 8.2 the  $B_\alpha$ 's are replaced by  $C_\alpha$ 's, one obtains the dual of the digraph shown for  $\hat{M}$  in Fig. 8.2 b). That is, the arrows are reversed. Thus, components 0 and 1 do not influence other components but may be influenced by them.

\*\*This is not automatically true. If it were, components 0 and 1 of  $\hat{M}$  shown in Fig. 8.2 b) would only influence themselves.

of other cells. It may thus be possible to view certain kinds of nonlinearities of a cell as a time variance of its spatial impulse response function. In such cases, a cellular automaton composed of such cells can be modeled by an LCA. Further investigations are needed to determine whether non-trivial nonlinearities can be modeled by non-trivial LCAs.



## Chapter 9

### Conclusion

Throughout the preceding chapters we have tried to bring together parts of traditionally rather separate areas: the theory of cellular automata, techniques from digital signal processing, and rudiments of the theory of partial differential equations and their numerical approximations. At the same time, we have tried to maintain a self-contained and rigorous level of description without obscuring what we feel are the basic facts about the class of systems involved. We have therefore not attempted to ground the ideas in the most general framework and have found it necessary to touch only very lightly on substantial bodies of existing and, in some cases, well known theory. We would like to close with a discussion of some of the issues that are likely to arise in more general settings and some of the results that can be obtained.

Since the entire discussion has been restricted to the case of autonomous cellular automata, we first briefly describe what can be expected when provisions are made for input and output. If an LCA has input and output variables such that, as a whole, it is a non-autonomous LSM (Sec. 2.5), then one can extend the notion of structural uniformity so that it is still possible to form parallel decompositions. If the input and output connections are such that from an I/O point of view, the LCA can be characterized as a linear operator with automorphism group  $S \times T$  (where  $S$  is its structure group and  $T$  is time), then the LCA is isomorphic to a set of completely independent LSMs each having input and output. Since the input, output, and state spaces

of these LSMs are the same as the corresponding spaces of a cell of the LCA, the results regarding minimality, controllability, and observability of the entire cellular automaton reduce to the corresponding questions for LSMs over these smaller spaces.

For example, if  $M = \langle S, K, H \rangle$  is a decomposable LCA which also has input space  $F(S, K)$  and output space  $F(S, K)$ , then given the appropriate uniformity of input/output connections,  $M$  is controllable if and only if it is 1-controllable (see Harrison [1969]). This means that if such a non-autonomous LCA is strongly connected, then by choosing the correct input configuration, any state can be taken to any other state in one time step. If, on the other hand,  $M$  has cell state set  $K^2$  and input and output spaces  $F(S, K^2)$ , then it takes at most two time steps to go from any state to any other.

We indicated in Sec. 3.6 and Sec. 5.6 that the framework used in this presentation can be generalized to include structures with uncountable structure groups. This direction of generalization will permit cellular automata and partial differential equations to be considered particular members of the same class of models. Using concepts from a systems theoretic view of the modeling and simulation process, it will be possible to explore relationships between models having a countable number of dimensions and those having an uncountable number. It is felt that this approach can lead to questions and results which have a character different than those motivated by the desire to solve uncountably dimensional problems by means of finite dimensional systems. In particular, it is hoped that such studies will increase our repertoire of formalisms for modeling both natural systems and computational processes.

## REFERENCES

1. Aho, A. V.; Hopcroft, J. E.; Ullman, J. D., The Design and Analysis of Computer Algorithms, Addison-Wesley, Reading, Mass., 1974.
2. Aladyev, V., " $\tau$ -grammars and their Languages", Izu. AN ESSR, Biol. 22, No. 4<sup>n</sup>, 1973.
3. \_\_\_\_\_, "Survey of Research in the Theory of Homogeneous Structures and their Applications", Mathematical Biosciences, 22, 1974.
4. Amoroso, S.; Cooper, G., "Tessellation Structures for Reproduction of Arbitrary Patterns", Journal of Computer and System Sciences, 5, 5, 1971.
5. Andrews, H. C., Computer Techniques in Image Processing, Academic Press, New York, 1970.
6. Ashby, W. R., An Introduction to Cybernetics, John Wiley & Sons, New York, 1963.
7. Beckmann, P., Orthogonal Polynomials for Engineers and Physicists, The Golem Press, Boulder, Colorado, 1973.
8. Bergland, G.D., "A Guided Tour of the Fast Fourier Transform", IEEE Spectrum, July 1969.
9. Boerner, H., Representation of Groups, North-Holland, Amsterdam, 1972.
10. Bracewell, R., The Fourier Transform and its Applications, McGraw-Hill, 1965.
11. Budden, F.J., The Fascination of Groups, Cambridge University Press, 1972.
12. Burks, A. W. (Ed.), Essays on Cellular Automata, University of Illinois Press, 1970.
13. \_\_\_\_\_, "Models of Deterministic Systems", Math. Sys. Theor., 8, 4, 1973.
14. Cairns, T. W., "On the Fast Fourier Transform on Finite Abelian Groups", IEEE Transactions on Computers, May 1971.
15. Chen, C. T., Introduction to Linear Systems Theory, Holt, Rinehart, and Winston, 1970.
16. Codd, E. F., Cellular Automata, Academic Press, New York, 1968.

17. Cooley, J. W.; Lewis, P. A. W.; Welch, P. D., "Application of the Fast Fourier Transform to Computation of Fourier Integrals, Fourier series, and Convolution Integrals", IEEE Trans. Audio Electroacoustics, Au-15, 2, 1967.
18. \_\_\_\_\_, "The Finite Fourier Transform," IEEE Trans. on Audio and Electroacoustics, Au-17, 2, 1969.
19. Curtis, C. W.; Reiner, I., Representation Theory of Finite Groups and Associative Algebras, Interscience Publishers, 1962.
20. de Boor, C., "The Method of Projections as Applied to the Numerical Solution of Two Point Boundary Value Problems Using Cubic Splines," Ph.D. Thesis, University of Michigan, 1966.
21. Duff, G.F.D.; Naylor, D., Differential Equations of Applied Mathematics, John Wiley & Sons, New York, 1966.
22. Fleck, A.C., "Isomorphism Groups of Automata," J. Assoc. Comput. Mach., 9, 1962
23. Foy, J., "A Computer Simulation of Impulse Conduction in Cardiac Muscle," Ph.D. Thesis, Computer & Communication Sciences, The University of Michigan, 1974.
24. Gardner, M., "The Fantastic Combinations of John Conway's New Solitaire Game 'Life'," Scientific American, Oct. 1970.
25. Gary, J., "The Numerical Solution of Partial Differential Equations," Manuscript, 1972.
26. Gold, B.; Rader, C. M., Digital Processing of Signals, McGraw-Hill, New York, 1969.
27. Goodman, E.; Weinberg, R.; Laing, R., "A Cell Space Embedding of Simulated Living Cells," Bio-Medical Computing, 2, 1971.
28. Grossman, I.; Magnus, W., Groups and their Graphs, Random House, 1964.
29. Harmuth, H. F., Transmission of Information by Orthogonal Functions, Springer-Verlag, New York, 1972.
30. Harrison, M. A., Lectures on Linear Sequential Machines, Academic Press, 1969.
31. Hartmanis, J.; Stearns, R. E., Algebraic Structure Theory of Sequential Machines, Prentice-Hall, Englewood Cliffs, N.J., 1966.
32. Hewitt, E.; Ross, K. A., Abstract Harmonic Analysis, Springer-Verlag, Berlin, 1963.

33. Holland, J., "A Universal Computer Capable of Executing an Arbitrary Number of Sub-Programs Simultaneously," Proceedings of the Eastern Joint Computer Conference, 1959.
34. \_\_\_\_\_, "Iterative Circuit Computers," Proc. of the 1960 Western Joint Computer Conf., 1960.
35. IEEE Trans. Audio Electroacoustics, Special issue on Fast Fourier Transform and its application to digital filtering and spectral analysis, Au-15, 2, 1967.
36. \_\_\_\_\_, Special issue on Fast Fourier Transform, Au-17, 2, 1969.
37. Jump, J. R., "A Note on the Iterative Decomposition of Finite Automata," Information and Control, 15, 1969.
38. Kalman, R. E.; Falb, P. L.; Arbib, M. A., Topics in Mathematical Systems Theory, McGraw-Hill, 1969.
39. Knuth, D. E., Seminumerical Algorithms, Addison-Wesley, Reading, Mass., 1969.
40. Lang, S., Algebraic Structures, Addison-Wesley, Reading, Mass. 1967.
41. Minnick, R. C., "A Survey of Microcellular Research," J. of the Assoc. for Comput. Mach., 14, 2, 1967.
42. Mortimer, J. A., "A Cellular Model for Mammalian Cerebellar Cortex," University of Michigan Technical Report, LCG-107, 1970.
43. Nicholson, P. J., "Algebraic Theory of Finite Fourier Transforms," J. of Computer and System Sciences, 5, 5, 1971.
44. Ostrand, T. J., "Pattern Reproduction in Tessellation Automata of Arbitrary Dimension," J. of Computer and System Science, 5, 6, 1971.
45. Roach, G. F., Green's Functions: Introductory Theory with Applications, Van Nostrand Reinhold Company, London, 1970.
46. Rosen, R., Dynamical System Theory in Biology, VI, Wiley-Interscience, 1970.
47. \_\_\_\_\_, "Biological Systems as Organizational Paradigms," Int. J. of General Systems, 1, 3, 1974.
48. Rotman, J. J., The Theory of Groups, Allyn and Bacon, Boston, 1965.
49. Rudin, W., Fourier Analysis on Groups, Wiley-Interscience, New York, 1962.

50. Schechter, M., Principles of Functional Analysis, Academic Press, New York, 1971.
51. Schwarz, R. J.; Friedland, B., Linear Systems, McGraw-Hill, New York, 1965.
52. Stroke, G. W., An Introduction to Coherent Optics and Holography, Academic Press, New York, 1969.
53. Thatcher, J., "Universality in the von Neumann Cellular Model," University of Michigan Technical Report, LCG-42, Sept. 1964.
54. von Neumann, J.; Burks, A. W. (Ed.), Theory of Self-Reproducing Automata, University of Illinois Press, Urbana, 1966.
55. Wagner, E.G., "On Connecting Modules Together Uniformly to Form a Modular Computer," IEEE Trans. on Electronic Computers, EC-15, 6, 1966.
56. Waltz, F., "Some Properties of Circulents," Univeristy of Michigan, Dept. of Electrical Engineering, SEL Technical Report No. 4, 1966.
57. Yamada, H.; Amoroso, S., "Tessellation Automata," Information and Control, 14, 3, 1969.
58. \_\_\_\_\_, "Structural and Behavioral Equivalences of Tessellation Automata," Information and Control, 18, 1, 1971.
59. Zeigler, B. P., "Toward a Formal Theory of Modelling and Simulation, I," or "On the Formulation of Problems in Modeling and Simulation within Mathematical Systems Theory," Proceedings of the Sixth International Congress on Cybernetics, Namur, Belgium, 1970.
60. \_\_\_\_\_, "Toward a Formal Theory of Modelling and Simulation, II," University of Michigan Technical Report, LCG-120, July 1971.
61. \_\_\_\_\_, "The Base Model Concept," Theory and Application of Variable Structure Systems, Mohler, R. R., Ruberti, A. (Eds.), Academic Press, New York, 1972.
62. \_\_\_\_\_, "A Conceptual Basis for Modelling and Simulation," Int. J. of General Systems, 1, 4, 1974.

