BINARY SEQUENCES WITH PRESCRIBED AUTOCORRELATIONS

Jorge Leite Martins de Carvalho

A thesis submitted for the

degree of Doctor of Philosophy

October 1977

Department of Computing and Control

Imperial College of Science and Technology

University of London

ABSTRACT

This thesis is concerned with the characterization of the class of autocorrelation functions of binary sequences, its connections with the class of covariance functions of binary discrete time stochastic processes, known as unit processes, and the design of algorithms for the generation of binary sequences with prescribed autocorrelations.

A geometric approach to the characterization of the class of autocorrelation functions of binary sequences is adopted. A detailed study of the properties of this class of functions is carried out, leading to a better understanding of its structure. Necessary and sufficient conditions are derived for a function to be an element of this class. Motivated by practical applications, special attention is given to the characterization of its finite dimensional projections. These turn out to be bounded convex polyhedra and their vertices have a number of important properties.

A new characterization of the class of covariance functions of discrete time unit processes is given and its relations with the class of autocorrelation functions of binary sequences are discussed.

Two classes of algorithms are proposed with proven global convergence and established rates of convergence

for the generation of binary sequences with any prescribed number of autocorrelation shifts.

Some computational results are also presented.

TO MANUELA

ACKNOWLEDGEMENTS

My sincerest thanks are owed to my supervisor Dr. J. M. C. Clark, who inspired this research and who gave me valuable guidance and encouragement.

I would also like to thank the staff and students of the Control Section for useful discussions, particularly Dr. M. H. Davis.

Financial support was received from the Portuguese Government Scholarship Authority.

Finally I wish to express my thanks to M. J. de Oliveira Baptista for her excellent typing.

CONTENTS

NOTATION

A Class of positive semi-definite sequences $(\rho_n)_{n \in \mathbb{Z}}$ with $\rho_o=1$

C Class of autocorrelation functions of binary sequences, that is $(\rho_n)_{n \in \mathbb{Z}} \in C$ if

$$\rho_n = \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N} s_k s_{k+n} \text{ , with } s_k \in \{-1,1\} \text{ for}$$

$k \geqslant 0$ and $s_k=0$ for $k<0$. Given that $\rho_n=\rho_{-n}$ for all n, we shall only consider the autocorrelation function defined for non-negative arguments, that is $(\rho_n)_{n \in \mathbb{N}_0}$.

Sometimes it will be convenient to establish a one-to-one correspondence between this class and a set of semi-infinite symmetric Toeplitz matrices as follows

$$\begin{bmatrix} \rho_o & \rho_1 & \rho_2 & \cdots & \rho_n & \cdots \\ \rho_1 & \rho_o & \rho_1 & \cdots & \rho_{n-1} & \cdots \\ \rho_2 & \rho_1 & \rho_o & \cdots & \rho_{n-2} & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \\ \rho_n & \rho_{n-1} & \rho_{n-2} & \cdots & \rho_o & \cdots \\ \vdots & \vdots & \vdots & & \vdots & \end{bmatrix}$$

C_n $\equiv {}^{\Pi}{}_n C$

$D_n^{\pm}$ Definition on section 4.3.1.

H Denotes convex hull, that is a class of linear combinations where the coefficients are all non-negative and add up to 1.

M^p Class of p-periodic matrices, that is matrices whose elements a_{ij} satisfy the relation $a_{ij} = a_{i+np,j+tp}$, $n,t \in \mathbb{Z}^+$.

$\mathbb{N}$ Set of positive integers.

$\mathbb{N}_0$ Set of positive integers and zero.

Π_n Denotes the projection of a matrix into the $(n \times n)$ matrix in the upper left corner.

P Subclass of autocorrelation functions of periodic binary sequences.

T Class of semi-infinite Toeplitz matrices.

Example: TU (that is $T \cap U$) is the class of stationary discrete-time unit covariances.

U Class of discrete time unit covariances that is $(\rho_{ij})_{i,j \in \mathbb{N}_0} \in U$ if $\rho_{ij} = EX_iX_j$ where $X_t(w) \in \{-1,1\}$ for all $w \in \Omega$ and for all $t \in \mathbb{N}_0$, and E denotes the expectation operator.

As above, the same can be said about mapping

these functions in a one-to-one fashion on to a set of semi-infinite symmetric matrices.

U_n $\equiv \Pi_n U$

$W_n^{\pm}$ Definition in section 4.3.1.

W_n^{i} Definition in section 4.3.1.

$\mathbb{Z}$ Set of positive and negative integers and zero.

CHAPTER 1

INTRODUCTION

In many engineering applications the need frequently arises for a signal with a prescribed autocorrelation function. In the identification of systems in the presence of disturbances it can be shown that the accuracy of the determination of the system characteristics is determined, to some extent, by the choice of the input signal. If the system under consideration is linear, time-invariant, stable, single input-single output, then the 'optimal' input is most conveniently characterized, for parametric identification purposes, in the time domain by its autocorrelation matrix.

The following simple example illustrates this situation. Consider a system described by the moving-average model

$$(1) \quad y(t) = a_o u_t + a_1 u_{t-1} + \ldots + a_{p-1} u_{t-p+1} + w_t \; ; \quad t=0, 1, \ldots, N$$

where w_t is gaussian white noise, and the coefficients $a_o, \ldots, a_{p-1}$ are to be estimated from N observations of the input and output.

Equation (1) can be written in matrix notation as

$$(2) \qquad \underline{Y} = X\underline{a} + \underline{W}$$

where $\underline{Y}$ is an N×1 vector of observations, X is an N×p matrix, and $\underline{a}$ is the vector of unknown coefficients $(a_0, a_1, \ldots, a_{p-1})$. The covariance matrix of the least squares estimator $\hat{\underline{a}}$ of $\underline{a}$ is

$$V(\hat{\underline{a}}) = \sigma^2 \left[X^T X\right]^{-1} ,$$

and if the input $(u_t)_{t \in \mathbb{N}_0}$ possesses an autocorrelation function $(\rho_n^u)_{n \in \mathbb{N}_0}$ with

$$\rho_n^u \triangleq \lim_{N \to \infty} \frac{1}{N} \sum_{t=0}^{N} u_t u_{t+n} ,$$

then $$\lim_{N \to \infty} NV(\hat{\underline{a}}) = \sigma^2 \begin{bmatrix} \rho_0^u & \rho_1^u & \cdots & \rho_{p-1}^u \\ & \ddots & & \vdots \\ & & \ddots & \rho_1^u \\ & & & \rho_0^u \end{bmatrix}^{-1}$$

where $(\rho_n^u)_{n \in \mathbb{N}_0}$ denotes the autocorrelation function of the input $(u_t)_{t \in \mathbb{N}_0}$.

We would like to draw attention at this point to the fact that we shall be using the words 'autocorrelation'

and 'covariance' to refer to different concepts. Their definitions can be found in the section dealing with notation. Sometimes in the literature the word auto-correlation means a normalized covariance, but this is by no means standard.

The covariance of the estimator obviously tends to 0 as $N \to \infty$. One of the objectives of experimental design is to increase the rate of convergence by manipulating the input of the system under test. The most widely used scalar measures of this 'rate of convergence' are

(3) Trace $\left[N\, W_W V(\underline{\hat{a}})\right]$, where W_W is some weighting matrix and

(4) $-\log\det\left[N\, V(\underline{\hat{a}})\right]$

Then the selection of the 'optimal' input becomes an optimization problem on the set of admissible inputs. However in our example it is clear that it is more convenient to carry out the optimization with respect to the input autocorrelation, because:

(a) (3) and (4) are convex functions of the input auto-correlation and therefore we can get a global optimum.

(b) The optimization is carried out in a space of much lower dimension, that is, $p \ll N$.

(c) If we optimize with respect to the input, we are

likely to get only locally optimizing inputs due to the non-convexity of the functions (3) and (4) with respect to $(u_t)_{t \in \mathbb{N}_0}$.

This has been suggested by several authors, namely G. C. Goodwin and R. L. Payne [3] and R. Mehra [6].

In the identification example given above we see that the information matrix depends only on the first p autocorrelation values of the input. This suggests that in general the required number of autocorrelation shifts might be connected with the 'memory' of the system. In fact R. L. Payne [7] has shown that the information matrix can be approximated to any desired accuracy by the first few autocorrelation shifts of the input. This is very important, because the selection of the optimal input autocorrelation then becomes an optimization problem on a subset of $\mathbb{R}^n$ - the set of the first n feasible autocorrelation shifts of the class of admissible inputs - where n is determined by the desired accuracy. Furthermore we only need to look at partial realizations of autocorrelation functions, that is, given a finite set of autocorrelation values construct a signal whose autocorrelation matches these values.

Very often the input signal is constrained to be binary, because of the relatively simple hardware

required for the generation, storing and processing of these signals. However the characterization of the class of autocorrelation functions of binary signals and methods for generating binary signals with prescribed autocorrelations was a previously unsolved problem either in continuous or discrete time.

This thesis will be concerned with the following problems: (i) The characterization of the class of autocorrelation functions of discrete time binary signals and in particular the class of their finite dimensional projections. (ii) The design of convergent algorithms for the generation of binary sequences with prescribed autocorrelations.

Our first attempt to solve these questions was algebraic, using the theory of linear recurring sequences over the finite field with two elements, GF(2). This fits quite naturally in the framework of control theory in the sense that it involves an inverse problem. In the theory of linear recurrent sequences over GF(2) we are given the coefficients of a linear recurrence and seek to compute the solutions. In our case a certain property of the solutions is prescribed (their autocorrelations) and we wish to determine the corresponding coefficients.

Support to an algebraic approach also seemed to

be given by the pioneering works of N. Zierler [11] and S. W. Golomb [2] on linear recurrent sequences, which led to the important discovery of the so called 'pseudo random binary noise'; that is, periodic binary sequences of period p with a constant out-of-phase autocorrelation equal to -1/p. Such a property is particularly useful in many engineering applications such as radar ranging and telecommunications. This type of signal is also particularly convenient to use if the impulse response of a linear system is to be determined using correlation methods, because its impulsive autocorrelation avoids the deconvolution of the Wiener-Hopf equation. However for parametric identification purposes a signal with an impulsive autocorrelation is not necessarily optimum, because the 'optimal' signal obviously depends upon the optimality criterion that is adopted.

The algebraic treatment was abandoned because of the difficulty in solving the inverse problem; that is to find the order and the coefficients of the linear recurrence over GF(2) having at least one solution with the prescribed autocorrelation. In the absence of a general solution what is usually done is to consider some sub-class of linear recurrences over GF(2) and then to study the autocorrelation properties

of the generated sequences. See for example [5] where the autocorrelation properties of binary sequences associated with non-primitive irreducible polynomials over GF(2) are studied. However such type of approach would only provide a partial answer to our problem.

Some early work of L. Shepp [8] on the class of autocorrelation functions of binary sequences helped us to gain insight into the structure of this class and strongly suggested that a geometric approach to this problem would be fruitful. This has proved to be right.

A characterization of the class of autocorrelation functions of binary sequences in terms of their finite dimensional projections has been obtained. It is shown that if $(\rho_n)_{n \in \mathbb{N}_0}$ is the autocorrelation of a binary sequence then, for all $m \in \mathbb{N}$, the vector $(\rho_0, \rho_1, \ldots, \rho_{m-1})$ lies in a polytope, that is a bounded convex polyhedron, whose vertices are autocorrelations of periodic binary sequences of periods not greater than 2^{m-1}. We denote this polytope by $\Pi_m C$. Conversely if $(a_0, a_1, \ldots, a_{m-1})$ belongs to this polytope then it can be extrapolated to an autocorrelation function $(\rho_n)_{n \in \mathbb{N}_0}$ of some binary sequence; in other words there exists a binary sequence $(s_n)_{n \in \mathbb{N}_0}$ such that

$$\rho_n = \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N} s_i s_{i+n}$$

and $a_j=\rho_j$, $0\leqslant j\leqslant m-1$. Furthermore it is also shown that $(\rho_n)_{n\in\mathbb{N}_0}$ can be chosen to be <u>periodic</u>. This is a rather surprising result, in contrast with the case of general positive semi-definite sequences, where it does not necessarily hold. Algorithms for the calculation of the vertices $\Pi_m C$ and associated periodic binary sequences are also given and have been used to calculate $\Pi_m C$, for some values of m. The calculation of the vertices of $\Pi_m C$ gave rise to some interesting mathematical questions, namely the problem of selecting the vertices of the convex hull of a finite number of points; this is treated in some detail.

A closely related class of functions, the class of covariance functions of discrete time binary processes, in short unit processes, is also characterized and its connections with the class of autocorrelation functions of binary sequences are discussed.

As far as the author is aware, the question of convergent procedures for the partial realization of autocorrelation functions of binary sequences has not been treated except in this thesis. A number of procedures have been suggested in the literature for the generation of binary signals with given auto-correlation or spectral properties, but no convergence proofs are available for them. See for example [1], [4] and [9].

Two conceptual algorithms with proved convergence to generate binary sequences with any prescribed number of autocorrelation shifts have been established. Experimental evidence suggests that the conditions imposed in the proposed conceptual algorithms can be relaxed in a way that leads to more easily implementable procedures that still seem to converge. This is the case of an algorithm we propose which is a simplification of one of our conceptual algorithms. It has worked very satisfactorily in a large number of examples but so far its converge has only been shown for the case where the prescribed number of autocorrelation values is not larger than three.

In addition to the application in systems identification that was described earlier we believe that our work will be of use in other areas. Our characterization of the class C in terms of its finite dimensional projections seems particularly suitable in the frequency analysis of binary signals when the power density spectrum has to be calculated from the autocorrelation function. As one in practice has always a finite measuring time this means that one has to face the truncation effects of the correlation function on the power density spectrum; see [10] for example. One way to overcome this problem is to extrapolate the auto-

correlation function. The extrapolation is usually done on the basis that the extrapolated values must give rise to a positive semi-definite function. However if we are dealing with binary signals, this condition is not sufficient: they must also lie inside $\Pi_m C$, for all m, which means that the knowledge of the polytopes $\Pi_m C$ is required in this case.

References

[1] A. van den Bos, Construction of Binary Multi-frequency Test Signals, IFAC Symposium on Identification, Prague, Czechoslovakia, 1967.

[2] S. W. Golomb and Others, Digital Communications with Space Applications, Prentice-Hall.

[3] G. Goodwin and R. L. Payne, Design and Characterization of Optimal Test Signals for Linear Single Input-Single Output System Parameter Estimation, Publication 73/1, Dept. of Computing and Control, Imperial College, London SW7.

[4] P. V. Indiresan and G. K. Uttaradhi, Iterative Method for Obtaining Good Aperiodic Binary Sequences, Journal of Optimization Theory and Applications, Vol. 7, No. 2, 1971.

[5] J. J. Lee and D. R. Smith, Families of Shift Register Sequences with Impulsive Correlation Properties, IEEE-Transactions on Information Theory, Vol. I T-20, March 1974.

[6] R. K. Mehra, Optimal Input Signals for Parameter Estimation in Dynamic Systems - Survey and New Results, IEEE-Transactions on Automatic Control, Vol. AC-19, No. 6, December 1974.

[7] R. L. Payne, Optimal Experiment Design for Dynamic Systems Identification, Ph.D. Dissertation, Imperial College, London SW7, 1974.

[8] L. Shepp, Private Communications.

[9] U. Somaini, Binary Sequences with Good Autocorrelation and Cross-correlation Properties, IEEE-Transactions on Aerospace and Electronic Systems, Vol. AES-11, No. 6, November 1975.

[10] B. Veltman, A. van den Bos and Others, Some Remarks on the Use of Autocorrelation Functions in the Analysis and Design of Signals, Proceedings, Nato Advanced Study Institute on Signal Processing, Loughborough, 1973.

[11] N. Zierler, Linear Recurring Sequences, J. Soc. Indust. Appl. Math., Vol. 7, No. 1, March 1959.

CHAPTER 2

THE CLASS OF AUTOCORRELATION FUNCTIONS OF BINARY SEQUENCES

2.1 Introduction

As described in the previous chapter, one is often faced in identification experiments with the problem of having to generate a binary sequence whose first m autocorrelation values match some $\underline{v}$, $\underline{v} \in \mathbb{R}^m$, where $\underline{v}$ is in general the result of some optimization problem.

There are other cases where one is given the autocorrelation vector $\underline{v} \in \mathbb{R}^m$ and wishes to extrapolate these to some further autocorrelation values or even an entire autocorrelation function.

In any case the knowledge of the sets $\Pi_m C$, $m \in \mathbb{N}$, is required. In this chapter we shall be concerned with the characterization of the sets $\Pi_m C$, and also the characterization of the class C.

As a result of this study the structure of the class C has been clarified.

To the author's belief the two most interesting results obtained are:

(i) $\Pi_m C$ is a polytope whose vertices are projections of

autocorrelation functions of periodic binary sequences of periods not greater than 2^{m-1}.

This property provides not only a feasible way for generating the vertices of $\Pi_m C$ (the number of periodic binary sequences with a given period is finite) but it will also enable to prove convergence for two of the algorithms described in chapter 4 for the partial realization of autocorrelation functions.

(ii) Given $\underline{v} \in \Pi_m C$ there always exists a periodic element of C, $(\rho_n)_{n \in \mathbb{N}_0}$, such that $\Pi_m \rho = \underline{v}$. This is a rather interesting result in contrast with the general case of positive semi-definite sequences for which it can be shown (see section 2.4) that points exist in $\Pi_m A$ which can not be extrapolated to a periodic element of A.

Attention is also paid to the calculation of the vertices of $\Pi_m C$. $\Pi_m C$ is a polytope and can therefore be described in terms of its vertices. Furthermore we shall see in chapter 4 that the periodic binary sequences associated with the vertices of $\Pi_m C$ can be used in the generation of binary sequences with autocorrelations matching $\Pi_m \rho$, for any given ρ in C. The selection of the vertices of $\Pi_m C$ gave rise to some interesting mathematical questions, namely the problem of selecting the vertices of the convex hull of a finite number of points. Some methods to solve this problem are proposed and some computational results are also presented.

2.2 Preliminaries

Definition: Given a binary sequence $(s_n)_{n \in \mathbb{N}_0}$, $s_n \in \{-1,1\}$ for all n, we define its autocorrelation $(\rho_j)_{j \in \mathbb{N}_0}$ by

$$\rho_j = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N} s_n s_{n+j} ,$$

if such a limit exists. It follows immediately that $-1 \leqslant \rho_j \leqslant 1$, for all j.

The autocorrelation can also be defined for negative shifts provided we set $s_n = 0$ for $n < 0$. However we shall only consider the autocorrelation function defined for non-negative arguments because $\rho_j = \rho_{-j}$, for all $j \in \mathbb{N}_0$.

Let C denote the class of autocorrelation functions of binary sequences. Consider $([-1,1], T_{\mathbb{R}})$ as a topological space; that is the interval $[-1,1]$ with the topology inherited from the usual topology of $\mathbb{R}$. Form the product space $([-1,1]^{\mathbb{N}}, T_{\mathbb{N}})$ where $T_{\mathbb{N}}$ is the product topology. Any element ρ in C is such that $\rho_o = 1$. Therefore C can be regarded as a subspace of $([-1,1]^{\mathbb{N}}, T_{\mathbb{N}})$ with the topology inherited from $T_{\mathbb{N}}$. All the topological notions of C will be with respect to this topology, unless otherwise specified. Now it is clear that convergence of a sequence of elements of C in the topology just defined is the same as pointwise convergence.

2.3 Properties of the Class C

The properties studied below will be used in our characterization of $\Pi_m C$ and also in establishing the bridge between the class C and the class of stationary discrete time unit covariances, that is TU.

The results stated in lemmas 1 and 3 have already been established by L. Shepp [4]. In view of their importance in the subsequent results the full proofs are given here.

Lemma 1

The class C is compact.

Proof

First it is shown that C is closed.

Suppose then that the family of autocorrelation functions, indexed by k, $(\rho_n^k)_{n \in \mathbb{N}_0}$ converges pointwise to a sequence $(\rho_n)_{n \in \mathbb{N}_0}$.

Let $(s_n^k)_{n \in \mathbb{N}_0}$ denote a binary sequence with autocorrelation $(\rho_n^k)_{n \in \mathbb{N}_0}$. By definition we have that for all $\varepsilon > 0$ and any positive integer p there exists a number $r_k(\varepsilon, p)$ such that

$$\left| \rho_n^k - \frac{1}{r} \sum_{m=0}^{r} s_m^k s_{m+n}^k \right| < \varepsilon \ , \ |n| \leqslant p \ , \ r > r_k(\varepsilon, p)$$

Now select a sequence of positive numbers $(\varepsilon_k)_{k \in \mathbb{N}_0}$

converging to 0 and define inductively another sequence $(n_k)_{k \in \mathbb{N}_0}$ of positive integers converging to ∞ satisfying:

(i) $(r_k(\varepsilon_k, k)/n_k)_{k \in \mathbb{N}} \longrightarrow 0$

(ii) $(n_{k-1}/n_k)_{k \in \mathbb{N}} \longrightarrow 0$

Construct a new binary sequence $(s_j)_{j \in \mathbb{N}_0}$ as follows:

$$s_j = \begin{cases} \text{arbitrary, for } j \leqslant n_1 \\ s^k_{j-n_k}, \text{ for } n_k < j \leqslant n_{k+1} \end{cases}$$

Now it is shown that this sequence has auto-correlation $(\rho_n)_{n \in \mathbb{N}_0}$.

Fix $n_0 \in \mathbb{N}_0$. From (i) and (ii) above and for N such that $n_k < N \leqslant n_{k+1}$ we have

$$\frac{1}{N} \sum_{m=0}^{N} s_m s_{m+n_0} = \frac{1}{N} \sum_{m=0}^{n_k - n_{k-1}} s^{k-1}_m s^{k-1}_{m+n_0} + \frac{1}{N} \sum_{m=0}^{N-n_k} s^k_m s^k_{m+n_0} + R(k)$$

where $R(k) \longrightarrow 0$ as k goes to ∞. In fact $|R(k)|$ is bounded by $n_{k-1}/N + 2kn_0/N$.

As k goes to ∞ we must still distinguish two cases:

a) $n_k < N \leqslant n_k + r_k(\varepsilon_k, k)$

$$\frac{1}{N} \sum_{m=0}^{N} s_m s_{m+n_0} = \rho_{n_0}^{k-1} + R'(k) \quad , \text{ where } \lim_{k \to \infty} R'(k) = 0.$$

Therefore

$$\lim_{N \to \infty} \frac{1}{N} \sum_{m=0}^{N} s_m s_{m+n_0} = \rho_{n_0}$$

b) $n_k + r_k(\varepsilon_k, k) < N \leqslant n_{k+1}$

$$\frac{1}{N} \sum_{m=0}^{N} s_m s_{m+n_0} = \frac{n_k - n_{k-1}}{N} \rho_{n_0}^{k-1} + \frac{N - n_k}{N} \rho_{n_0}^{k} + R''(k) \quad ,$$

where $\lim_{k \to \infty} R''(k) = 0$. Also $\lim_{k \to \infty} \frac{N - n_{k-1}}{N} \longrightarrow 1$. Therefore

$$\lim_{N \to \infty} \frac{1}{N} \sum_{m=0}^{N} s_m s_{m+n_0} = \rho_{n_0} \quad .$$

This shows that C is closed. C is also compact, because it is a closed subset of a compact space, namely $[-1,1]^{\mathbb{N}}$.

This ends the proof of lemma 1.

Let $(\rho_n)_{n \in \mathbb{N}_0}$ be the autocorrelation function of the binary sequence $(s_n)_{n \in \mathbb{N}_0}$. We show next that for any given $\delta > 0$ and $n_0 \in \mathbb{N}_0$ there exists a periodic binary sequence $(u_n)_{u \in \mathbb{N}_0}$ with an autocorrelation $(\lambda_n)_{n \in \mathbb{N}_0}$ and such that

$$|\rho_n - \lambda_n| < \delta \quad , \quad |n| < n_0 .$$

Lemma 2

P is dense in C.

Proof

First we show that any point in C is an accumulation point of P.

Let $(\rho_n)_{n \in \mathbb{N}_0}$ be the autocorrelation of the sequence $(s_n)_{n \in \mathbb{N}_0}$. We have that given $\delta > 0$ and $n_0 \in \mathbb{N}$, there exists $\overline{N}(\delta, n_0)$ such that

$$(1) \qquad \left|\rho_n - \frac{1}{N} \sum_{k=0}^{N} s_k s_{k+n}\right| < \delta/4 \quad , \quad |n| \leqslant n_0 \text{ and } N \geqslant \overline{N}.$$

We now define a periodic binary sequence $(u_n)_{n \in \mathbb{N}_0}$ of period p, where $p = \max(\overline{N}, n_0\ 4/\delta)$, as follows:

$$u_i = s_i \qquad i = 0, 1, 2, \ldots, p-1.$$

$$(2) \qquad \left|\rho_n - \frac{1}{p} \sum_{k=0}^{p-1} u_k u_{k+n}\right| = \left|\rho_n - \frac{1}{p} \sum_{k=0}^{p-n-1} s_k s_{k+n} - \frac{1}{p} \sum_{k=p-n}^{p-1} u_k u_{k+n}\right|$$

From (1) and from the choice of p we have

$$\left| \rho_n - \frac{1}{p} \sum_{k=0}^{p-n-1} s_k s_{k+n} \right| < \delta/2$$

From (2) we then have

$$\left| \rho_n - \frac{1}{p} \sum_{k=0}^{p-1} u_k u_{k+n} \right| \leqslant \delta/2 + \delta/4 \quad , \; |n| \leqslant n_0$$

We have shown that C is contained in the closure of P. However $C \supset P$ and C is closed. Therefore C= closure of P

This ends the proof of lemma 2.

Another very important property of the class C is as follows:

Lemma 3

The class C is convex.

Proof

We have already shown that the class C is closed. Therefore to show it is convex it suffices to show that, for any two elements of C, $(\rho'_n)_{n \in \mathbb{N}_0}$ and $(\rho''_n)_{n \in \mathbb{N}_0}$, the function $(\rho_n)_{n \in \mathbb{N}_0}$ defined by

$$\rho_n = \frac{1}{2} (\rho'_n + \rho''_n)$$

is again an element of C. The proof will be made by construction.

Let $(s'_n)_{n \in \mathbb{N}_0}$ and $(s''_n)_{n \in \mathbb{N}_0}$ be binary sequences with autocorrelation ρ' and ρ'' respectively. Define a new binary sequence $(s_n)_{n \in \mathbb{N}_0}$ as follows:

$$s_n = \begin{cases} s'_{n-j(j-1)/2} , & j(j-1) \leqslant n < j^2 \\ \\ s''_{n-j(j-1)/2-j} , & j^2 \leqslant n < j(j-1) + 2j \end{cases} , \quad j = 1,2,\ldots.$$

Now it is shown that the autocorrelation function of this sequence is $(\rho_n)_{n \in \mathbb{N}_0}$.

Fix $m \in \mathbb{N}_0$ and assume $j(j-1) \leqslant N < j^2$. Then we have

$$\frac{1}{N} \sum_{n=0}^{N} s_n s_{n+m} = \frac{1}{N} \sum_{n=0}^{N-j(j-1)/2} s'_n s'_{n+m} + \frac{1}{N} \sum_{n=0}^{j(j-1)/2} s''_n s''_{n+m} + R(j)$$

where $|R(j)|$ is bounded by $4jm/N$. As $j \to \infty$, and N as above, we have

$$\frac{1}{N} \sum_{n=0}^{N} s_n s_{n+m} \longrightarrow \frac{1}{2} (\rho'_m + \rho''_m) = \rho_m$$

since $R(j) \longrightarrow 0$, and $j(j-1)/2N \longrightarrow 1/2$.

The same holds if $j^2 \leqslant N < j(j+1)$.

This ends the proof of lemma 3.

The next lemma tells us that the class C can be studied by considering only zero mean binary sequences.

Lemma 4

Given a binary sequence $(s_n^*)_{n \in \mathbb{N}_0}$ with autocorrelation function $(\rho_n^*)_{n \in \mathbb{N}_0}$, there always exists a binary sequence $(s_n)_{n \in \mathbb{N}_0}$ such that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N} s_i = 0$$

and with autocorrelation $(\rho_n^*)_{n \in \mathbb{N}_0}$.

Proof

Refer to the proof of lemma 3 and set $s_n' = s_n^*$ and $s_n'' = -s_n^*$, for all $n \in \mathbb{N}_0$. Then it immediately follows that the resulting sequence has autocorrelation $(\rho_n^*)_{n \in \mathbb{N}_0}$ and zero mean.

q.e.d.

The convexity of the class C will play an important role in the establishment of convergent algorithms for the partial realization of autocorrelation functions.

The properties stated in lemmas 1-3 also hold for the class of unit covariances and are easier to establish for this class. With such important properties in common it is not surprising that the classes C and U are equivalent; this is shown in chapter 3.

2.4 Geometric Characterization of $\Pi_n C$

As we have mentioned in the introduction to the thesis, there are situations where the knowledge of $\Pi_n C$ is required, namely in the selection of optimal input autocorrelation functions. Therefore a geometric characterization of these sets is invaluable.

We start by establishing one of the main results in this chapter. Loosely speaking it says that $\Pi_n P$ is contained in a polytope, that is a bounded convex polyhedron whose vertices are autocorrelations of periodic binary sequences of periods not greater than 2^{m-1}.

Theorem 1

$$H\left(\Pi_m\left(\bigcup_{p=1}^{2^{m-1}} M^p \cap P\right)\right) \supset \Pi_m P.$$

In the proof of this theorem we shall make use of the next two lemmas:

Lemma 1

Let $(\rho_n)_{n \in \mathbb{N}_0}$ be the autocorrelation function of a periodic binary sequence $(s_n)_{n \in \mathbb{N}_0}$ of even period p, and such that $\rho_{p/2} = -1$. Then

$$\rho_m = \frac{2}{p} \sum_{i=0}^{p/2-1} s_i s_{i+m} \text{ , for any } m \in \mathbb{N}_0 .$$

Proof of lemma 1

$\rho_{p/2} = -1 \Longrightarrow s_i = -s_{i+p/2}$, for all i. Therefore

$$\rho_m = \frac{1}{p}\sum_{i=0}^{p-1} s_i s_{i+m} = \frac{1}{p}\left\{\sum_{i=0}^{p/2-1} s_i s_{i+m} + \sum_{i=p/2}^{p-1} s_i s_{i+m}\right\} =$$

$$= \frac{1}{p}\left\{\sum_{i=0}^{p/2-1} s_i s_{i+m} + \sum_{i=p/2}^{p-1} (-s_{i-p/2})(-s_{i-p/2+m})\right\} =$$

$$= \frac{2}{p}\left\{\sum_{i=0}^{p/2-1} s_i s_{i+m}\right\} .$$

q.e.d.

Lemma 2

Let $(s_n)_{n \in \mathbb{N}_0}$ be an element of C. Define the m×m dyadic matrices

$$E_s^{m,i} = (s_i, \ldots, s_{i+m-1})^T (s_i, \ldots, s_{i+m-1}).$$

If for some integers q and p we have

$$\Pi_{m-1} E_s^{m,q} \neq \Pi_{m-1} E_s^{m,k+q} \quad , \; k=1,\ldots,p-1, \text{ and}$$

$$\Pi_{m-1} E_s^{m,q} = \Pi_{m-1} E_s^{m,p+q} \quad , \text{ then}$$

$$\sum_{i=q}^{p+q-1} E_s^{m,i} = p \begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{m-1} \\ & \ddots & & \vdots \\ & & \rho_1 & \\ & & & \rho_0 \end{bmatrix} \quad , \text{ where } \rho_i, \; i=0,\ldots,m-1$$

are the first m autocorrelation values of a periodic

binary sequence of period p or 2p depending on $(s_q,\ldots,s_{m+q-1})$ being equal to or the negative of $(s_{q+p},\ldots,s_{q+p+m-1})$ respectively.

Proof of lemma 2

Without any loss of generality assume q=0. Two situations must be considered:

(i) $(s_o,\ldots,s_{m-2}) = (s_p,\ldots,s_{p+m-2})$

(ii) $(s_o,\ldots,s_{m-2}) = -(s_p,\ldots,s_{p+m-2})$

(i) Define a periodic binary sequence $(y_n)_{n\in\mathbb{N}_0}$ of period p as follows:

$$y_n = s_n \ , \ 0\leqslant n\leqslant p-1$$

From our assumption we also have $y_n = s_n$, $0\leqslant n\leqslant p+m-2$. Denote the autocorrelation function of $(y_n)_{n\in\mathbb{N}_0}$ by $(\rho_n)_{n\in\mathbb{N}_0}$. Then

$$\begin{bmatrix} \rho_o & \rho_1 & \cdots & \rho_{m-1} \\ & \ddots & & \vdots \\ & & & \rho_1 \\ & & & \rho_o \end{bmatrix} = \frac{1}{p}\sum_{i=0}^{p-1} \begin{bmatrix} y_i \\ \vdots \\ y_{i+m-1} \end{bmatrix} \begin{bmatrix} y_i & \cdots & y_{i+m-1} \end{bmatrix} =$$

$$= \frac{1}{p}\sum_{i=0}^{p-1} \begin{bmatrix} s_i \\ \vdots \\ s_{i+m-1} \end{bmatrix} \begin{bmatrix} s_i & \cdots & s_{i+m-1} \end{bmatrix}$$

(ii) This case can be proved similarly, using lemma 1.

q.e.d.

We are now in a position to give the proof of theorem 1.

Proof of theorem 1

All we need to show is that if $(\rho_n)_{n\in\mathbb{N}_0}$ is the autocorrelation function of a periodic binary sequence $(s_n)_{n\in\mathbb{N}_0}$, of period p, then there exists a finite number of periodic binary sequences $(s_n^i)_{n\in\mathbb{N}}$, $i= 1,2,\ldots,N$, of periods not greater than 2^{m-1} such that $(\rho_o,\ldots,\rho_{m-1}) \in H\{(\rho_o^1,\ldots,\rho_{m-1}^1),\ldots,(\rho_o^N,\ldots,\rho_{m-1}^N)\}$ where $(\rho_n^i)_{n\in\mathbb{N}_0}$ denotes the autocorrelation of $(s_n^i)_{n\in\mathbb{N}_0}$.

If $p\leqslant 2^{m-1}$ the result follows. So consider the case where $p> 2^{m-1}$. By definition we have

$$\begin{bmatrix} \rho_o & \rho_1 & \cdots & \rho_{m-1} \\ & \ddots & & \vdots \\ & & \ddots & \rho_1 \\ & & & \rho_o \end{bmatrix} = \frac{1}{p}\sum_{i=0}^{p-1} \begin{bmatrix} s_i \\ \vdots \\ s_{i+m-1} \end{bmatrix} \begin{bmatrix} s_i & \cdots & s_{i+m-1} \end{bmatrix} =$$

$$= \frac{1}{p}\,(E_s^{m,o},\ldots,E_s^{m,p-1}).$$

If we consider the first $2^{m-2}+1$ elements of the sum in brackets on the right hand side we get for some j and q, $0\leqslant j,q\leqslant p-1$

$$\Pi_{m-1}E_s^{m,j} = \Pi_{m-1}E_s^{m,j+q} \quad \text{with } 0<j+q\leqslant 2^{m-2} \quad \text{and}$$

$$\Pi_{m-1}E_s^{m,j} \neq \Pi_{m-1}E_s^{m,j+k} \qquad \text{for } 1 \leqslant k \leqslant q-1 .$$

Set $p_1=q$. From lemma 2 we have that

$$\sum_{i=0}^{p_1-1} E_s^{m,j+i} = p_1 \begin{bmatrix} \rho_0^1 & \rho_1^1 & \cdots & \rho_{m-1}^1 \\ & & & \vdots \\ & & & \rho_1^1 \\ & & & \rho_0^1 \end{bmatrix}$$

where $(\rho_n^1)_{n \in N_0}$ is the autocorrelation function of a periodic binary sequence $(s_n^1)_{n \in N_0}$ defined as follows:

(i) if $(s_j,\ldots,s_{j+m-2}) = (s_{j+p_1},\ldots,s_{j+p_1+m-2})$ we set the period of $(s_n^1)_{n \in N_0}$ equal to p_1 and define $s_n^1=s_n$ for $j<n<j+p_1-1$.

(ii) if $(s_j,\ldots,s_{j+m-2}) = -(s_{j+p_1},\ldots,s_{j+p_1+m-2})$ we set the period of $(s_n^1)_{n \in N_0}$ equal to $2p_1$ and define $s_n^1=s_n$ for $j \leqslant n \leqslant j+p_1-1$, and $s_n^1=-s_{n-p_1}^1$ for $j+p_1 \leqslant n \leqslant j+2p_1-1$.

As far as the sum of the remaining matrices is concerned, we can write

$$E_s^{m,1}+\ldots+E_s^{m,j-1}+E_s^{m,j+p_1}+\ldots+E_s^{m,p} =$$

$$= E_s^{m,j+p_1}+\ldots+E_s^{m,p}+E_s^{m,1}+\ldots+E_s^{m,j-1}$$

$$= \sum_{i=j+p_1}^{p+j-1} \begin{bmatrix} s_i \\ \vdots \\ s_{i+m-1} \end{bmatrix} \begin{bmatrix} s_i & \cdots & s_{i+m-1} \end{bmatrix}$$

because $(s_n)_{n \in N_0}$ has period p.

For the sake of simplicity of presentation define a new binary periodic sequence $(u_n)_{n \in \mathbb{N}_0}$ as follows:

(i) if $(s_{j+p_1},\ldots,s_{j+p_1+m-1}) = (s_{j+p},\ldots,s_{p+j+m-1})$ set the period of $(u_n)_{n \in \mathbb{N}_0}$ equal to $(p-p_1)$ and define

$$u_i = s_{j+p_1+i} \quad , \; i=0,1,\ldots,p-p_1-1$$

(ii) if $(s_{j+p_1},\ldots,s_{j+p_1+m-1}) = -(s_{j+p},\ldots,s_{p+j+m-1})$ then we set the period of $(u_n)_{n \in \mathbb{N}_0}$ equal to $2(p-p_1)$ and let

$$u_i = s_{j+p_1+i} \quad , \; i=0,1,\ldots,p-p_1-1 \quad \text{and}$$

$$u_i = -u_{i-p+p_1} \quad , \; i=p-p_1,\ldots,2(p-p_1)-1 .$$

We can then write

$$\begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{m-1} \\ & \ddots & & \vdots \\ & & & \rho_1 \\ & & & \rho_0 \end{bmatrix} = \frac{p_1}{p} \begin{bmatrix} \rho_0^1 & \rho_1^1 & \cdots & \rho_{m-1}^1 \\ & \ddots & & \vdots \\ & & & \rho_1^1 \\ & & & \rho_0^1 \end{bmatrix} +$$

$$+ \frac{1}{p} \sum_{i=0}^{p-p_1-1} \begin{bmatrix} u_i \\ \vdots \\ u_{i+m-1} \end{bmatrix} \begin{bmatrix} u_i & \cdots & u_{i+m-1} \end{bmatrix}$$

We can again apply the same procedure to the second term on the right hand side of this equation and obtain

$$\begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{m-1} \\ & & & \vdots \\ & & & \rho_1 \\ & & & \rho_0 \end{bmatrix} = \frac{p_1}{p} \begin{bmatrix} \rho_0^1 & \rho_1^1 & \cdots & \rho_{m-1}^1 \\ & & & \vdots \\ & & & \rho_1^1 \\ & & & \rho_0^1 \end{bmatrix} +$$

$$+ \frac{p_2}{p} \begin{bmatrix} \rho_0^2 & \rho_1^2 & \cdots & \rho_{m-1}^2 \\ & & & \vdots \\ & & & \rho_1^2 \\ & & & \rho_0^2 \end{bmatrix} + \frac{1}{p} \sum_{i=0}^{p-p_1-p_2-1} \begin{bmatrix} v_i \\ \vdots \\ v_{i+m-1} \end{bmatrix} \begin{bmatrix} v_i & \cdots & v_{i+m-1} \end{bmatrix},$$

where $(\rho_i^2)_{i \in \mathbb{N}_0}$ is the autocorrelation of a periodic binary sequence of period p_2 or $2p_2$, $p_2 \leqslant 2^{m-2}$, and $(v_i)_{i \in \mathbb{N}_0}$ is a periodic binary sequence of period $(p-p_1-p_2)$ or $2(p-p_1-p_2)$.

And so on until, for some $N \in \mathbb{N}$, we have $p - \sum_{i=1}^{N-1} p_i \leqslant 2^{m-2}$, since each p_i is greater than 0.

Denote the sequence we end up with by $(s_i^N)_{i \in \mathbb{N}_0}$ and set $p_N = p - \sum_{i=1}^{N-1} p_i$. Two situations need still to be considered:

(i) $(s_i^N)_{i \in \mathbb{N}_0}$ has period p_N or $2p_N$.

(ii) The period of $(s_i^N)_{i \in \mathbb{N}_0}$ is a factor of p_N or $2p_N$.

In both cases

$$\frac{1}{p_N} \sum_{i=0}^{p_N-1} \begin{bmatrix} s_i^N \\ \vdots \\ s_{i+m-1}^N \end{bmatrix} \begin{bmatrix} s_i^N & \cdots & s_{i+m-1}^N \end{bmatrix} = \begin{bmatrix} \rho_0^N & \rho_1^N & \cdots & \rho_{m-1}^N \\ & & & \vdots \\ & & & \rho_1^N \\ & & & \rho_0^N \end{bmatrix}$$

where $(\rho_j^N)_{j \in \mathbb{N}_0}$ denotes the autocorrelation function of $(s_i^N)_{i \in \mathbb{N}_0}$. At this stage we have

$$\begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{m-1} \\ & \ddots & & \vdots \\ & & \ddots & \rho_1 \\ & & & \rho_0 \end{bmatrix} = \sum_{i=1}^{N} \frac{p_i}{p} \begin{bmatrix} \rho_0^i & \rho_1^i & \cdots & \rho_{m-1}^i \\ & \ddots & & \vdots \\ & & \ddots & \rho_1^i \\ & & & \rho_0^i \end{bmatrix}$$

where $p = \sum_{i=1}^{N} p_i$ and $(\rho_n^i)_{n \in \mathbb{N}_0}$, $i=1,2,\ldots,N$, are autocorrelation functions of periodic binary sequences of period not greater than 2^{m-1}.

This completes the proof of theorem 1.

We are now in a position to give a geometric characterization of $\mathbb{H}_m C$, for all m, $m \in \mathbb{N}$.

Theorem 2

$\mathbb{H}_m C$ is a polytope, whose vertices $V_i \triangleq (v_0^i, \ldots, v_{m-1}^i)$ are given by

$$v_j^i \triangleq \frac{1}{p_i} \sum_{n=0}^{p_i-1} s_n^i s_{n+j}^i \quad , \quad j=0,1,2,\ldots,m-1$$

where $(s_n^i)_{n \in \mathbb{N}_0}$ is a periodic binary sequence of period $p_i \leqslant 2^{m-1}$.

Proof

We have already shown that C is convex and compact in the product topology. Therefore the set $\Pi_m C$ is convex and compact in the usual topology of $\mathbb{R}^m$. Then it immediately follows that

$$H(\Pi_m(\bigcup_{p=1}^{2^{m-1}} M^p \cap P)) \subset \Pi_m C.$$

Note that $H(\Pi_m(\bigcup_{p=1}^{2^{m-1}} M^p \cap P))$ is closed being the convex hull of a finite number of points. From theorem 1 we have

$$H(\Pi_m(\bigcup_{p=1}^{2^{m-1}} M^p \cap P)) \supset \Pi_m P.$$

However P is dense in C (lemma 2, section 2.3). Therefore closure $(\Pi_m P) = \Pi_m C$.

But this implies that

$$H(\Pi_m(\bigcup_{p=1}^{2^{m-1}} M^p \cap P)) \supset \Pi_m C.$$

$\bigcup_{p=1}^{2^{m-1}} M^p \cap P$ contains only a finite number of uniformly bounded elements. Therefore

$$H(\Pi_m(\bigcup_{p=1}^{2^{m-1}} M^p \cap P)) \text{ is a polytope.}$$

This completes the proof of theorem 2.

We now state a very interesting result which is a corollary of the previous theorem.

Theorem 3

A point $(a_0,a_1,\ldots,a_{m-1})$ in ${}^{\mathbb{H}}{}_m C$ can always be extrapolated to a periodic autocorrelation function of a binary sequence.

We wish to draw attention to the fact that $(\rho_n)_{n\in\mathbb{N}_0}$ being a periodic element of C does not imply that it is the autocorrelation of a periodic binary sequence. For example it can be shown that the following function

$$\rho_n = \begin{cases} 1, & n \text{ even} \\ 0, & n \text{ odd} \end{cases}$$

is the autocorrelation function of a binary sequence, although there is no periodic binary sequence with such an autocorrelation. If there was one then it should have period 2. However none of the four periodic binary sequences of period 2 has that autocorrelation.

Proof of theorem 3

From theorem 2 we have that

$$(a_0,a_1,\ldots,a_{m-1}) = \sum_{i=1}^{N} \alpha_i(\rho_0^i,\ldots,\rho_{m-1}^i)$$

where $\sum \alpha_i = 1$, $\alpha_i \geqslant 0$, and $(\rho_n^i)_{n\in\mathbb{N}_0}$ is a periodic autocorrelation.

Let $\rho_j = \sum_{i=1}^{N} \alpha_i \rho_j^i$ for all $j \in \mathbb{N}_0$. Then $(\rho_j)_{j \in \mathbb{N}_0}$ is periodic, with $\rho_j = a_j$ for $0 \leqslant j \leqslant m-1$, and from the convexity of C it is also in C.

The result stated in theorem 3 is rather surprising in contrast with the case of general positive semi-definite sequences where it does not necessarily hold. This can be shown as follows:

The extreme points of the set of possible values of ρ_1 and ρ_2 where $(\rho_n)_{n \in \mathbb{N}_0}$ is a positive semi-definite sequence with $\rho_o = 1$ are the locus defined by (cos w, cos 2w) for $-\Pi \leqslant w \leqslant \Pi$. See figure 11. Suppose w_o is irrational in this range. Then cos w_o, cos $2w_o$ can only be the second and third values of the sequence $(\cos nw_o)_{n \in \mathbb{N}_0}$. This follows from Bochner's theorem which states that for any positive semi-definite sequence $(\rho_n)_{n \in \mathbb{N}_0}$, with $\rho_o = 1$

$$\rho_n = \int_{-\Pi}^{\Pi} \cos nw \, dG(w)$$

for some unit measure on $[-\Pi, \Pi]$; so if G were not Dirac measure at w_o then (ρ_1, ρ_2) would differ from (cos w_o, cos $2w_o$). Hence (cos w_o, cos $2w_o$) has no periodic extension.

In theorem 2 we gave a characterization of the finite dimensional projections of C which is of interest in practical problems. For the sake of completeness we now give a complete characterization of the class C.

Theorem 4

$(\rho_n)_{n\in\mathbb{N}_0} \in C$ if and only if

$$\begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{m-1} \\ & \searrow & \searrow & \vdots \\ & & & \rho_1 \\ & & & \rho_0 \end{bmatrix} \in H\left(\Pi_m\left(\bigcup_{p=1}^{2^{m-1}} M^p \cap P\right)\right), \text{ for all } m\in\mathbb{N}.$$

Proof

The necessity of the condition follows from theorem 2. To prove sufficiency we can use theorem 3. Then for each m we have a periodic autocorrelation $(\rho_i^m)_{i\in\mathbb{N}_0}$ such that $\rho_i^m=\rho_i$, $i\leqslant m-1$. But this means that $(\rho_n)_{n\in\mathbb{N}_0}$ is the pointwise limit of a sequence of elements of C. Therefore $(\rho_n)_{n\,c\mathbb{N}_0} \in C$, because C is closed (lemma 1, section 2.3).

2.5 Calculation of the Vertices of $\Pi_m C$

The importance of these vertices has already been pointed out in the introduction to this chapter. The previous theorems have revealed that if $(a_0,a_1,\ldots,a_{m-1})$ is a vertex of $\Pi_m C$ then there exists a periodic binary

sequence $(s_n)_{n \in N_0}$ of period p, $p \leqslant 2^{m-1}$ such that

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_{m-1} \\ & & & \vdots \\ & & & a_1 \\ & & & a_0 \end{bmatrix} = \frac{1}{p} \sum_{i=0}^{p-1} \begin{bmatrix} s_i \\ \vdots \\ s_{i+m-1} \end{bmatrix} \begin{bmatrix} s_i & \cdots & s_{i+m-1} \end{bmatrix} \qquad (1)$$

Therefore a possible way of constructing the vertices and associated sequences with the smallest possible period is to construct all the periodic binary sequences of periods not greater than 2^{m-1} (there is only a finite number of them) and their first m autocorrelation values and then select the vertices of the convex hull of this finite set of points.

However we only need to consider a subset of these sequences as the next lemma shows:

Lemma 1

If in the right hand side of expression (1) above there are two matrices with equal $(m-1)\times(m-1)$ upper left corners, then either $(a_0, a_1, \ldots, a_{m-1})$ is not a vertex of $\Pi_m C$ or there exists a binary sequence with smaller period with these autocorrelation values.

Proof

If there are two such matrices then we can express (1) as a convex combination of two autocorrelation matrices by lemma 2, section 2.4. If these matrices are distinct then $(a_0, a_1, \ldots, a_{m-1})$ is not a vertex. If they are equal this means that there exists a sequence with period less than p and with its first m autocorrelation values equal to $(a_0, a_1, \ldots, a_{m-1})$.

q.e.d.

Bearing in mind this result an algorithm has been implemented to generate a set of periodic binary sequences of periods not greater than 2^{m-1} and their fist m autocorrelation values such that their convex hull is $\Pi_m C$. This is done in appendix A.

The calculation of the vertices of the convex hull of a finite number of points in $\mathbb{R}^m$ gives rise to a variety of problems, particularly when the number of points is 'large'.

Suppose that one is given a finite set of points $p_1, \ldots, p_N$ and wishes to select the vertices of their convex cover $H\{p_1, \ldots, p_N\}$. This polytope will be referred as X. Three selection methods will now be proposed. The discussion of their relative merits is given later.

I) Coordinate Rotation

Some results from convex set theory enable us to select, by simple inspection, vertices of X. For example: (i) The point with the largest or the smallest ith component, i=1,2,...,m, is a vertex of X ; (ii) The farthest point from the origin is a vertex of X.

Although not all the vertices satisfy these conditions in general we can still select them using fact (i) above by a suitable set of rotations of the axis; each vertex will satisfy (i) for at least one such rotation. It can easily be shown that at least one such set of rotations exists as follows: for each vertex v of X take a supporting hyperplane Π_v of X such that $\Pi_v \cap X = v$. Let one of the axis become orthogonal to this hyperplane. Then vertex v satisfies condition (i) for this axis (we are assuming an orthogonal basis).

II) Simplex Decomposition

This method relies on the fact that the vertices of an n-dimensional polytope lie in the intersection of at least n extreme supporting hyperplanes (or faces). This condition is particularly convenient for us, because we already have a finite set of points from X containing its vertices. By 'extreme' supporting hyperplanes we

mean supporting hyperplanes whose intersection with X is a polytope with one dimension less than X. We call this intersection a face.

A straightforward way of constructing these extreme supporting hyperplanes is to consider all possible combinations of n points at a time from $\{p_1,\ldots,p_N\}$ and for those defining an hyperplane testing if it is a supporting hyperplane of X. If yes it will be also an extreme supporting hyperplane.

We now give a procedure to compute the extreme supporting hyperplanes that does not require us to consider all possible hyperplanes defined by $\{p_1,\ldots,p_N\}$ and that substantially reduces the required amount of work. In this method X is obtained as the last element of a finite (nested) sequence of increasing polytopes defined in terms of their boundaries. This procedure is based on a paper by Degtyar and Finkel'shteyn [2] where a method to decompose a polytope as a union of disjoint simplices is described.

To avoid minor technicalities assume $H\{p_1,\ldots,p_N\}$ is a solid m-dimensional polyhedron, where m is the dimension of the vector p_i, $i=1,\ldots,N$.

We now describe the algorithm:

Step 1 Set K=1.

Construct a simplex through m+1 of the given points and denote it by polytope 1. Let us call its faces the elements of the boundary. For K>1 these boundary elements will not be in general faces but n-1 simplices that are disjoint subsets of the faces of polytope K. In an attempt to enclose as many points of $\{p_1,\ldots,p_N\}$ as possible, to minimize the required amount of work, it is suggested that points are selected with as large or as small components as possible.

Set $J = \{p_1,\ldots,p_N\} \setminus \{\text{Set of vertices of polytope 1}\}$

Step 2 Test if the first element of J is an exterior point of polytope K. If not remove it from J and repeate the same procedure for the next element of J and so on (The boundary of each intermediate polytope is known. Therefore to test if a point belongs to one of them it is only required to verify a set of linear inequalities). In the course of this, two situations are then possible: (i) A point of J is eventually found lying outside polytope K, say p_q. Then we define polytope (K+1) = H(polytope $K \cup p_q$), remove point

p_q from J and go to step 3. (ii) All the points of J are points of polytope K. This means that polytope K = H $\{p_1,\ldots,p_N\}$ and therefore the construction is finished.

Let us say that a collection of boundary elements are affinely independent if they each lie in distinct affinely independent hyperplanes. Then the vertices of H $\{p_1,\ldots,p_N\}$ are those boundary element vertices that are common to m affinely independent boundary elements.

Step 3 Here the boundary of polytope (K+1) is constructed. We start by separating the elements of the boundary of polytope K into two disjoint classes:

Class A - This class contains those elements of the boundary defining an hyperplane which produces a closed half space containing polytope K and p_q.

Class B - Contains all the remaining elements of the boundary.

The elements of the boundary of polytope (K+1) are the elements of the boundary of polytope K in class A and the (m-1) dimensional simplices constructed as follows:

Each of them is defined by p_q and (m-1) vertices of an element of the boundary of polytope K in

class A and in the intersection with an element in class B.

Step 4 Set K = K+1 and go to step 2.

III) Distance Minimization

This procedure relies on the following property: Given a set of points $\{p_1,\ldots,p_N\}$ we have that p_i is a vertex of $H\{p_1,\ldots,p_N\}$ if and only if the distance from p_i to $H\{p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_N\}$ is greater than zero.

To compute the distance from a point to the convex hull of others it is required to solve, in principle, a constrained minimization problem - in fact a geometric programming problem as follows:

$$\min f(\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_N) = \Big\| p_i - \sum_{\substack{j=1\\ j\neq i}}^{N} \alpha_j p_j \Big\|$$

subject to the constraints $\sum_{\substack{j=1\\ j\neq i}}^{N} \alpha_j = 1$, $\alpha_j \geqslant 0$ for all j.

Fortunately the constrained optimization can be avoided by means of the following transformation:

$$\alpha_j = \beta_j^2 \Big/ \Big(\sum_{\substack{n=1\\ n\neq i}}^{N} \beta_n^2 \Big)$$

and then the problem becomes

$$\text{minimize } f(\beta_1,\ldots,\beta_{i-1},\beta_{i+1},\ldots,\beta_N) = \left\| p_i - \sum_{\substack{j=1\\ j\neq i}}^{N} \frac{\beta_j^2}{\sum_{\substack{n=1\\ n\neq i}}^{N} \beta_n^2}\, p_j \right\| .$$
$$\text{over } \mathbb{R}^{N-1}$$

Discussion of the proposed methods

In the first method there is the problem of constructing a suitable set of rotations of the axis. We do not know of any way of doing it without making use of supporting hyperplanes. But if supporting hyperplanes need to be considered then there is no point in using this method because the second one is particularly suitable for that purpose.

We believe that the best way to appreciate the elegance of Degtyar's and Finkel'shteyn's idea used in the second method is to compare it with the rough procedure of having to construct all the hyperplanes defined by $\{p_1,\ldots,p_N\}$ and test which of them are actually supporting hyperplanes of $H\{p_1,\ldots,p_N\}$. In Degtyar's procedure the only subsets of $\{p_1,\ldots,p_N\}$ we consider are the ones whose points define an hyperplane containing a face of the intermediate polytopes. Furthermore each polytope in this sequence strictly contains its predecessor. Therefore

those combinations of points of $\{p_1,\ldots,p_N\}$ defining hyperplanes slicing polytopes already constructed are no longer possible and this is particularly significant when the intermediate polytopes get bigger and bigger. Furthermore at every stage the method provides sufficient conditions for a given set of points to define a supporting hyperplane of the polytope concerned. Therefore no time is wasted in the construction of 'unnecessary' hyperplanes. Furthermore if at each step one tries to construct a polytope as large as possible it is more likely to enclose more points of $\{p_1,\ldots,p_N\}$ and therefore reducing the number of polytopes required to reach $H\{p_1,\ldots,p_N\}$. We have not tested this method numerically and therefore no accurate comparison can be made with the third method which we have used in the generation of the results given in the last section.

The third method is particularly attractive, because of its simplicity of implementation and speed of execution. We have used it in conjunction with the algorithm given in appendix A to generate the vertices of $\Pi_m C$ for some values of m. This method performed very well together with a particular type of minimization algorithm (and some speed up features) using a 'pattern search' technique which does not require gradient evaluations. The function minimization is performed by constructing a sequence $(c_n)_{n\in N}$ of points

of H $\{p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_N\}$ converging to the nearest point of p_i. A well known result in optimization theory states that if a point p lies outside a convex set S, then there exists a unique point in S, say s_o, such that

$$\|p - s_o\| \leqslant \|p - s\| \quad , \quad \forall s \in S.$$

Furthermore the hyperplane passing through s_o and orthogonal to vector $p-s_o$ is a supporting hyperplane of J. This result suggested the inclusion of the following 'speed up' feature in this algorithm: Every n_o iterations, $n_o \in N$, it is checked if the hyperplane orthogonal to vector c_n-p_i, at p_i, produces an open half space containing H $\{p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_N\}$. If p_i actually lies outside the convex hull this eventually becomes true, and therefore there is no need to carry on with the minimization procedure. This feature was introduced in the algorithm and it was found that it greatly reduced the required amount of iterations when p_i was a vertex. n_o should be fixed initially. The reason why we do not make $n_o=1$ is because this test is a waste of time when the point actually lies inside the convex hull. Therefore its choice is the result of a compromise depending on our 'feeling' about the point under test.

Obviously the minimization procedure becomes slower as the dimension of the space increases, and there will

be a stage where it will be convenient to divide $H\{p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_N\}$ into a union of as many disjoint convex hulls as necessary to make the minimization procedure feasible in each of them. Needless to say that as soon as a point is found not to be a vertex it should be immediately discarded from the definition of the convex hull because it is the number of these points that determines the dimension of the space we are optimizing over.

If N is very large and only a few of the given points are actually vertices of their convex cover then the second method might be worth considering. However for small values of N the third method is undoubtly much faster than the second.

As far as the accuracy of method 3 is concerned, we have that when it decides that a point is a vertex, a separating hyperplane is actually produced. For points 'very' close to the boundary of $H\{p_1,\ldots,p_N\}$ the decision will depend on the prescribed limits of accuracy.

2.6 Connections with the Frequency Domain

Although the first n autocorrelation shifts do not completely characterize the discrete time signal they provide us with valuable information about the number of frequencies it contains. In fact it can be shown [3] that the n×n matrix

$$R_n \triangleq \begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{n-1} \\ & \ddots & & \vdots \\ & & \ddots & \rho_1 \\ & & & \rho_0 \end{bmatrix}$$

where $(\rho_k)_{k \in N_0}$ denotes the autocorrelation of the signal $(u_k)_{k \in N_0}$, is <u>positive</u> definite if and only if the support of the spectral distribution of $(u_k)_{k \in N_0}$ has at least n points, that is, the signal contains at least n frequencies. The signal is then said to be persistently exciting of order n.

This notion is useful in connection with identification problems where it is necessary to have some conditions on the input to get consistent estimates. See for example [1]. Particularly in the case of parametric identification it is required to have persistent excitation of some finite order, depending on the number of parameters to be identified.

2.7 Computational Results

In this section the vertices of $\Pi_m C$, and associated periodic binary sequences are given for values of m up to 6. These results have been generated by a computer program that makes use of method 3 as described in section 2.5.

Denote each element of $\Pi_m C$ by $(\rho_0, \rho_1, \ldots, \rho_{m-1})$. Because ρ_0 is always equal to 1 we shall implicitly assume all the points of $\Pi_m C$ in the linear manifold $\rho_0 = 1$ and therefore regard them as elements of $\mathbb{R}^{m-1}$.

The shapes of $\Pi_3 C$ and $\Pi_4 C$ are depicted in figures 11 and 12 respectively.

It can be seen that the vertices of $\Pi_m C$, $m \leqslant 5$, are also arquitype covariances, that is of the form $\rho_n = \frac{2}{\pi} \text{ arc sin cos } 2\pi n\lambda$, $-\frac{1}{2} \leqslant \lambda \leqslant \frac{1}{2}$. The corresponding values of λ are also indicated. 'p' denotes the period of the periodic extrapolation of the vertex, with the shortest period.

The results are now given in the tables below.

m	Vertices of $\Pi_m C$	λ	p	Periodic extrapolation : $(\rho_n)_{n \in N_0}$	Associated periodic sequence
2	1	0	1	(1,1,...)	(1,...)
	-1	1/2	2	(1,-1,1,...)	(1,-1,...)
3	(1,1)	0	1	(1,1,...)	(1,...)
	(0,-1)	1/4	4	(1,0,-1,0,1,...)	(1,1,-1,-1,...)
	(-1,1)	1/2	2	(1,-1,1,...)	(1,-1,...)
4	(1,1,1)	0	1	(1,1,...)	(1,...)
	(1/3,-1/3,-1)	1/6	6	(1,1/3,-1/3,-1,-1/3,1/3,1,...)	(1,1,1,-1,-1,-1,...)
	(0,-1,0)	1/4	4	(1,0,-1,0,1,...)	(1,1,-1,-1,...)
	(-1/3,-1/3,1)	1/3	3	(1,-1/3,-1/3,1,...)	(1,1,-1,...)
	(-1,1,-1)	1/2	2	(1,-1,1,...)	(1,-1,...)
5	(1,1,1,1)	0	1	(1,1,...)	(1,...)
	(.5,0,-.5,-1)	1/8	8	(1,.5,0,-.5,-1,-.5,0,.5,1,...)	(1,1,1,1,-1,-1,-1,-1,...)
	(1/3,-1/3,-1,-1/3)	1/6	6	(1,1/3,-1/3,-1,-1/3,1/3,1,...)	(1,1,1,-1,-1,-1,...)
	(0,-1,0,1)	1/4	4	(1,0,-1,0,1,...)	(1,1,-1,-1,...)
	(-1/3,-1/3,1,-1/3)	1/3	3	(1,-1/3,-1/3,1,...)	(1,1,-1,...)
	(-.5,0,.5,-1)	3/8	8	(1,-.5,0,.5,-1,.5,0,-.5,1,...)	(1,-1,1,-1,-1,1,-1,1,...)
	(-1,1,-1,1)	1/2	2	(1,-1,1,...)	(1,-1,...)

m	Vertices of $\Pi_m C$	λ	p	Periodic extrapolation : $(\rho_n)_{n \in N_0}$	Associated periodic sequence
	(1,1,1,1,1)	0	1	(1,1,...)	(1,...)
	(.6,.2,-.2,-.6,-1)	1/10	10	(1,.6,.2,-.2,-.6,-1,-.6,-.2,.2,.6,1,...)	(1,1,1,1,1,-1,-1,-1,-1,-1,...)
	(.5,0,-.5,-1,-.5)	1/8	8	(1,.5,0,-.5,-1,-.5,0,.5,1,...)	(1,1,1,1,-1,-1,-1,-1,...)
	(1/3,-1/3,-1,-1/3,1/3)	1/6	6	(1,1/3,-1/3,-1,-1/3,1/3,1,...)	(1,1,1,-1,-1,-1,...)
	(.2,-.6,-.6,.2,1)	1/5	5	(1,.2,-.6,-.6,.2,1,...)	(1,1,1,-1,-1,...)
	(0,-1,0,1,0)	1/4	4	(1,0,-1,0,1,...)	(1,1,-1,-1,...)
6	(-.2,-.6,.6,.2,-1)	3/10	10	(1,-.2,-.6,.6,.2,-1,.2,.6,-.6,-.2,1,...)	(1,1,-1,1,1,-1,-1,1,-1,-1,...)
	(-1/3,-1/3,1,-1/3,-1/3)	1/3	3	(1,-1/3,-1/3,1,...)	(1,1,-1,...)
	(-.5,0,.5,-1,.5)	3/8	8	(1,-.5,0,.5,-1,.5,0,-.5,1,...)	(1,1,-1,1,-1,-1,1,-1,...)
	(-.6,.2,.2,-.6,1)	2/5	5	(1,-.6,.2,.2,-.6,1,...)	(1,1,-1,1,-1,...)
	(-1,1,-1,1,-1)	1/2	2	(1,-1,1,...)	(1,-1,...)
	(1/3,-1/3,-1/3,-1/3,1/3)	–	6	(1,1/3,-1/3,-1/3,-1/3,1/3,1,...)	(1,1,1,1,-1,-1,...)
	(-1/3,-1/3,1/3,-1/3,-1/3)	–	6	(1,-1/3,-1/3,1/3,-1/3,-1/3,1,...)	(1,1,-1,1,-1,-1,...)
	(-1/7,3/7,-1/7,-1/7,3/7)	–	7	(1,-1/7,3/7,-1/7,-1/7,3/7,-1/7,1,...)	(1,1,1,1,-1,1,-1,...)
	(1/7,3/7,1/7,-1/7,-3/7)	–	14	(1,1/7,3/7,1/7,-1/7,-3/7,-1/7,-1,-1/7,-3/7,-1/7,1/7,3/7,1/7,1,...)	(1,1,1,1,1,-1,1,-1,-1,-1,-1,-1,1,-1,...)

References

[1] K. J. Åström and P. Eykhoff, System Identification-- A Survey, Automatica, Vol. 7, 1971.

[2] V. U. Degtyar and M. ya Finkel'shteyn, Classification of Algorithms Based on Construction of Convex Hulls of Sets, Engineering Cybernetics, Vol. 12, No. 2, March-April 1974.

[3] L. Ljung, Characterization of the Concept of Persistently Exciting in the Frequency Domain, Report 7119, Lund Institute of Technology, Sweden.

[4] L. A. Shepp, Private Communications.

CHAPTER 3

THE CLASS OF UNIT COVARIANCES

3.1 Introduction

In 1955, B. Mcmillan [2] gave a characterization of the class of unit covariances. However, the conditions of his characterization are not easy to verify.

A more explicit characterization is given in this chapter that clarifies the structure of the class U. The key argument in the proof of this new characterization is the fact that U is compact in the product topology. This is shown by means of a result in functional analysis known as the Alaoglu theorem.

The realizations of any discrete time stochastic binary process, in short a unit process, are binary sequences. It is therefore natural to raise the question about the links between the class of unit covariances (ensemble averages) and the class of autocorrelation functions of binary sequences (time averages). The latter has already been studied in the previous chapter. We shall see that the sub-class of stationary discrete time unit covariances is equivalent to the class of auto-correlations of binary sequences.

The connections between the class of clipped

gaussian processes and the class of unit processes are also discussed.

3.2 Characterization of the Class of Unit Covariances

In 1955 B. Mcmillan [2] characterized the class U as follows:

Theorem 1

$(\rho_{i,j})_{i,j\in\mathbb{N}} \in U$ if and only if $\rho_{i,i}=1$ for all $i\in\mathbb{N}$ and $\sum_{i=1}^{m}\sum_{j=1}^{m}\rho_{i,j}a_{i,j}\geq 0$, $\forall m\in\mathbb{N}$ and all corner-positive matrices $\{a_{ij}\}$, $i,j=1,\ldots,m$, where a matrix $\{a_{i,j}\}$ is corner-positive if $\sum_{i=1}^{m}\sum_{j=1}^{m}a_{ij}x_ix_j\geq 0$ for every sequence $(x_1,\ldots,x_m)$ with $x_i=\pm 1$, $i=1, 2,\ldots, m$.

We are unaware of Mcmillan's original proof which was never published. However, L. Shepp gave an elegant proof in [4], which is reproduced in appendix E.

For the sake of completeness we mention a paper by E. Masry [1] where a characterization of a subclass of unit covariances associated with renewal processes is given.

Our characterization of the class U is now given:

Theorem 2

$(\rho_{i,j})_{i,j \in \mathbb{N}} \in U$ if and only if $\{\rho_{i,j}\}_m \in H\left\{E_1^m, \ldots, E_{2^m-1}^m\right\}$ for all $m \in \mathbb{N}$, where $\{\rho_{i,j}\}_m$ is an $m \times m$ matrix with (i,j)th element equal to $\rho_{i,j}$, and $E_i^m = e_i^t e_i$, $e_i = (x_1, \ldots, x_j, \ldots, x_m)$, $x_i = \pm 1$.

An immediate consequence of this theorem is that $\Pi_m U$ is a polytope, for all m. For example, for $m=3$ we have that $\Pi_3 U$ is the convex hull of the 'points':

$$E_1^4 = \begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1&1&1\end{bmatrix} = \begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix} \qquad E_2^4 = \begin{bmatrix}1\\1\\-1\end{bmatrix}\begin{bmatrix}1&1&-1\end{bmatrix} = \begin{bmatrix}1&1&-1\\1&1&-1\\-1&-1&1\end{bmatrix}$$

$$E_3^4 = \begin{bmatrix}1\\-1\\1\end{bmatrix}\begin{bmatrix}1&-1&1\end{bmatrix} = \begin{bmatrix}1&-1&1\\-1&1&-1\\1&-1&1\end{bmatrix} \qquad E_4^4 = \begin{bmatrix}1\\-1\\-1\end{bmatrix}\begin{bmatrix}1&-1&-1\end{bmatrix} = \begin{bmatrix}1&-1&-1\\-1&1&1\\-1&1&1\end{bmatrix}$$

The matrices E_i^m are in fact vertices of their convex hull. If we regard each of them as an element of $\mathbb{R}^{m^2}$ we have that they lie on the surface of a sphere with radius m^2 and centre at $(0, \ldots, 0)$. $\Pi_m U$ is a convex and closed set. This means that the class of unit covariances is convex and closed under pointwise limits (product topology).

Another interesting property of $\Pi_m U$ is that it is a neighbourly polytope that is a bounded convex polyhedron

where any pair of vertices is connected by an edge. This is shown in appendix B.

Theorem 1 and theorem 2 have to be necessarily equivalent because they characterize the same object. We now show this is in fact true.

All the results about cones and polarity to be used in this proof can be found in [5].

Let's start by establishing a one-to-one correspondence between the class of m×m matrices and $\mathbb{R}^{m^2}$. This enables us to identify the trace of the product of two matrices with the inner product of the associated vectors.

Denote by $(CP)^m$ the class of m×m corner-positive matrices. From the definition of a corner-positive matrix we have that $(-CP)^m$ is the polar cone of

$$\left\{E^m_1, \ldots, E^m_{2^m-1}\right\}, \quad \forall\, m \in \mathbb{N}.$$

The polar cone S^p of a set $S \subseteq \mathbb{R}^n$ is defined as

$$\left\{y \in \mathbb{R}^n \mid y^T x \leq 0 \ , \ \forall\, x \in S\right\}.$$

From theorem 1 we have that the set of matrices $\{\rho_{i,j}\}_m$ is the polar cone of $(-CP)^m$, $\forall\, m \in \mathbb{N}$. Therefore $\{\rho_{i,j}\}_m$ belongs to the polar of the polar of $\left\{E^m_1, \ldots, E^m_{2^m-1}\right\}$, $\forall m \in \mathbb{N}$, which can be shown [5] to be the conical (or positive) hull of $\left\{E^m_1, \ldots, E^m_{2^m-1}\right\}$.

This implies that

$$\{\rho_{i,j}\}_m = \sum_{l=1}^{2^{m-1}} \alpha_l E_l^m \quad , \quad \alpha_l \geq 0 \ , \ \forall l .$$

However, the main diagonal elements of the matrices on both sides of this equality are equal to 1. This implies that we must also have

$$\sum_{l=1}^{2^{m-1}} \alpha_l = 1 \ .$$

But this means that $\{\rho_{i,j}\}_m \in H\left\{E_1^m, \ldots, E_{2^{m-1}}^m\right\}$, $\forall m \in \mathbb{N}$, which is precisely the statement of theorem 2.

q.e.d.

Next we show that U is compact. This result will be of use in the direct proof of theorem 2. The topological concepts concerning the class U are referred to the product topology. Similar considerations to those made in chapter 2 about the topology of the class of autocorrelation functions of binary sequences can be made here.

Lemma

The class U is compact.

Proof

We start the proof of this lemma by showing that U is closed.

Suppose therefore (ρ^n) is a sequence in U converging pointwise to a limit f.

Consider the sample space of binary sequences $\Omega = \{-1, 1\}^{\mathbb{N}}$. Endowed with the product discrete topology it is compact Hausdorff.

Since each $\rho^n \in U$ there exist probability measures P^n on the corresponding Borel field such that

$$\rho^n_{i,j} = \int X_i X_j dP^n \quad ,$$

X_i being the ith coordinate of X.

Each P^n determines a positive linear functional I^n of unit norm ($I^n(1)=1$) on $C(\Omega)$. Since the set of positive linear functionals of unit norm is a closed subset of the unit ball, it is compact in the weak-star topology by Alaoglu's theorem [3] and so there exists a subsequence from (I^n) converging to a limit I.

By the Riesz representation theorem [3] there is a probability measure P such that $I(g) = \int g dP$ for all g in $C(\Omega)$.

Since $X_i X_j$ is continuous and ρ^n converges pointwise to f we have

$$f_{ij} = \int X_i X_j dP \quad .$$

But this means that f is a unit covariance.

U is compact because it is a closed subset of $[-1, 1]^{\mathbb{N}}$, which is a compact space.

This ends the proof of the lemma.

An immediate consequence of this lemma is

Corollary

The subclass of stationary unit covariances, TU, is compact.

We are now in a position to give a proof of theorem 2.

Proof of theorem 2

The necessity of both conditions is obvious since $\rho_{i,i} = E(X_i^2) = E(1) = 1$, and from the definition of covariance we have

$$\rho_{i,j} = EX_iX_j = \sum_{k=1}^{2^{m-1}} \alpha_k^m x_i^k x_j^k \quad , \; i,j \leqslant m \; ,$$

where α_k^m is the sum of the probabilities of symmetric realizations of the vector of the first m random variables that is $(X_1, \ldots, X_m)$.

We now prove sufficiency.

If $\{\rho_{i,j}\}_m \in H\left\{E_1^m, \ldots, E_{2^{m-1}}^m\right\}$ then we have that the (semi-infinite) matrix

$$\bar{R}^m \triangleq \begin{bmatrix} \{\rho_{i,j}\}_m & \cdots & \{\rho_{i,j}\}_m & \cdots \\ \vdots & & \vdots & \\ \{\rho_{i,j}\}_m & \cdots & \{\rho_{i,j}\}_m & \cdots \\ \vdots & & \vdots & \\ \{\rho_{i,j}\}_m & \cdots & \{\rho_{i,j}\}_m & \cdots \\ \vdots & & \vdots & \end{bmatrix}$$

also belongs to the convex hull of the (semi-infinite) matrices

$$\bar{E}_i^m \triangleq \begin{bmatrix} E_i^m & \cdots & E_i^m & \cdots \\ \vdots & & \vdots & \\ E_i^m & \cdots & E_i^m & \cdots \\ \vdots & & \vdots & \end{bmatrix} \quad , \; i=1, 2, \ldots, 2^{m-1} \; .$$

This means that $\bar{R}^m$ is the covariance matrix of a periodic unit process of period m, whose realizations are the first lines, and their negatives, of the semi-infinite matrices $\bar{E}_i^m$, with the sum of their probabilities given by the coefficients α_k^m of the above convex combination.

Denote one such process by $(\bar{X}_n^m)_{n \in \mathbb{N}}$ and the elements of the matrix $\bar{R}^m$ by $\bar{\gamma}_{k,l}^m$. By construction we have $\bar{\gamma}_{l,k}^m = \rho_{i,j}$, $l=i \bmod m$, $k=j \bmod m$.

Then if conditions of theorem 2 are satisfied we have that, $\forall l,k$

$$\bar{\gamma}^{m}_{l,k} \longrightarrow \rho_{l,k} \qquad \text{as } m \to \infty .$$

But it has already been shown that the class U is closed under pointwise limits. Therefore $(\rho_{i,j})_{i,j \in \mathbb{N}}$ is a unit covariance.

This ends the proof of the theorem.

In the previous chapter it has been shown that the class of periodic autocorrelation functions of binary sequences was dense in C. The next theorem shows that a similar result holds in TU (subclass of stationary unit covariances).

Theorem 3

The class of discrete time stationary periodic unit covariances is dense in TU.

Proof

From the proof of sufficiency of theorem 2 it follows that a stationary unit covariance is the pointwise limit of a sequence of periodic unit covariances but not necessarily stationary.

Let $(\rho_n)_{n \in \mathbb{N}_0}$ denote the given stationary unit covariance and by using the same procedure as in the

proof of theorem 2 construct a sequence of periodic unit covariances $(\bar{\gamma}^m_{i,j})_{i,j \in \mathbb{N}}$ of period m such that $(\bar{\gamma}^m_{i,j}) \longrightarrow (\rho_n)_{n \in \mathbb{N}_0}$ (pointwise convergence) as $m \to \infty$. Denote by $(\bar{X}^m_n)_{n \in \mathbb{N}}$ a periodic unit process with covariance $(\bar{\gamma}^m_{i,j})_{i,j \in \mathbb{N}}$.

By 'randomizing time' we derive a stationary periodic process $(X^m_n)_{n \in \mathbb{N}}$ from $(\bar{X}^m_n)_{n \in \mathbb{N}}$ as follows:

Let $X^m_n(w,s) = \bar{X}^m_{n+s}(w)$, where S is an uniformly distributed random variable on $X = \{0, 1, \ldots, m-1\}$ and independent of $(\bar{X}^m_n)_{n \in \mathbb{N}}$.

Then we have for given $i,j \in \mathbb{N}$:

$$E(X^m_i X^m_j) = \int_{\Omega \times X} X^m_i(w,s)\, X^m_j(w,s)\, d\mu =$$

$$= \int_X \int_\Omega \bar{X}^m_{i+s}(w)\, \bar{X}^m_{j+s}(w)\, d\mu_\Omega d\mu_X \text{ (by independence)} =$$

$$= \frac{1}{m} \sum_{s=0}^{m-1} \bar{\gamma}^m_{i+s,j+s}.$$

But the last expression is an average over m consecutive terms of a 'sequence' of period m. Therefore the process $(X^m_n)_{n \in \mathbb{N}}$ is <u>stationary</u>.

Now assume without any loss of generality that $m > j > i$. Then we have

$$\frac{1}{m} \sum_{s=0}^{m-1} \bar{\gamma}_{i+s,j+s} = \frac{1}{m} \sum_{s=0}^{m-1} \bar{\gamma}^m_{s+1,s+1+j-i} \quad \text{(by stationarity)}$$

$$= \frac{1}{m}\left[\sum_{s=0}^{m-1-(j-1)} \bar{\gamma}^m_{s+1,s+1+j-i} + \sum_{s=m-(j-i)}^{m-1} \bar{\gamma}^m_{s+1,s+1+j-i}\right]$$

$$= \frac{1}{m}\left[(m-j+i)\rho_{j-i} + (j-i)\rho_{m-j+i}\right]$$

This reveals that, for any given $i,j \in \mathbb{N}$, $E(X_i^m X_j^m) \longrightarrow \rho_{j-i}$, as $m \to \infty$.

The proof of theorem 3 is complete.

3.3 Relations Between the Class of Unit Covariances and the Class of Autocorrelation Functions of Binary Sequences

Our starting point in establishing the links between C and U is the following

Lemma 1

The autocorrelation function of a periodic binary sequence is also a stationary unit covariance.

Proof

Let $(\rho_n)_{n \in \mathbb{N}_0}$ denote the autocorrelation function of the periodic binary sequence $(s_n)_{n \in \mathbb{N}_0}$ of period p, that is

$$\rho_n = \frac{1}{p}\sum_{i=0}^{p-1} s_{i+n}\, s_i \ .$$

Now define a sample space Ω comprising all the ith shifted versions of $(s_n)_{n \in \mathbb{N}_0}$, denoted by $(s_k^i)_{k \in \mathbb{N}_0}$, $i=1,\ldots, p$, and assign probability $1/p$ to each of them.

Let $(X_k)_{k \in \mathbb{N}}$ be a stochastic process defined on this space, where each X_k is a coordinate variable in Ω, that is $X_k(s^i)=s_k^i$, $s^i \in \Omega$. The covariance of this process is given by

$$E(X_k X_{k+l}) = \frac{1}{p} \sum_{i=1}^{p} s_k^i \, s_{k+l}^i = \frac{1}{p} \sum_{i=1}^{p} s_{k+i} \, s_{k+l+i} = \rho_l$$

This completes the proof of lemma 1.

The realizations of a unit process with a periodic covariance are periodic binary sequences a.s.. Furthermore if the process is stationary we have that the probability of realizations that are a shifted version of each other is equal. This gives rise to the following result

Lemma 2

Every stationary periodic unit covariance is a convex combination of autocorrelations of periodic binary sequences.

Proof

Let $(X_n)_{n \in \mathbb{N}}$ denote a stationary periodic unit process of period p and covariance $(\rho_n)_{n \in \mathbb{N}_0}$.

Partition the sample space of this process into disjoint events A_i, $i=1, 2, \ldots, t$, where each A_i comprises all the realizations that are a shifted version of each other, and denote their period by p_i.

Denote each sequence in A_i by $(x_n^{i,j})_{n \in \mathbb{N}}$, $j=1, \ldots, p_i$ and such that $x_n^{i,j} = x_{n+j-1}^{i,1}$.

Then we can write

$$\rho_k = \int_{A_1+\cdots+A_t} X_1 X_{1+k} d\mu = \sum_{i=1}^{t} \left\{ \frac{\mu(A_i)}{p_i} \sum_{j=1}^{p_i} x_1^{i,j} \, x_{1+k}^{i,j} \right\}$$

$$= \sum_{i=1}^{t} \left\{ \frac{\mu(A_i)}{p_i} \sum_{j=1}^{p_i} x_{1+j-1}^{i,1} \, x_{1+k+j-1}^{i,1} \right\}$$

where $\sum_{i=1}^{t} \mu(A_i) = 1$, $\mu(A_i) \geqslant 0$.

But $\frac{1}{p_i} \sum_{j=1}^{p_i} x_{1+j-1}^{i,1} \, x_{1+k+j-1}^{i,1} = \rho_k^i$

where $(\rho_n^i)_{n \in \mathbb{N}_0}$ is the autocorrelation of any sequence in A_i. Therefore

$$\rho_k = \sum_{i=1}^{t} \mu(A_i) \rho_k^i \, , \quad \forall k \in \mathbb{N}_0 \, ,$$

that is, $(\rho_n)_{n \in \mathbb{N}_0}$ is a convex combination of autocorrelation functions of periodic binary sequences.

This ends the proof of lemma 2.

The properties of the classes C and TU we have established so far enable us to establish an important relation between these two classes as follows:

Theorem

The classes C and TU are equivalent.

This means that $(\rho_n)_{n \in \mathbb{N}_0}$ is a discrete time stationary unit covariance if and only if there exists a binary sequence $(s_n)_{n \in \mathbb{N}_0}$ such that, $\forall n \in \mathbb{N}_0$

$$\rho_n = \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N} s_i \, s_{i+n} \, .$$

A word of warning: the above theorem is not an ergodicity statement.

Proof

It has been shown in chapter 2 that P is dense in C. This result, together with lemma 1 and the fact that both C and TU are convex and closed, imply that $C \subset TU$.

From lemma 2 we have that the class of stationary periodic unit covariances is contained in C. But theorem 3 in section 3.2 states that this class is dense in TU. Therefore $TU \subset C$.

This ends the proof of the theorem.

Once we have shown that C and TU are equivalent and given that the generation of the vertices of $\Pi_m U$ is extremely simple, it is natural to raise the question why should we take then the burden of having to generate the vertices of $\Pi_m C$?

That is necessary mainly because:

(i) The binary sequences associated with the vertices of $\Pi_m C$ can be used as 'generators' of binary sequences with prescribed autocorrelations.

(ii) The characterization in terms of the matrices E_i^m does not provide the answer to the question:

Given a finite set of values $(a_0, a_1, \ldots, a_{m-1})$ such that

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_{m-1} \\ & & & \vdots \\ & & & a_1 \\ & & & a_0 \end{bmatrix} \in H\left\{E_1^m, \ldots, E_{2^m-1}^m\right\}$$

is there any $(\rho_n)_{n \in \mathbb{N}_0} \in C$ such that $\rho_n = a_n$, $0 \leqslant n \leqslant m-1$?

Unless we can show that

$$\Pi_m TU \stackrel{?}{=} \Pi_m T \Pi_m U.$$

Obviously $\Pi_m TU \subset \Pi_m T \Pi_m U$. Although a considerable effort has been made we have not yet been able to prove if the reverse inclusion is true. We strongly believe that the reverse inclusion might also be true. We have

verified that it is true for values of m up to 4 and we can find no arguments contradicting this assumption.
(iii) To extrapolate the autocorrelation function of a binary signal. We already know that if a signal is binary it is not sufficient that the extrapolated values give rise to a positive semi-definite sequence. It must also lie inside $\Pi_m C$, for all m, which means the knowledge of the polytopes $\Pi_m C$ is required.
(iv) If an optimization problem needs to be carried out in $\Pi_m C$, for some m, for example in an identification experiment, as described in chapter 1, then the knowledge of the boundary of $\Pi_m C$ is required and it can be defined in terms of the vertices.

3.4 Examples of Unit Covariances

(i) A Markov Process

This example is due to B. Mcmillan.

Let $(X_n)_{n \in \mathbb{N}}$ be a two-state Markov process with $X_n \in \{-1,1\}$, $\forall n$, and

$$p(X_n | X_{n-1}) = \begin{cases} p \ , \text{ if } X_n \neq X_{n-1} \\ (1-p) \ , \text{ if } X_n = X_{n-1} \end{cases}$$

$$P(X_n=1) = 1/2 \ .$$

Then $(X_n)_{n \in \mathbb{N}}$ is a stationary process and then we write $EX_mX_{m+n}=\rho_n$.

Now we compute the covariance function of this process.

Assume $n>0$.

$\rho_n = E(X_oX_n) = p(X_oX_n=1) - p(X_oX_n=-1)$

$p(X_oX_n=1)$ = prob. of the trajectories $(x_o,\ldots, x_n)$ with an even number of jumps, that is

$$p(X_oX_n=1) = \sum_{\substack{l=o \\ (l \text{ even})}}^{\mu} \binom{n}{l} p^l(1-p)^{n-l} ,$$

where $\mu = \begin{cases} n, \text{ if } n \text{ is even} \\ n-1, \text{ if } n \text{ is odd} \end{cases}$

Similarly for $p(X_oX_n=-1)$. Therefore

$$\rho_n = \sum_{k=o}^{n} \binom{n}{k} (-p)^k (1-p)^{n-k} .$$

If we assume $n \in \mathbb{Z}$ then we can write $\rho_n = \alpha^{|n|}$, where $\alpha=(1-2p)$.

In figure 11 we plot the locus of (ρ_1,ρ_2), $-1 \leqslant \alpha \leqslant 1$ together with $\Pi_3 U$.

(ii) The Arc-Sine Law

Let $(X_n)_{n \in \mathbb{N}}$ be a Gaussian process with covariance function $(\rho_n)_{n \in \mathbb{N}_0}$. Define a new process $(X_n^*)_{n \in \mathbb{N}}$ by

$$X_n^* = \begin{cases} 1 \text{ , if } x_n > 0 \\ -1 \text{ , if } x_n \leqslant 0 \end{cases}$$

X^* is called a 'Clipped Gaussian Process' and its covariance $(\rho_n^*)_{n \in \mathbb{N}_0}$ is given by the well known arc-sine law:

$$\rho_n^* = \frac{2}{\pi} \text{ arc } \sin \frac{\rho_n}{\rho_0}$$

It is well known that the class of covariances of gaussian distributed weakly stationary processes with $EX_n^2=1$, $EX_n=0$ for all n is equal to the class of positive semi-definite sequences $(\gamma_n)_{n \in \mathbb{N}_0}$, with $\gamma_0=1$.

The Bochner-Herglotz representation theorem states that for each such sequence there exists a unique F such that

$$\gamma_n = \int_0^1 \cos 2\pi n\lambda \, dF(\lambda) \quad \text{for all } n.$$

This means that the extreme elements of this class are of the form $(\cos 2\pi n\lambda)_{n \in \mathbb{N}_0}$, $0 \leqslant \lambda \leqslant 1$.

For example the parametric equations of the boundary of the two-dimensional projections of the

elements of this class are

$$x_1 = \cos 2\pi\lambda$$
$$x_2 = \cos 4\pi\lambda \quad , \ 0 \leqslant \lambda \leqslant 1.$$

This parabola is also depicted in figure 11.

Of interest is the fact that the locus of the points $P(\lambda) \triangleq (p_1(\lambda), p_2(\lambda))$, where

$$p_1(\lambda) = \frac{2}{\pi} \text{ arc sin } (\cos 2\pi n\lambda)$$
$$p_2(\lambda) = \frac{2}{\pi} \text{ arc sin } (\cos 4\pi n\lambda) \quad , \ 0 \leqslant \lambda \leqslant \tfrac{1}{2}$$

is the boundary of $\Pi_3 TU$. Furthermore we have that

$$P(0) = (1,1)$$
$$P(1/4) = (0,-1)$$
$$P(1/2) = (-1,1)$$

that is, the vertices of $\Pi_3 TU$ occur for $\lambda \in \{0, 1/4, 1/2\}$.

This naturally gives rise to the question: Are the elements of TU clipped gaussian processes?

For $\Pi_4 C$ we still have that the $2/\pi$ arc sin image of the line L defined by

$$L \equiv \begin{cases} x_1 = \cos 2\pi\lambda \\ x_2 = \cos 4\pi\lambda \\ x_3 = \cos 6\pi\lambda \end{cases} \quad , \ 0 \leqslant \lambda \leqslant \tfrac{1}{2}.$$

is a polygonal curve whose vertices are the vertices

of $\Pi_4 C$. The vertices occur for values of $\lambda \in \{0, 1/6, 1/4, 1/3, 1/2\}$. But this does not show that $\Pi_4 C$ is the image of the convex hull of L because $2/\Pi$ arc sin (.) is a non-linear function.

In fact the following counter-example shows that the image of L by $2/\Pi$ arc sin (.) is strictly contained in $\Pi_4 C$.

Consider the following point on the boundary of $\Pi_4 C$:

$$1/3(1,1,1) + 2/3(0,-1,0) = (1/3,-1/3,1/3).$$

Its inverse image is

$$(\sin \Pi/6, -\sin \Pi/6, \sin \Pi/6) = (.5,-.5,.5).$$

But

$$\det \begin{bmatrix} 1 & .5 & -.5 & .5 \\ & 1 & .5 & -.5 \\ & & 1 & .5 \\ & & & 1 \end{bmatrix} = -1.6875$$

which means that there is no positive semi-definite sequence $(\rho_n)_{n \in \mathbb{N}_0}$ with

$$\rho_0 = 1$$
$$\rho_1 = .5$$
$$\rho_2 = -.5$$
$$\rho_3 = .5$$

Therefore we have that the class of clipped gaussian processes is strictly contained in the class of stationary unit covariances.

References

[1] E. Masry, On Covariance Functions of Unit Processes, SIAM J. Appl. Math., Vol. 23, No. 1, July 1972.

[2] B. Mcmillan, History of a Problem, J. Soc. Indust. Appl. Math., Vol. 3, No. 3, Sept. 1955.

[3] H. L. Royden, Real Analysis, Macmillan.

[4] L. A. Shepp, Covariances of Unit Processes, Proc. of Working Conference on Stochastic Processes, Santa Barbara, California, 1967.

[5] Stoer and Witzgall, Convexity and Optimization in Finite Dimensions, Vol. I, Springer-Verlag.

CHAPTER 4

ALGORITHMS FOR THE GENERATION OF BINARY SEQUENCES WITH PRESCRIBED AUTOCORRELATIONS

4.1 Introduction

In this chapter we shall look at the following problem: Given a vector of autocorrelation values $(a_0, a_1, \ldots, a_{m-1})$ in ${}_{\mathbb{R}}C_m$ we wish to construct a binary sequence $(s_n)_{n \in \mathbb{N}_0}$ such that

$$a_k = \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N} s_i \, s_{i+k} \quad , \quad 0 \leqslant k \leqslant m-1 .$$

Basically two different types of algorithms will be presented which appear to be original.

The question of generating binary sequences with prescribed autocorrelations was previously unsolved, although it arises very often in engineering practice; in the introduction to this thesis we have already seen an example where such a situation occurs.

A problem related to this is discussed, in terms of the frequency domain, in a paper by van den Bos [1] where he presents a procedure to construct binary periodic signals with their power distributed over

selected frequencies in a prescribed way. He uses a trial and error procedure, but no convergence proof is given for this method, nor is a characterization for the class of feasible spectral distributions.

When we started looking into this problem it seemed to us quite natural to solve it by means of the theory of shift register sequences or linear recurrent sequences over GF(2), that is the finite field with two elements. Although this approach would restrict us to periodic sequences, this is no obstacle given that the class of autocorrelation functions of these sequences is dense in C. Presented this way it is an inverse problem: One wishes to find the order and the coefficients of an equation in order that it has at least one solution with a prescribed property. In order to solve this problem we need to know:

(i) If there is any linear recurrence with such a property.

(ii) If the answer to (i) is yes, how can we determine the order and the coefficients of that recurrence, without having to go through all possible combinations?

The inverse problem seems very difficult to answer in an algebraic framework and, as far as we know, it has not been yet solved.

What is usually done is to solve a direct problem: Given a certain class of polynomials over GF(k) one studies the autocorrelation properties of the sequences generated by the elements of that class. Such type of approach has been very fruitful in some areas. For example N. Zierler [2] and S. W. Golomb [3] have discovered the so called 'pseudo random binary noise' that is periodic binary sequences with a constant out of phase autocorrelation equal to $-1/p$ where p denotes the period of the sequence, when studying the properties of the sequences associated with primitive polynomials over a finite field. This became perhaps the most important and well known subclass of binary sequences. We can also mention another study on these lines of thought made by J. Lee [4] and [5] on the autocorrelation properties of the sequences associated with non-primitive irreducible polynomials over GF(2) which have direct application as codes. However an attempt along these lines would only provide a partial answer to our problem.

Our experience has shown that the theoretical problems involved with the generation of unit covariances are somewhat easier than the ones concerning the realization of autocorrelation functions. One may think then of generating unit processes with prescribed covariances, because we have already seen that the classes C and U are equivalent, and then use their realizations

as the desired binary sequences. However this is not quite so, because the prescribed covariance might give rise to a non-ergodic process, which means that the time averages do not equal the ensemble averages. Take for example the unit covariance $(\rho_n)_{n \in \mathbb{N}_0}$ such that:

$$\rho_n = \begin{cases} 0, & n \text{ odd} \\ 1, & n \text{ even} \end{cases} .$$

There are only four possible realizations of a unit process with this covariance, as follows:

$\gamma_1 = 1, 1, 1, \ldots$ with period 1
$\gamma_2 = -1, -1, -1, \ldots$ " " 1
$\gamma_3 = 1, -1, 1, -1, \ldots$ " " 2
$\gamma_4 = -1, 1, -1, 1, \ldots$ " " 2

and all with probability 1/4. However the 'autocorrelations' of these realizations, that is, their time averages $(\rho_n^i)_{n \in \mathbb{N}_0}$ i=1, 2, 3, 4 are

$$\rho_n^1 = \rho_n^2 = 1 \text{ , for all n.}$$

$$\rho_n^3 = \rho_n^4 = \begin{cases} 1, & n \text{ even} \\ -1, & n \text{ odd.} \end{cases}$$

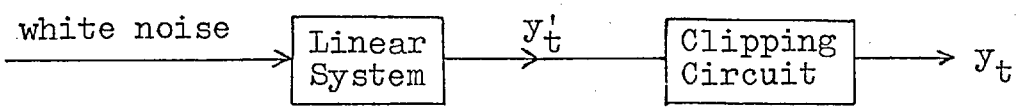
This means that there is nc realization of the above process with autocorrelation equal to its

covariance. This is sometimes overlooked in the literature and can lead to wrong conclusions. That is the case of a method based on the so called arc sine law and a factorization theorem that is often suggested for the generation of binary sequences with prescribed autocorrelations. It is as follows:

Given the desired autocorrelation function $(\rho_n)_{n \in \mathbb{N}_0}$, another function $(\rho'_n)_{n \in \mathbb{N}_0}$ is constructed by setting $\rho'_n = \sin \pi \rho_n/2$. Then if $(\rho'_n)_{n \in \mathbb{N}_0}$ is a positive semi--definite sequence all it is required is a normal stochastic stationary process with covariance $(\rho'_n)_{n \in \mathbb{N}_0}$. Such a process can be obtained as the output of a linear dynamical system driven by Gaussian white noise provided $(\rho'_n)_{n \in \mathbb{N}_0}$ gives rise to a rational spectral density in e^{iw}. See for example Astrom [6]. The spectral density $\phi'(w)$ of a process with covariance $(\rho'_n)_{n \in \mathbb{N}_0}$ is defined as

$$\phi'(w) = \frac{1}{2\pi} \sum_{n=-\infty}^{+\infty} e^{-inw} \rho'_n \quad .$$

This procedure is summarized in the diagram below:



The output of the clipping circuit is defined as

$$y_t = \begin{cases} 1 \text{ , if } y'_t \geqslant 0 \\ -1 \text{ , if } y'_t < 0 \end{cases}$$

and its covariance $(\rho_n)_{n \in \mathbb{N}_0}$ is related with the input covariance $(\rho'_n)_{n \in \mathbb{N}_0}$ by

$$\rho_n = \frac{2}{\pi} \text{ arc sin } \frac{\rho'_n}{\rho'_0}$$

as we have seen in chapter 3.

From what has been said above we see that this procedure is only suitable to realize unit covariance functions and not autocorrelations of binary sequences. Furthermore it requires $(\rho'_n)_{n \in \mathbb{N}_0}$ to be a positive semi-definite sequence where $\rho'_n = \sin \pi\rho_n/2$ and $(\rho_n)_{n \in \mathbb{N}_0}$ is the prescribed autocorrelation. As we have seen in one of the examples given in chapter 3, this is not always true and therefore only a subclass of unit covariances can be realized by this method.

We now propose a convergent algorithm for the generation of binary sequences with prescribed autocorrelations.

4.2 Type 1 Algorithm

This method takes advantage of the convexity and compactness of $\Pi_m C$ and the properties of its vertices. It is shown that the prescribed autocorrelation can be realized by a suitable interspersion of the binary sequences whose

autocorrelations are vertices of $\Pi_m C$.

We know that

$$(a_0, a_1, \ldots, a_{m-1}) = \sum_{i=1}^{N} \alpha_i (\rho_0^i, \rho_1^i, \ldots, \rho_{m-1}^i), \qquad (*)$$

with $\alpha_i \geqslant 0$, $\sum_{i=1}^{N} \alpha_i = 1$, $N \leqslant m+1$, and that there exist periodic binary sequences $(s_n^i)_{n \in \mathbb{N}_0}$ with autocorrelations $(\rho_n^i)_{n \in \mathbb{N}_0}$ and period $p_i \leqslant 2^{m-1}$, $i=1, 2, \ldots, N$.

For the moment assume that the coefficients of the above convex combination are rational numbers. Then we can write $\alpha_i = p_i/q_i$, where p_i and q_i are positive integers. Denote by M the least common multiple of $q_1, \ldots, q_N$.

Set $\alpha_i = \bar{\alpha}_i/M$. Define a new sequence $(s_n)_{n \in \mathbb{N}_0}$ by interspersing the sequences $(s_n^i)_{n \in \mathbb{N}_0}$, setting for $j=1,2,\ldots$

$$s_n = \begin{cases} s^1_{n-(M-\bar{\alpha}_1)L/M} & , \ L \leqslant n \leqslant L+\bar{\alpha}_1 j \\ s^2_{n-(M-\bar{\alpha}_2)L/M-\bar{\alpha}_1 j} & , \ L+\bar{\alpha}_1 j < n \leqslant L+(\bar{\alpha}_1+\bar{\alpha}_2) j \\ \vdots & \\ s^N_{n-(M-\bar{\alpha}_N)L/M-j(\bar{\alpha}_1+\ldots+\bar{\alpha}_{N-1})} & , \ L+j\sum_{i=1}^{N-1}\bar{\alpha}_i < n \leqslant U \end{cases} \qquad (**)$$

where $L=Mj(j-1)/2$ and $U=Mj(j+1)/2$. Recall that $\sum_{i=1}^{j} i = j(j+1)/2$.

Now we show that the sequence so constructed has the desired autocorrelation.

Take $T \in \mathbb{N}$ such that

$$Mj(j-1)/2 + j \sum_{i=1}^{k-1} \bar{\alpha}_i < T \leqslant Mj(j-1)/2 + j \sum_{i=1}^{k} \bar{\alpha}_i$$

Then we can write

$$\frac{1}{T} \sum_{n=1}^{T} s_n s_{n+t} = \frac{1}{T} \left\{ \sum_{n=1}^{\bar{\alpha}_1 j(j+1)/2} s_n^1 s_{n+t}^1 + \cdots \right.$$

$$\cdots + \sum_{n=1}^{\bar{\alpha}_{k-1} j(j+1)/2} s_n^{k-1} s_{n+t}^{k-1} + \sum_{n=1}^{\bar{\alpha}_k j(j-1)/2 + A} s_n^k s_{n+t}^k +$$

$$\left. + \sum_{n=1}^{\bar{\alpha}_{k+1} j(j-1)/2} s_n^{k+1} s_{n+t}^{k+1} + \cdots + \sum_{n=1}^{\bar{\alpha}_N j(j-1)/2} s_n^N s_{n+t}^N \right\} + O(1),$$

with $0 \leqslant t \leqslant m-1$, and $A = T - Mj(j-1)/2 - j \sum_{i=1}^{k-1} \bar{\alpha}_i$.

In $O(1)$ we include the effect of cross product terms of the form $s_n^i s_k^l$, $i \neq l$.

For a given j (and T as above) $|O(1)|$ is bounded by $2mNj/T \cong 2jm^2/T$ which goes to zero at the rate of $1/j$, when $j \to \infty$, because T grows at the rate of j^2.

Since $\bar{\alpha}_i j(j+1)/2T \xrightarrow[T \to \infty]{} \bar{\alpha}_i/M$, and

$$\bar{\alpha}_i (j-1)j/2T \xrightarrow[T \to \infty]{} \bar{\alpha}_i/M ,$$

with $1 \leqslant i \leqslant N$ and T as above, we have that

$$\lim_{T \to \infty} \frac{1}{T} \sum_{n=1}^{T} s_n s_{n+t} = \frac{\bar{\alpha}_1}{M} \rho_t^1 + \ldots + \frac{\bar{\alpha}_N}{M} \rho_t^N = \alpha_1 \rho_t^1 + \ldots + \alpha_N \rho_t^N$$

which means that $(s_n)_{n \in \mathbb{N}_0}$ has the desired autocorrelation.

q.e.d.

So far we have assumed that the coefficients α_i, $i=1, 2, \ldots, N$, in the above expression (*) were rational numbers. For the case where some of them are irrational the problem can be solved by approximating the irrational coefficients by rational ones, the degree of approximation depending on the desired accuracy, and then by applying the above method with these coefficients.

Although this method is a feasible procedure for practical implementation algorithms of the next type have the advantage of being easier to implement, particularly on a computer and of not making use of any *a priori* knowledge of the structure of $\Pi_m C$. They are iterative procedures where k consecutive terms of the desired sequence are selected at each iteration, with k being some fixed integer. Furthermore, loosely speaking, they are constructed in a way that converges as fast as possible. Convergence in the method above might be rather slow. Think for example of situations where there is at least one $\alpha_i = p_i/q_i$ with p_i and q_i 'very large'.

4.3 Type 2 Algorithms

In this section we shall discuss iterative procedures for constructing binary sequences with prescribed autocorrelations. For didactic reasons we shall discuss first a procedure convergent in a subset of $\Pi_3 C$, and then show how it can be generalized not only to the entire class $\Pi_3 C$ but also to $\Pi_m C$, for any $m \in \mathbb{N}$.

4.3.1 An algorithm for generating binary sequences with prescribed autocorrelations that is convergent on a subset of $\Pi_3 C$

We need some preliminary considerations before we proceed.

Given a binary sequence $(s_n)_{n \in \mathbb{N}_0}$ its autocorrelation function $(\rho_k)_{k \in \mathbb{N}_0}$ has been defined as

$$\rho_k = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N} s_n s_{n+k} \quad .$$

Define

$$c_T^k = \frac{1}{T} \sum_{n=0}^{T-k} s_n s_{n+k} \quad , \quad k, T \in \mathbb{N}.$$

Then we can write

$$\underline{c}_{n+1} = \frac{n}{n+1} \underline{c}_n + \frac{1}{n+1} \underline{u}_{n+1}$$

where $\underline{c}_n \equiv (c_n^1, c_n^2)$ and $\underline{u}_n \equiv (u_n, v_n) \equiv (s_n s_{n-1}, s_n s_{n-2})$. We note the sequence $(\underline{c}_n)_{n \in \mathbb{N}}$ can be specified in terms of $(u_n)_{n \in \mathbb{N}}$, because $v_n = u_n u_{n-1}$.

Define

$$W_{m,n} = \frac{1}{m} \sum_{i=1}^{m} \underline{u}_{n+i} \equiv \left(\frac{1}{m} \sum_{i=1}^{m} u_{n+i}, \frac{1}{m} \sum_{i=1}^{m} v_{n+i} \right).$$

With the substitution $v_n = u_n u_{n-1}$, $W_{m,n}$ can be regarded as a function of $u_n, \ldots, u_{n+m}$.

Let $W_{m,n}^+$ denote the restriction of $W_{m,n}$ to $u_n = +1$ and $W_{m,n}^-$ its restriction to $u_n = -1$.

Obviously the range of $W_{m,n}^+$ is the same as that of $W_{m,1}^+$ and we denote it by D_m^+. Define D_m^- similarly.

The appearance of D_1^+, D_2^+, D_3^+ and D_4^+ is shown in figures 1, 2, 3 and 4. We have used the fact that $v_{n+1} = u_n u_{n+1}$ to generate these figures. Then the rule is: If step n is to the right ($u_n = 1$) step (n+1) is NE or SW. If step n is to the left ($u_n = -1$) then step (n+1) is NW or SE. It also follows that D_m^- is the reflection of D_m^+ about the vertical axis.

Let W_m^+ denote the convex hull of D_m^+ and W_m^- that of D_m^-. Let W_m^i denote the intersection of W_m^+ with W_m^-. The appearance of W_2^i, W_3^i and W_4^i is depicted in figures 5, 6 and 7.

From the definition of $W_{m,n}$ it follows that the convex hull of its range tends to $\Pi_3 C$ as m goes to ∞, because the vertices of $\Pi_3 C$ are autocorrelations of

periodic binary sequences of period not greater than 4. Therefore W_m^i also tends to $\Pi_3 C$ as m tends to ∞.

We now establish the algorithm.

Suppose $\underline{\rho} \equiv (\rho_1, \rho_2)$ is the prescribed vector of autocorrelations. Let $(\underline{c}_{2n})_{n \in \mathbb{N}}$ be the sequence generated by the expression

$$\underline{c}_{2n+2} = \frac{n}{n+1} \underline{c}_{2n} + \frac{1}{2n+2} (\underline{u}_{2n+1} + \underline{u}_{2n+2})$$

where u_{2n+1} and u_{2n+2} are chosen to minimize the Euclidean norm of $(\underline{c}_{2n+2} - \underline{\rho})$, given that the values of $\underline{c}_{2n}$ and u_{2n} have been previously determined.

Proposition

If $\underline{\rho} \in W_2^i$ then $(\underline{c}_n)_{n \in \mathbb{N}}$ converges to $\underline{\rho}$, for all initial values, at the rate of $O(n^{-\frac{1}{2}})$.

In the proof of the proposition we shall make use of the following lemma:

Lemma

Consider the triangle $\{a,b,c\}$ and a point d on side (b,c). Then $\min.(\|c-a\|^2, \|b-a\|^2) \leqslant \|d-a\|^2 + \|b-c\|^2$.

Proof

$$\|b-a\|^2 = \|d-a\|^2 + \|b-d\|^2 + 2(d-a)^T(b-d)$$

$$\|c-a\|^2 = \|d-a\|^2 + \|c-d\|^2 + 2(d-a)^T(c-d) \ .$$

But either $(d-a)^T(b-d) \leqslant 0$ or $(d-a)^T(c-d) \leqslant 0$. Since $\|c-d\| \leqslant \|b-c\|$ and $\|b-d\| \leqslant \|b-c\|$ the result follows.

We are now in a position to prove convergence for the algorithm. Since

$$\| \underline{c}_{n+1} - \underline{c}_n \| \leqslant \left(2 + \| \underline{c}_n \|\right)/(n+1)$$

we need consider only the convergence of the subsequence $(\underline{c}_{2n})_{n \in \mathbb{N}}$.

Proof

Assume that $u_{2n}=+1$ and therefore that $\underline{c}_{2n+2} \in \left(\frac{n}{n+1}\, \underline{c}_{2n} + \frac{1}{n+1}\, W_2^+\right)$. We shall denote the vector addition of sets in the usual way, that is for example $\underline{a} + B = \left\{ \underline{a} + \underline{b} : \text{all } \underline{b} \in B \right\}$. Two possibilities must be considered:

(i) $\underline{\rho} \notin \frac{n}{n+1}\, \underline{c}_{2n} + \frac{1}{n+1}\, W_2^+$

Let $\underline{x}$ denote the intersection of the line segment $(\underline{\rho}, \underline{c}_{2n})$ with the edge of $n/(n+1)\underline{c}_{2n} + 1/(n+1)W_2^+$ which is closest to $\underline{\rho}$ (see figure 8).

Let V denote the closest vertex to $\underline{\rho}$ on this edge.

Then by the lemma

$$\|V-\underline{\rho}\|^2 \leqslant \|\underline{x}-\underline{\rho}\|^2 + d^2/(n+1)^2$$

where d is the diameter of W_2^+. (Note that $d/(n+1)$ is the diameter of $n/(n+1)\underline{c}_{2n} + 1/(n+1)W_2^+$). But from the definition of $\underline{c}_{2n+2}$ we have that $\underline{c}_{2n+2}$ is either V or a point closer to $\underline{\rho}$, namely a vertex of $n/(n+1)\underline{c}_{2n} + 1/(n+1)W_2^+$ or $n/(n+1)\underline{c}_{2n}$.

Then

$$\|\underline{c}_{2n+2} - \underline{\rho}\|^2 \leqslant \|\underline{x} - \underline{\rho}\|^2 + d^2/(n+1)^2 \ .$$

Now $\|\underline{x} - \underline{\rho}\| \leqslant n/(n+1) \|\underline{c}_{2n} - \underline{\rho}\|$; this follows from the fact that $\underline{x}$ is the closest point to $\underline{\rho}$ of $n/(n+1)\underline{c}_{2n} + 1/(n+1)W_2^i$ on the line joining $\underline{\rho}$ to $\underline{c}_{2n}$ and that the point $\underline{\rho}+n/(n+1)(\underline{c}_{2n} - \underline{\rho})$ on this line lies inside $n/(n+1)\underline{c}_{2n} + 1/(n+1)W_2^i$. Therefore

$$\|\underline{c}_{2n+2} - \underline{\rho}\|^2 \leqslant n^2/(n+1)^2 \|\underline{c}_{2n} - \underline{\rho}\|^2 + d^2/(n+1)^2 \quad (*)$$

(ii) $\underline{\rho} \in n/(n+1)\underline{c}_{2n} + 1/(n+1)W_2^+$

In this case we have $\|\underline{c}_{2n+2} - \underline{\rho}\|^2 \leqslant d^2/(n+1)^2$.

The same arguments hold if $u_{2n}=-1$ and $\underline{c}_{2n+2} \in \left(n/(n+1)\underline{c}_{2n} + 1/(n+1)W_2^-\right)$; in particular inequality (*) still holds.

Therefore we have for all $\underline{c}_{2n}$ and all $\underline{\rho} \in W_2^i$

$$\begin{aligned}(n+1)^2 \|\underline{c}_{2n+2} - \underline{\rho}\|^2 &\leqslant n^2 \|\underline{c}_{2n} - \underline{\rho}\|^2 + d^2 \\ &\leqslant (n-1)^2 \|\underline{c}_{2n-2} - \underline{\rho}\|^2 + 2d^2 \\ &\leqslant \cdots \\ &\leqslant \|\underline{c}_2 - \underline{\rho}\|^2 + (n-1)d^2\end{aligned}$$

and therefore $\|\underline{c}_{2n} - \underline{\rho}\|$ converges to zero as n goes to ∞ at the rate of $O(n^{-\frac{1}{2}})$.

q.e.d.

4.3.2 A convergent algorithm to generate binary sequences with any prescribed number of autocorrelations

If the vector of prescribed autocorrelations has dimension greater than two, it is still possible to implement iterative procedures leading to a convergent solution on the same lines of thought as above.

Before proceeding we need to extend some of our previous definitions to $\mathbb{R}^k$, for k greater than two.

Let W_m , $m \in \mathbb{N}$, denote a function with domain $\{-1,1\}^{m+k}$ defined as follows:

$$W_m = \frac{1}{m} \sum_{j=1}^{m} \underline{u}_{j+k-1} \quad ,$$

where $\underline{u}_i = (s_i s_{i-1}, s_i s_{i-2}, \ldots, s_i s_{i-k})$, and $s_i \in \{-1,1\}$ for all i.

Denote by W_m^j the restriction of W_m to the jth possible value of $(s_{k-1}, s_{k-2}, \ldots, s_o)$; this vector can only assume 2^k distinct values. Denote by D_m^j the range of W_m^j and let Ω_m^j denote the convex hull of D_m^j. Finally set

$$\Omega_m = \bigcap_{j=1}^{2^k} \Omega_m^j$$

In the generalization of the algorithm of section 4.3.1 to the k-dimensional case, $k>2$, we shall make use of the following results:

Lemma 1

The convex hull of the range of W_m converges to $\Pi_{k+1}C$ as m tends to infinity.

Proof

The result follows from the fact that $\Pi_{k+1}C$ is a polytope whose vertices are autocorrelations of periodic binary sequences of periods not greater than 2^k as it has been shown in chapter 2.

Lemma 2

Consider a polytope P lying in an hyperplane in $\mathbb{R}^k$ with vertices $\underline{v}_1, \ldots, \underline{v}_p$. Let $\underline{x}$ denote a point in P and $\underline{y}$ any other point in $\mathbb{R}^k$. Assume $\min\left(\|\underline{y} - \underline{v}_1\|, \ldots, \|\underline{y} - \underline{v}_p\|\right) = \|\underline{y} - \underline{v}_{i_0}\|$. Then

$$\|\underline{y} - \underline{v}_{i_0}\|^2 \leqslant \|\underline{y} - \underline{x}\|^2 + d^2$$

, where d is the diameter of P.

Proof

All we need to show is that there exists a vertex $\underline{v}$ of P such that

$$<(\underline{v} - \underline{x}),(\underline{y} - \underline{x})> \geqslant 0.$$

If not $<(\underline{v}_i - \underline{x}),(\underline{y} - \underline{x})> < 0$ for all $\underline{v}_i$ which would imply $<(\underline{y} - \underline{x}),(\underline{y} - \underline{x})> < 0$, for all y in P, which is impossible since $\underline{x} \in P$.

Then we have that $\|\underline{v} - \underline{y}\|^2 \leqslant \|\underline{x} - \underline{y}\|^2 + \|\underline{v} - \underline{x}\|^2$ and the result follows.

Given a binary sequence $(s_n)_{n \in \mathbb{N}_0}$ define

$$c_T^k = \frac{1}{T} \sum_{n=0}^{T-k} s_n s_{n+k} \quad , \text{ for } k, T \in \mathbb{N},$$

and set $\underline{c}_m \equiv (c_m^1, \ldots, c_m^k)$.

The algorithm of section 4.3.1 is now extended to the k-dimensional case, $k > 2$.

Let $\underline{\rho} \equiv (\rho_1, \rho_2, \ldots, \rho_k)$ be the prescribed auto-correlation vector. Let $(s_n)_{n \in \mathbb{N}_0}$ be the binary sequence generated by

$$\underline{c}_{m_0}(n+1) = \frac{n}{n+1} \underline{c}_{m_0} n + \frac{1}{m_0 n + m_0} \left(\sum_{j=1}^{m_0} (s_{m_0 n+j} s_{m_0 n+j-1}, \ldots \right.$$

$$\left. \ldots, s_{m_0 n+j} s_{m_0 n+j-k}) \right)$$

where $s_{m_0 n+1}, \ldots, s_{m_0 n+m_0}$ are chosen to minimize the euclidean norm of $\underline{c}_{m_0}(n+1) - \underline{\rho}$ given that $\underline{c}_{m_0} n$ and $s_0, s_1, \ldots, s_{m_0 n}$ have been previously determined.

Proposition

If $\underline{\rho} \in \Omega_{m_0}$ than $(\underline{c}_n)_{n \in \mathbb{N}}$ converges to $\underline{\rho}$ for all initial values, at the rate of $O(n^{-\frac{1}{2}})$.

Proof

Since $\|\underline{c}_{n+1} - \underline{c}_n\| \leqslant (m_0\|\underline{c}_n\| + m_0 k)/(n+m_0)$ for $1 < m_0$, we need to consider only the convergence of the subsequence $(\underline{c}_{m_0 n})_{n \in \mathbb{N}}$.

Then assume that $(s_{m_0 n}, \ldots, s_{m_0 n+1-k})$ takes its jth possible value, $1 \leqslant j \leqslant 2^k$ and therefore that

$$\underline{c}_{m_0 n+m_0} \in \left(\frac{n}{n+1}\,\underline{c}_{m_0 n} + \frac{1}{n+1}\,\Omega^j_{m_0}\right);$$

the addition of sets is defined as in section 4.3.1.

Two possibilities must be considered:

(i) $\underline{\rho} \notin \left(\frac{n}{n+1}\,\underline{c}_{m_0 n} + \frac{1}{n+1}\,\Omega^j_{m_0}\right)$

Let $\underline{x}$ denote the intersection of the line segment $(\underline{\rho}, \underline{c}_{m_0 n})$ with the boundary of $n/(n+1)\underline{c}_{m_0 n} + 1/(n+1)\,\Omega^j_{m_0}$ which is closer to $\underline{\rho}$ and let $\underline{v}$ denote the closest vertex to $\underline{\rho}$ on the face containing $\underline{x}$. Then by lemma 2 we have that

$$\|\underline{\rho} - \underline{v}\|^2 \leqslant \|\underline{x} - \underline{\rho}\|^2 + d_j^2/(n+1)^2$$

where d_j is the diameter of $\Omega^j_{m_0}$. (Note that $d_j/(n+1)$ is the diameter of $n/(n+1)\underline{c}_{m_0 n} + 1/(n+1)\,\Omega^j_{m_0}$).

From the definition of $\underline{c}_{m_0 n}$ we have that $\underline{c}_{m_0 n+m_0}$ is either $\underline{v}$ or another point of $D^j_{m_0}$ closer to $\underline{\rho}$. Therefore

$$\|\underline{c}_{m_0 n+m_0} - \underline{\rho}\|^2 \leqslant \|\underline{x} - \underline{\rho}\|^2 + d_j^2/(n+1)^2 \quad .$$

We also have that

$\|\underline{x} - \underline{\rho}\| \leqslant \frac{n}{n+1} \|\underline{c}_{m_0 n} - \underline{\rho}\|$; this follows from the fact that $\underline{x}$ is the closest point to $\underline{\rho}$ of $n/(n+1)\underline{c}_{m_0 n} + 1/(n+1)\,\Omega^j_{m_0}$ on the line joining $\underline{\rho}$ to $\underline{c}_{m_0 n}$ and that the point $\underline{\rho} + n/(n+1)(\underline{c}_{m_0 n} - \underline{\rho})$ on this line lies inside $n/(n+1)\underline{c}_{m_0 n} + 1/(n+1)\,\Omega^j_{m_0}$, because $\underline{\rho} \in \Omega^j_{m_0}$ by assumption. Then

$$\|\underline{c}_{m_0 n+m_0} - \underline{\rho}\|^2 \leqslant n^2/(n+1)^2 \,\|\underline{c}_{m_0 n} - \underline{\rho}\|^2 + d_j^2/(n+1)^2$$

(ii) $\underline{\rho} \in \left(\frac{n}{n+1}\,\underline{c}_{m_0 n} + \frac{1}{n+1}\,\Omega^j_{m_0}\right)$

In this case we have $\|\underline{c}_{m_0 n+m_0} - \underline{\rho}\|^2 \leqslant d_j^2/(n+1)^2$. Therefore we have for all $\underline{c}_{m_0 n}$ and $\underline{\rho} \in \Omega_m$

$$\begin{aligned}(n+1)^2 \,\|\underline{c}_{m_0 n+m_0} - \underline{\rho}\|^2 &\leqslant n^2 \|\underline{c}_{m_0 n} - \underline{\rho}\|^2 + d^2 \\ &\leqslant (n-1)^2 \|\underline{c}_{m_0 (n-1)} - \underline{\rho}\|^2 + 2d^2 \\ &\vdots \\ &\leqslant \|\underline{c}_{m_0} - \underline{\rho}\|^2 + (n-1)d^2\end{aligned}$$

where $d = \max(d_1, \ldots, d_{2^k})$, and therefore $\|\underline{c}_{m_0 n} - \underline{\rho}\|$ converges to zero at the rate of $O(n^{-\frac{1}{2}})$.

q.e.d.

This algorithm can be used to generate binary sequences with the prescribed vector of autocorrelations anywhere in the <u>interior</u> of $\Pi_k C$, for any k, provided one selects m_0 such that Ω_{m_0} contains $\underline{\rho}$; by lemma 1 we have that this is always possible.

We found that in practice we can simplify this procedure by selecting only one new term of the binary sequence at each iteration without convergence being lost. However the rate of convergence decreases as the number of prescribed autocorrelations increases. This is apparent in the two numerical examples of application of this algorithm given in appendix C. In the two dimensional example the 'error' at the 100th iteration is already smaller than the error at the 450th iteration in the five dimensional case.

The simplicity of the one step ahead iterative algorithm and the satisfactory way it works in practice, particularly for small dimensions, are enough to motivate a study of its convergence properties. However such a study is difficult because the number of search directions at each iteration is limited to two and because the criterion being used in the selection of the terms of the sequence, the minimization of an euclidean norm, gives rise to a complicated behaviour. To avoid these difficulties we have decided to modify the selection criterion of the terms of the sequence

and this has enabled us to establish a first order iterative algorithm with a linear rate of convergence if the prescribed autocorrelation vector is in the interior of $\Pi_3 C$. This is done in the next section.

4.3.3 One step ahead iterative methods

An infinite number of 'n-steps ahead' iterative procedures can be obtained by changing the selection criterion of the terms of the sequence. However, they all have certain basic features in common and we shall discuss next some of them for the case where n=1.

Set $\underline{a} \equiv (1,1)$, $\underline{b} \equiv (-1,1)$, $\underline{c} \equiv (-1,-1)$, $\underline{d} \equiv (1,-1)$ and $\underline{e} \equiv \frac{1}{2}(c+d)$.

Given the first p+1 terms of a binary sequence $(s_n)_{n \in \mathbb{N}_0}$ we have already defined

$$\underline{c}_{p+1} = \frac{p}{p+1}\,\underline{c}_p + \frac{1}{p+1}\,\underline{u}_{p+1} \qquad (*)$$

where $\underline{u}_p \equiv (s_p s_{p-1}, s_p s_{p-2})$, $\underline{c}_p \equiv (c_p^1, c_p^2)$ and

$$c_p^k = \frac{1}{p} \sum_{n=0}^{p-k} s_n s_{n+k} \quad , \; k \in \{1,2\}.$$

Expression (*) can be rewritten as

$$\underline{c}_{p+1} - \underline{c}_p = \frac{1}{p+1}\,(\underline{u}_{p+1} - \underline{c}_p)$$

which means that the change $(\underline{c}_{p+1} - \underline{c}_p)$ can only take the values $\frac{1}{p+1}(\underline{a} - \underline{c}_p)$, $\frac{1}{p+1}(\underline{b} - \underline{c}_p)$, $\frac{1}{p+1}(\underline{c} - \underline{c}_p)$ or $\frac{1}{p+1}(\underline{d} - \underline{c}_p)$ (see figure 9) and that $\| \underline{c}_{p+1} - \underline{c}_p \|$ is bounded by $2\sqrt{2}/(p+1)$ if $\underline{c}_p \in H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$.

Given that at each iteration only one term of the sequence is selected, there are only two possible choices:

(i) if $s_p s_{p-1} = +1$, then $\underline{u}_{p+1}$ can only take the value $\underline{a}$ or $\underline{c}$.

(ii) if $s_p s_{p-1} = -1$, then $\underline{u}_{p+1}$ is either $\underline{b}$ or $\underline{d}$.

We also have that

$\underline{u}_p = \underline{a}$ (or $\underline{b}$) implies $\underline{u}_{p+1} \in \{\underline{a},\underline{c}\}$ (or $\in \{\underline{b},\underline{d}\}$)

$\underline{u}_p = \underline{c}$ (or $\underline{d}$) implies $\underline{u}_{p+1} \in \{\underline{b},\underline{d}\}$ (or $\in \{\underline{a},\underline{c}\}$).

This shows that the occurrences of the values $\underline{c}$ and $\underline{d}$ for $\underline{u}_p$ are interlaced: between two consecutive values of p for which $\underline{u}_p = \underline{c}$ there is exactly one value of p for which $\underline{u}_p = \underline{d}$, and vice-versa. We also have that $\underline{c}_p = \frac{1}{p}(p_a\underline{a} + p_b\underline{b} + p_c\underline{c} + p_d\underline{d})$ where p_a, p_b, p_c and p_d are the number of occurrences of the respective values $\underline{a}$, $\underline{b}$, $\underline{c}$ and $\underline{d}$ in the first p steps, and by the argument just given we have that p_c and p_d differ by at most one. Therefore we have the following result:

Proposition

The triangle $H\{\underline{a},\underline{b},\underline{e}\}$ is a domain of attraction for any 'one-step ahead' iterative algorithm.

We now give a convergent iterative algorithm for generating binary sequences with at most two prescribed autocorrelation values. As we shall see, this method has the advantage, besides its simplicity, of having a linear rate of convergence.

Denote by $\underline{\rho}$ the bi-dimensional vector of prescribed autocorrelations and by S_a and S_b the lines $(\underline{\rho},\underline{a})$ and $(\underline{\rho},\underline{b})$ respectively. These lines will act as switching lines for the sequence $(\underline{c}_n)_{n \in \mathbb{N}}$ generated by the algorithm. Let A and C denote the half planes produced by S_b containing respectively $\underline{a}$ and $\underline{c}$, and B and D denote the half planes produced by S_a containing $\underline{b}$ and $\underline{d}$ respectively.

Now the algorithm is as follows:

The binary sequence $(s_n)_{n \in \mathbb{N}_0}$ is constructed term by term. The first two terms are selected at random. When p terms have already been selected s_{p+1} is chosen so that

(i) $\underline{u}_{p+1} = \underline{a}$ if $\underline{c}_p \in C$ or $\underline{u}_{p+1} = \underline{c}$ if $\underline{c}_p \in A$ for the case where $s_p s_{p-1} = 1$

(ii) $\underline{u}_{p+1} = \underline{b}$ if $\underline{c}_p \in D$ or $\underline{u}_{p+1} = \underline{d}$ if $\underline{c}_p \in B$ for the case where $s_p s_{p-1} = -1$

This is a first order algorithm if we regard the 'state' as being $(c_p^1, c_p^2, s_p s_{p-1})$. The 'state space' can be taken to be the disjoint union of two squares of side 2 on each of which $s_p s_{p-1}$ takes a constant value.

The following results will be of use in the proof of convergence of the algorithm:

Lemma 1

An upper bound on the minimum value of m such that $\sum_{n=n_i}^{m} \frac{1}{n+a} > \lambda$, $\lambda, a \in \mathbb{R}$, is given by the smallest integer greater than $(n_i+a)e^{\lambda} - (1+a)$.

Proof

We have that

$$\sum_{i=n_i}^{m} \frac{1}{i+a} > \int_{n_i-1}^{m} \frac{1}{x+a+1}\, dx \quad , \quad \text{and}$$

$$\int_{n_i-1}^{m} \frac{1}{x+a+1}\, dx = \log \frac{m+a+1}{n_i+a}$$

The result follows.

Lemma 2

If two consecutive terms of the sequence $(\underline{c}_n)_{n \in \mathbb{N}}$, generated by this algorithm, say $\underline{c}_{n_0}$ and $\underline{c}_{n_0-1}$, lie on opposite half planes produced by one of the lines S_a or S_b and $\underline{c}_{n_0-1} \in H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$, then for all n greater than n_0 the distance from $\underline{c}_n$ to that line is not greater than $4\sqrt{2}/n_0$, provided $\underline{\rho}$ is an interior point of $H\{\underline{a},\underline{b},\underline{e}\}$.

Proof

We shall prove the result for line S_b. By symmetry the same result holds for S_a. Please refer to figure 10.

From the definition of $(\underline{c}_p)_{p \in \mathbb{N}}$ we have that if $\underline{c}_{n_0} \in H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$ then $\underline{c}_n$ is also in $H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$ for all n greater than n_0. We also know that

$\|\underline{c}_{p+1} - \underline{c}_p\| = \|\underline{u}_{p+1} - \underline{c}_p\|/(p+1)$ where $\underline{u}_p \in \{\underline{a},\underline{b},\underline{c},\underline{d}\}$ for all p. Therefore if $\underline{c}_p \in H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$ then

$\|\underline{c}_{p+1} - \underline{c}_p\| \leqslant 2\sqrt{2}/(p+1)$ since $2\sqrt{2}$ is the diameter of $H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$.

First let us assume that S_b is crossed from half plane A into half plane C at stage n_0, that is $\underline{c}_{n_0-1}$ is in A and c_{n_0} is in C. Then the distance from $\underline{c}_{n_0}$ to S_b is less than $2\sqrt{2}/n_0$. But from the definition of the algorithm it follows that $\underline{u}_{n+1}$ does not take the value $\underline{c}$ while $\underline{c}_n$ is in half plane C and takes at most once the value $\underline{d}$; in this case $\underline{c}_n$ may be driven away from S_b but by not more than $4\sqrt{2}/n_0$, if $\underline{d}$ lies in half plane C. However the occurrence of the values $\underline{a}$ or $\underline{b}$ for $\underline{u}_n$ bring $\underline{c}_n$ closer to S_b.

Now assume that S_b is crossed at stage n_0 from C into A. The only possibility that needs to be discussed is when $\underline{d}$ is in half plane A and $\underline{c}_{n_0}$ falls in AB. While $\underline{c}_n$ remains in AB, $\underline{u}_n$ takes the values $\underline{c}$ and $\underline{d}$ alternatively and from the corollary in appendix D we have that

the associated $\underline{c}_n$'s will lie in two lines intersecting at the point $\underline{e}=\frac{1}{2}(\underline{c}+\underline{d})$. This situation is illustrated in figure 10. Therefore if $\underline{\rho} \in H\{\underline{a},\underline{b},\underline{e}\}$ the distance from $\underline{c}_n$ to S_b, for all $n \geqslant n_0$ and $\underline{c}_n$ in A will be at most $2\sqrt{2}/n_0$. This ends the proof of lemma 2.

In the discussion of convergence of this algorithm it will be convenient to adopt the following definition of neighbourhood:

Definition

Given a positive real number δ, we define a δ-neighbourhood of $\underline{\rho}$ as the set of points within a distance of less than δ from lines S_a and S_b.

We are now in a position to prove convergence for the algorithm. To simplify matters suppose that $\underline{c}_1$ is an interior point of $H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$. This ensures that in all the subsequent iterations $\underline{c}_n$ is also inside the square $H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$ and therefore less than $2\sqrt{2}$ away from any of its vertices. Although the case in which $\underline{c}_1$ is outside $H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$ is uninteresting, it is easy to show that, for all $\underline{c}_1$, $\underline{c}_n$ eventually falls in $H\{\underline{a},\underline{b},\underline{c},\underline{d}\}$ provided the prescribed auto-correlation is neither $\underline{a}$ nor $\underline{b}$.

Proposition

Suppose $\underline{\rho}$ is an interior point of H $\{\underline{a},\underline{b},\underline{e}\}$ and $\underline{c}_1$ is an interior point of the square H $\{\underline{a},\underline{b},\underline{c},\underline{d}\}$. Given $\delta > 0$ the sequence $(\underline{c}_n)_{n \in \mathbb{N}}$ generated by the algorithm remains in a δ-neighbourhood of $\underline{\rho}$ for all n greater than $k_1/\delta + k_2$, where k_1 and k_2 are constants depending only on $\underline{\rho}$.

Proof

Take $n_0 \in \mathbb{N}$ such that $n_0 > 4\sqrt{2}/\delta$. We now study the behaviour of the algorithm after iteration n_0. Recall that we have the square H $\{\underline{a},\underline{b},\underline{c},\underline{d}\}$ divided in four regions namely AB, AD, BC and CD.

While $\underline{c}_n$, $n>n_0$, remains in the same switching region as $\underline{c}_{n_0}$ three possibilities must be considered:

(i) $\underline{u}_{n+1}$ constantly equal to $\underline{a}$

(ii) $\underline{u}_{n+1}$ constantly equal to $\underline{b}$

(iii) $\underline{u}_{n+1}$ assuming the values $\underline{c}$ and $\underline{d}$ alternatively.

From appendix D it follows that in (i) and (ii) the points $\underline{c}_n$, $n>n_0$, lie on a straight line containing $\underline{a}$ or $\underline{b}$, respectively, and that in (iii) the iterations of the form $\underline{c}_{2n}$, $n>n_0/2$ lie in a straight line containing vertex $\underline{e}$. Furthermore it follows from the definition of the algorithm that the region containing $\underline{c}_{n_0}$ does not contain the vertex which the subsequent $\underline{c}_n$'s are heading

to, provided $\underline{\rho}$ is neither $\underline{a}$ nor $\underline{b}$ nor $\underline{e}$, which is true by assumption. Then we have that there exists a positive ε_1,depending only on $\underline{\rho}$ such that

$$\|\underline{c}_{n+2} - \underline{c}_n\| > \varepsilon_1/(n+2) \quad , \; n>n_0 .$$

But this implies that at least one of the lines S_a or S_b will be crossed at some later stage, say n_1, because the series $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent. From lemma 1 we have

$$\sum_{n=1}^{m} \frac{\varepsilon_1}{2n+n_0} > 2\sqrt{2} \quad \text{if} \quad m > (1 + \frac{n_0}{2})\, e^{4\sqrt{2}/\varepsilon_1} - (1 + \frac{n_0}{2})$$

and therefore n_1 will be less than $(2+n_0)\, e^{4\sqrt{2}/\varepsilon_1} - 2$.

In case (i) we have that S_b was crossed at stage n_1, and therefore $\underline{c}_{n_1}$ falls either in AB or AD. If $\underline{c}_{n_1}$ falls in AD then $\underline{u}_{n_1+1} = \underline{c}$ and then $\underline{u}_n$, $n>n_1+1$, will remain constantly equal to $\underline{b}$; then by similar arguments we have that S_a will be crossed at a stage not later than $(2+n_1)\, e^{4\sqrt{2}/\varepsilon_2} - 2$ where ε_2 plays a role similar to ε_1. If $\underline{c}_{n_1}$ falls in AB then $\underline{u}_n$, $n>n_1$, takes the values $\underline{c}$ and $\underline{d}$ alternatively until AB is left either by crossing S_a or S_b. A similar discussion applies for case (ii).

While $\underline{c}_n$ remains in half plane B the values $\underline{c}$, $\underline{d}$ and $\underline{a}$ occur for $\underline{u}_n$ in the following sequence: $\ldots,\underline{c},\underline{d},\underline{c},\underline{d},\underline{a},\underline{a},\ldots,\underline{a},\underline{c},\underline{d},\underline{c},\underline{d},\ldots,\underline{c},\underline{d},\underline{a},\ldots,\underline{a},\underline{c},\underline{d},\ldots$ the values

of $\underline{a}$ occurring when $\underline{c}_n$ falls in region BC.

We also know that $(\underline{c}_{n+1} - \underline{c}_n) = (\underline{a} - \underline{c}_n)/(n+1)$ if $\underline{u}_{n+1} = \underline{a}$ and $(\underline{c}_{n+2} - \underline{c}_n) = 2(\underline{e} - \underline{c}_n)/(2+n)$ if $\underline{u}_{n+1} = \underline{c}$ and $\underline{u}_{n+2} = \underline{d}$.

Since $\underline{\rho}$ is an interior point of $H\{\underline{a},\underline{b},\underline{e}\}$, there exists a positive ε_3, depending only on $\underline{\rho}$ such that the projections of $(\underline{e} - \underline{c}_n)$ and $(\underline{a} - \underline{c}_n)$ on S_b along edge $(\underline{a},\underline{e})$ will be greater than ε_3.

By the same reasoning as above we have that S_a will then be crossed at a stage not later than $(2+n_1)\ e^{4\sqrt{2}/\varepsilon_3} - 2$.

It then follows that at an iteration not later than $n_0 e^{8\sqrt{2}/\varepsilon} + 2(e^{4\sqrt{2}/\varepsilon} - 1)$, $\varepsilon = \min(\varepsilon_1, \varepsilon_2, \varepsilon_3)$, $\underline{c}_n$ will have crossed lines S_a and S_b, and by lemma 2 it will remain inside a neighbourhood of $\underline{\rho}$ of size $4\sqrt{2}/n_0$.

This ends the proof of proposition.

Comments

(i) We have assumed in this proposition that $\underline{\rho}$ was an interior point of $H\{\underline{a},\underline{b},\underline{e}\}$. However it can be shown on similar lines that the algorithm still converges if $\underline{\rho}$ lies on the edges $(\underline{a},\underline{e})$ or $(\underline{b},\underline{e})$.

(ii) In contrast to the general multi-step method of section 4.3.2 , the rate of convergence is asymptoticaly $O(1/n)$ rather than $O(1/n^{\frac{1}{2}})$.

(iii) The rate of convergence of this algorithm is global in the sense that it holds for all δ and not just for δ sufficiently small. It has also an obvious 'inverse form':

$$\| c_n - \rho \| \leqslant \min \left\{ 2\sqrt{2}, \frac{k_1}{|n - k_2|} \right\}.$$

References

[1] A. van den Bos, Construction of Binary Multi-frequency Test Signals, IFAC Symposium on Identification, Prague, Czechoslovakia, 1967.

[2] N. Zierler, Linear Recurring Sequences, J. Soc. Indust. Appl. Math., Vol. 7, No. 1, March 1959.

[3] S. W. Golomb and Others, Digital Communications with Space Applications, Prentice-Hall, 1964.

[4] J. J. Lee, Digital Sequences with Special Correlation Properties, Ph.D. Thesis, State University of New York at Stony Brook, 1971.

[5] J. J. Lee and D. R. Smith, Families of Shift Regiter Sequences with Impulsive Correlation Properties, IEEE Transactions on Information Theory, Vol. IT-20, March 1974.

[6] K. J. Åström, Introduction to Stochastic Control, Academic Press.

CHAPTER 5

FURTHER RESEARCH

In this thesis a study of the autocorrelation and covariance properties of deterministic and stochastic binary sequences has been made within a geometric framework.

The inverse problem of generating binary sequences with prescribed autocorrelations has also been treated.

It would be of interest to extend these results to the more general case of generating sets of sequences not only with prescribed autocorrelations but also with cross-correlations. We have already pointed out in the introduction to the thesis that the input signals have a significant bearing upon the achievable accuracy in identification experiments, and that for single input-single output linear systems the optimal input is most conveniently characterized by its autocorrelation matrix. However if the system under consideration has several inputs it can be shown that the 'optimal' input signals are not only characterized by their autocorrelations but also by their crosscorrelations. We also believe that the generation of signals with prescribed autocorrelations and crosscorrelations will have important applications in areas like telecommunications.

The extension of the characterization of the class of autocorrelation functions to the class of cross-correlation functions appears to offer few problems but the important question of determining the set of feasible crosscorrelations for a given set of fixed autocorrelation shifts does not seem so clear.

Also of interest is the determination of the extreme elements of the class of covariance functions of continuous time unit processes. It has been shown that such a class is convex and compact. Therefore any element in this class can be represented as a convex combination of extreme elements. The question of uniqueness of such a representation deserves further investigation, but judging from L. Shepp's work it is highly unlikely that this unique representation exists.

APPENDIX A

An algorithm is given to generate a set of points whose convex hull is $\Pi_m C$, for any value of m.

Algorithm

Step 1 Select the desired value of m and define $X_{l,k} \triangleq x_{l,k}(x_{l,k}, \ldots, x_{l,k+m-2})$, $l,k \in \mathbb{N}$. Set $i=1$ and $x_{i,l}=1$, $l=1,\ldots,m-1$. Go to step 3.

Step 2 Construct X(i,1) such that $X(i,1) \neq X(k,1)$, $1 \leqslant k < i$.

Step 3 Set n=m

Step 4 Set $j=n-m+2$ and $x_{i,n}=x_{i,j}$

Step 5 Compare X(i,j) with X(i,k), $1 \leqslant k < j$. There are two possible cases:

5.1 $X(i,j) \neq X(i,k)$, $1 \leqslant k < j$. Then set $n=n+1$ and go to step 4.

5.2 $X(i,j) = X(i,k_0)$, for some k_0, $1 \leqslant k_0 < j$. Two situations must again be considered

5.2.1 If $k_0 > 1$ go to step 6.

5.2.2 If $k_0=1$ the vector $\frac{1}{j-1}\sum_{k=1}^{j-1} x_{i,k}\,(x_{i,k},\ldots,\ldots,x_{i,k+m-1})$ satisfies the given necessary conditions in lemma 1 (section 2.5) to be a vertex of $\Pi_m C$.

Comment: The components of this vector are also the first m autocorrelation values of the periodic binary sequence, of period j-1, with j-1 consecutive terms equal to $(x_{i,1},\ldots, x_{i,j-1})$ or of the periodic sequence, of period 2(j-1), with 2(j-1) consecutive terms equal to $(x_{i,1},\ldots, x_{i,j-1}, -x_{i,1},\ldots, -x_{i,j-1})$ depending on $x_{i,1}$ being equal to or the negative of $x_{i,j}$.

Step 6 Two situations are possible

6.1 If $x_{i,j}=x_{i,n}$ set $x_{i,n}=-x_{i,j}$, and go to step 5.

6.2 Otherwise look for the greatest $l, m\leqslant l<n$ such that $x_{i,l}=x_{i,l-m+2}$. Two cases must again be considered:

6.2.1 If there is no such l go to step 7.

6.2.2 If such an l exists set n=l, $x_{i,n}=-x_{i,j}$ and go to step 5.

Step 7 Set i=i+1. If $i>2^{m-2}$ stop; otherwise go to step 2.

Comment: When step 7 is reached all possible periodic binary sequences starting with $x_{i,1},\ldots, x_{i,m-1}$ and

satisfying the conditions given in lemma 1 (section 2.5) have been considered. Therefore a new starting value should be selected.

APPENDIX B

Theorem

$\Pi_n U$ is a neighbourly polytope, for all $n \in \mathbb{N}$.

Proof

$\Pi_2 U$ is a line segment and $\Pi_3 U$ is a tetrahedron and therefore neighbourly polytopes. We shall prove the result by induction. Fix $n > 3$. Fix $i, j \leqslant n$, $i \neq j$.

Let $V_{ij}^{+} \triangleq \Pi_n U \cap \{M : m_{ij}=1\}$. Then precisely half of the 2^{m-1} vertices of $\Pi_n U$ lie on V_{ij}^{+}; the remainder lie strictly to one side of the hyperplane $\{M : m_{ij}=1\}$ and are in $V_{ij}^{-} = \Pi_n U \cap \{M : m_{ij}= -1\}$. Now define the projection operator Π^i such that

$\Pi^i : \Pi_n U \longrightarrow \Pi_{n-1} U$,

that is the linear map obtained by removing the ith row and column from a matrix $R \in \Pi_n U$. Then we have

$\Pi^i E_1^n = E_k^{n-1}$, some $1 \leqslant k \leqslant 2^{n-2}$.

Furthermore $V_{ij}^{+} \cap (\Pi^i)^{-1} E_k^{n-1} = E_1^n$. Hence Π^i maps the 2^{n-2} vertices of V_{ij}^{+} onto the 2^{n-2} vertices of $\Pi_{n-1} U$ in a one-to-one fashion.

Therefore under the assumption that $\Pi_{n-1} U$ is neighbourly V_{ij}^{+} is also and by a similar argument so is V_{ij}^{-}.

We note at this point that the edges of V^+_{ij} are also edges of $\Pi_n U$ because V^+_{ij} (and V^-_{ij}) is the result of the intersection of $\Pi_n U$ with a supporting hyperplane. Furthermore any pair E^n_k, E^n_l of vertices of $\Pi_n U$ lie either in V^+_{ij} or V^-_{ij} for some i,j , $j \neq i$.

From the above arguments it follows that the line segment connecting any pair of vertices of $\Pi_n U$ is an edge of $\Pi_n U$. Since $\Pi_3 U$ is neighbourly it follows by induction that $\Pi_n U$ is also neighbourly for $n > 3$.

q.e.d.

APPENDIX C

In this appendix two numerical examples of application of an iterative algorithm, similar to the one described in section 4.3.2, are presented; in this algorithm only one new term of the sequence is selected at each iteration.

The cost at iteration n is defined as $\|\underline{c}_n - \underline{\rho}\|$ with $\underline{c}_n$ and $\underline{\rho}$ as defined in chapter 4.

```
EXAMPLE NO.1
 THE PRESCRIBED AUTOCORRELATION VECTOR IS
A( 1)=  .1000000
A( 2)=  .6000000
ITERATION OF ORDER    6
C( 1)=  .5000000
C( 2)=  .3333333
COST=  .4807402
ITERATION OF ORDER    7
C( 1)=  .2857143
C( 2)=  .1428571
COST=  .4934262
ITERATION OF ORDER    8
C( 1)=  .1250000
C( 2)=  .2500000
COST=  .3508917
ITERATION OF ORDER    9
C( 1)=         0
C( 2)=  .3333333
COST=  .2848001
ITERATION OF ORDER   10
C( 1)= -.1000000
C( 2)=  .4000000
COST=  .2828427
ITERATION OF ORDER   15
C( 1)=  .1333333
C( 2)=  .4666667
COST=  .1374369
ITERATION OF ORDER   20
C( 1)=  .0500000
C( 2)=  .5000000
COST=  .1118034
ITERATION OF ORDER   30
C( 1)=  .0333333
C( 2)=  .5333333
COST=  .0942809
```

```
ITERATION OF ORDER   40
C( 1)=  .0750000
C( 2)=  .5500000
COST=  .0559017
ITERATION OF ORDER   50
C( 1)=  .1000000
C( 2)=  .5600000
COST=  .0400000
ITERATION OF ORDER   60
C( 1)=  .0833333
C( 2)=  .5666667
COST=  .0372678
ITERATION OF ORDER   70
C( 1)=  .1000000
C( 2)=  .5714286
COST=  .0285714
ITERATION OF ORDER   80
C( 1)=  .0875000
C( 2)=  .5750000
COST=  .0279508
ITERATION OF ORDER   90
C( 1)=  .1000000
C( 2)=  .5777778
COST=  .0222222
ITERATION OF ORDER  100
C( 1)=  .0900000
C( 2)=  .5800000
COST=  .0223607
 THE GENERATED SEQUENCE IS
  1. -1. -1. -1. -1. -1.  1. -1.  1. -1.  1.  1.  1.  1.  1.
  1.  1. -1.  1. -1.  1.  1.  1.  1.  1.  1. -1.  1. -1.  1.
 -1. -1. -1. -1. -1. -1. -1.  1. -1.  1. -1.  1.  1.  1.  1.
  1.  1.  1. -1.  1. -1.  1.  1.  1.  1.  1.  1. -1.  1. -1.
  1. -1. -1. -1. -1. -1. -1. -1.  1. -1.  1. -1. -1. -1. -1.
 -1. -1.  1. -1.  1. -1.  1.  1.  1.  1.  1.  1. -1.  1.
 -1.  1.  1.  1.  1.  1.  1. -1.  1. -1.


EXAMPLE NO.2
 THE PRESCRIBED AUTOCORRELATION VECTOR IS
A( 1)= -.2000000
A( 2)=  .2000000
A( 3)= -.2000000
A( 4)=  .2000000
A( 5)=-1.0000000
ITERATION OF ORDER   10
C( 1)=  .3000000
C( 2)=  .2000000
C( 3)= -.1000000
C( 4)= -.2000000
C( 5)= -.3000000
COST=  .9539392
ITERATION OF ORDER   50
C( 1)= -.1000000
C( 2)=  .2000000
C( 3)= -.1800000
C( 4)=  .1200000
C( 5)= -.8200000
COST=  .2218107
```

```
ITERATION OF ORDER  100
C( 1)= -.1500000
C( 2)=  .2000000
C( 3)= -.1900000
C( 4)=  .1600000
C( 5)= -.9100000
COST=  .1109054
ITERATION OF ORDER  150
C( 1)= -.1666667
C( 2)=  .2000000
C( 3)= -.1933333
C( 4)=  .1733333
C( 5)= -.9400000
COST=  .0739369
ITERATION OF ORDER  200
C( 1)= -.1750000
C( 2)=  .2000000
C( 3)= -.1950000
C( 4)=  .1800000
C( 5)= -.9550000
COST=  .0554527
ITERATION OF ORDER  250
C( 1)= -.1800000
C( 2)=  .2000000
C( 3)= -.1960000
C( 4)=  .1840000
C( 5)= -.9640000
COST=  .0443621
ITERATION OF ORDER  300
C( 1)= -.1833333
C( 2)=  .2000000
C( 3)= -.1966667
C( 4)=  .1866667
C( 5)= -.9700000
COST=  .0369685
ITERATION OF ORDER  350
C( 1)= -.1857143
C( 2)=  .2000000
C( 3)= -.1971429
C( 4)=  .1885714
C( 5)= -.9742857
COST=  .0316872
ITERATION OF ORDER  400
C( 1)= -.1875000
C( 2)=  .2000000
C( 3)= -.1975000
C( 4)=  .1900000
C( 5)= -.9775000
COST=  .0277263
ITERATION OF ORDER  450
C( 1)= -.1888889
C( 2)=  .2000000
C( 3)= -.1977778
C( 4)=  .1911111
C( 5)= -.9800000
COST=  .0246456
ITERATION OF ORDER  500
C( 1)= -.1900000
C( 2)=  .2000000
C( 3)= -.1980000
C( 4)=  .1920000
C( 5)= -.9820000
COST=  .0221811
```

THE GENERATED SEQUENCE IS

1.	-1.	1.	1.	1.	1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.
1.	-1.	1.	1.	1.	-1.	1.	-1.	-1.	-1.	1.	-1.	1.	1.	1.
-1.	1.	-1.	-1.	-1.										

APPENDIX D

Lemma

Let $(s_n)_{n \in \mathbb{N}_0}$ be a periodic sequence of period p and autocorrelation function $(\rho_n)_{n \in \mathbb{N}_0}$. Then for any given $m \in \mathbb{N}$ the sequence $(\underline{c}_{k_0+np})_{n \in \mathbb{N}_0}$ lies on a straight line passing through $\underline{\rho}$ for any fixed k_0, $1 \leqslant k_0 \leqslant p$, where $\underline{\rho} \equiv (\rho_1, \ldots, \rho_m)$, $\underline{c}_n \equiv (c_n^1, \ldots, c_n^m)$ and

$$c_{n+1}^i = \frac{n}{n+1} c_n^i + \frac{s_{n+1-i}\, s_{n+1}}{n+1} .$$

Proof

All we need to show is that for all $n \in \mathbb{N}$ there exists $\lambda \in \mathbb{R}$ such that $(\underline{c}_{k_0+np} - \underline{c}_{k_0}) = \lambda(\underline{\rho} - \underline{c}_k)$.

From the definition of c_n^i we can write

$$c_{k_0+np}^i - c_k^i = \frac{\sum_{t=k_0+1}^{k_0+np} s_t s_{t-i} - npc_k^i}{k_0+np} = \frac{np}{k_0+np}(\rho_i - c_{k_0}^i),$$

because $(s_n)_{n \in \mathbb{N}_0}$ has period p and by definition

$$\frac{1}{p} \sum_{t=k_0}^{k_0+p} s_t s_{t-i} = \rho_i .$$

The result then follows.

Corollary

If $(s_n)_{n \in \mathbb{N}_0}$ is 'eventually periodic' that is there exist natural numbers n_0 and p such that $s_{np+m} = s_m$ for all $m \geqslant n_0$ and $n \in \mathbb{N}_0$, then $(\underline{c}_{k_0+np})_{n \in \mathbb{N}_0}$ also lies in a straight line passing through $\underline{\rho}$, provided $n_0 \leqslant k_0 < n_0+p$.

Proof

We have seen that

$$c^i_{k_0+np} - c^i_k = \frac{p}{k_0+np} \frac{\sum_{t=k_0+1}^{k_0+np} s_t s_{t-i}}{p} - \frac{np c^i_{k_0}}{k_0+np}$$

If $k_0 \geqslant n_0$ then $\frac{1}{p} \sum_{t=k_0+1}^{k_0+np} s_t s_{t-i} = n\rho_i$

The result follows.

APPENDIX E

Theorem (B. Mcmillan)

$\rho \in U$ if and only if $\rho(0)=1$ and

$$\sum_{m=1}^{N} \sum_{n=1}^{N} \rho(n-m) A_{mn} \geqslant 0,$$

for every N, and all corner-positive matrices $\{A_{mn}\}$, $m,n=1,\ldots N$, where a matrix $\{A_{mn}\}$ is corner-positive if

$$\sum_{m=1}^{N} \sum_{n=1}^{N} A_{mn} \varepsilon_m \varepsilon_n \geqslant 0$$

for every sequence $(\varepsilon_1,\ldots,\varepsilon_N)$, with $\varepsilon_i=\pm 1$, $i=1, 2,\ldots,N$.

L. Shepp's proof of this theorem is now reproduced below.

Proof

The necessity of both conditions stated in Mcmillan's theorem is easily proved, since

$$\rho(0) = E(X_n^2) = E(1) = 1$$

and

$$\sum_{m=1}^{N} \sum_{n=1}^{N} \rho(n-m) A_{mn} = E\left(\sum_{m=1}^{N} \sum_{n=1}^{N} A_{mn} X_m X_n \right) \geqslant 0$$

The last inequality is an immediate consequence of the fact that $\{A_{mn}\}$ is corner-positive.

To prove sufficiency of Mc Millan's conditions, let us consider a given function $\rho(n)$, $n=0, \pm 1,\ldots$, and ask the question: under what conditions is it possible

to define a unit process $\{X_n\}$ with covariance function ρ ? Since the covariance function determines the two-dimensional distributions $P\{X_m=\varepsilon_1, X_n=\varepsilon_2\}$, the question can be expressed in the equivalent form: given the one-dimensional and two-dimensional distributions, when can an extension be made to the higher dimensional distributions?

It will be useful to formulate the question in still another way. Let us regard X_n as a coordinate variable in the space, Ω of all infinite sequences of ± 1's; that is $\Omega = \{(\ldots, x_{-1}, x_0, x_1, \ldots)\}$, $x_i = \pm 1$, is the set of all possible realizations of a unit process. Then our goal is to find a probability measure P on the space Ω such that

$$E(X_m X_n) = \int X_m X_{m+n} dP = \rho(n) \qquad (1)$$

where ρ is given. Let $C(\Omega)$ be the set of all continuous functions on Ω. The operator E is defined by ρ for those functions in $C(\Omega)$ which are of the form $X_m X_n$. If we can find the desired probability measure P, then E will be extended to all of $C(\Omega)$. On the other hand, the converse of this statement follows from the Riesz Representation Theorem, which asserts that if E is a functional defined on $C(\Omega)$ which is linear and positive ($g \geq 0 \Rightarrow E(g) \geq 0$), with $E(1)=1$, then E is given by an integral with respect to a probability measure. Thus

our goal is to determine under what conditions one can extend an operator E, which is defined by (1) on the set G of functions $X_m X_{m+n}$, to a positive, linear functional defined on all of $C(\Omega)$.

We shall use the following elegant result which I first discovered in a paper by H. G. Kellelerer[1].

Lemma 1 Let E be a functional defined on a subset C of $C(\Omega)$, where $1 \in G$ and $E(1)=1$. Then a linear positive extension of E to $C(\Omega)$ exists if and only if for any $g_1,\ldots,g_N \in G$ and any real numbers $\alpha_1,\ldots,\alpha_N$,

$$\sum_{i=1}^{N} \alpha_i g_i \geqslant 0 \Longrightarrow \sum_{i=1}^{N} \alpha_i E(g_i) \geqslant 0 \qquad (2)$$

The necessity of this condition is clear. To prove sufficiency, let M be the linear space spanned by G; i.e., the space $M = \{\sum \alpha_i g_i\}$ of all linear combinations of functions in G. Define E on M as follows:

$$E(\sum \alpha_i g_i) = \sum \alpha_i E(g_i) \qquad (3)$$

Note that E is well-defined on M; i.e., independent of representation, since if $\sum \alpha_i g_i = 0$, then $\sum \alpha_i E(g_i) = 0$.

[1] H. G. Kellerer, 'Verteilungsfunctionen mit Gegebenen Marginalverteilunger', Zeitscrift fur Wahrscheinlichkeitstheorie, 3, 1964, pp. 247-270.

E satisfies the conditions

a) E is linear on M

b) $|E(g)| \leqslant \|g\|$, where $\|g\| = \sup_{\text{over } \Omega} |g|$

Condition (a) is an immediate consequence of the definition (3). To prove condition (b), note that $\pm\sum\alpha_i g_i \leqslant 1 . \sup_{\text{over}\,\Omega} |\sum\alpha_i g_i|$, and since $1 \in G$, it follows from (2) that

$$E(\pm\textstyle\sum\alpha_i g_i) \leqslant \sup|\sum\alpha_i g_i| E(1) = \|\sum\alpha_i g_i\|$$

or $$\left|E\left(\textstyle\sum\alpha_i g_i\right)\right| \leqslant \|\sum\alpha_i g_i\|$$

From conditions (a) and (b), it follows by the Hahn-Banach theorem that there is an extension of E on M to a linear functional $\bar{E}$ on $C(\Omega)$, with $|\bar{E}g| \leqslant \|g\|$, $\forall g \in C(\Omega)$. But the latter condition implies that $\bar{E}$ is positive on $C(\Omega)$, for if $0 \leqslant g \leqslant 1$, then

$$1/2 - \bar{E}(g) = \bar{E}(1/2-g) \leqslant \|1/2-g\| \leqslant 1/2 \Rightarrow \bar{E}(g) \geqslant 0$$

We now return to the proof of McMillan's theorem. In this case, $G = \{X_i X_j\}$. Furthermore, $1 \in G$ since $X_n^2 = 1$, and $E(1) == \rho(0) = 1$. Now, the condition $\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_{ij} X_i X_j \geqslant 0$ for all ± 1 values of the X_i is equivalent to the condition that $\{\alpha_{ij}\}$ is a corner-

-positive matrix. But then condition (2) of the lemma becomes

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij}X_iX_j \geqslant 0 \Longrightarrow \sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij}E(X_iX_j)$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij}p(i-j) \geqslant 0$$

But this is exactly McMillan's condition.

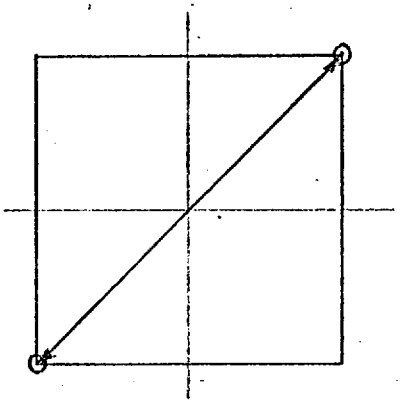


Figure 1 : D_1^+

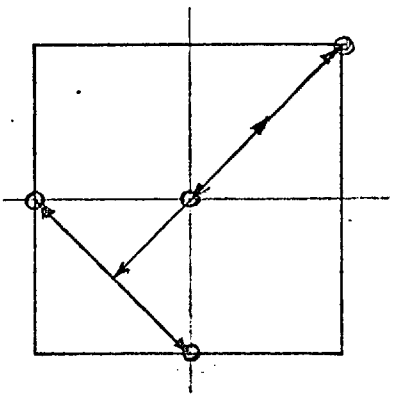


Figure 2 : D_2^+

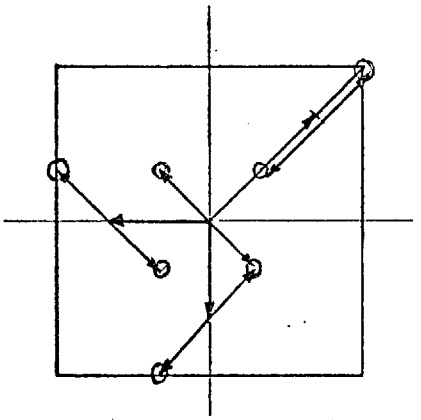


Figure 3 : D_3^+

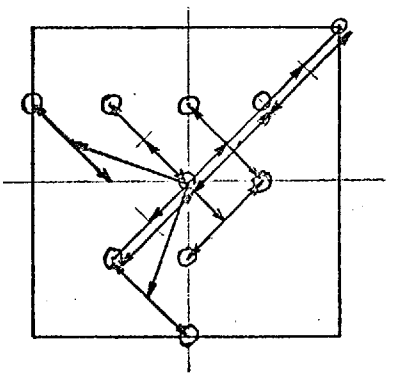


Figure 4 : D_4^+

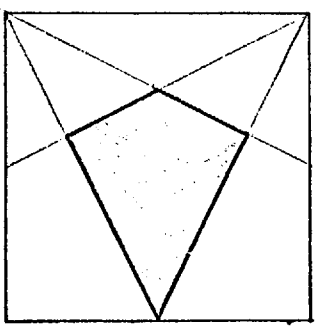


Figure 5 : W_2^i

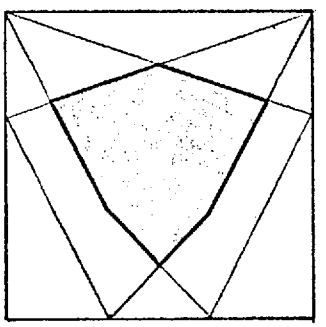


Figure 6 : W_3^i

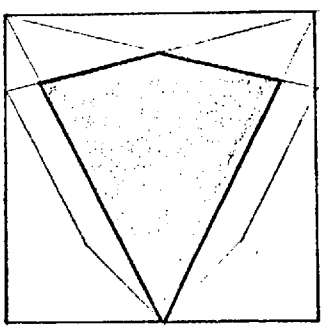


Figure 7 : W_4^i

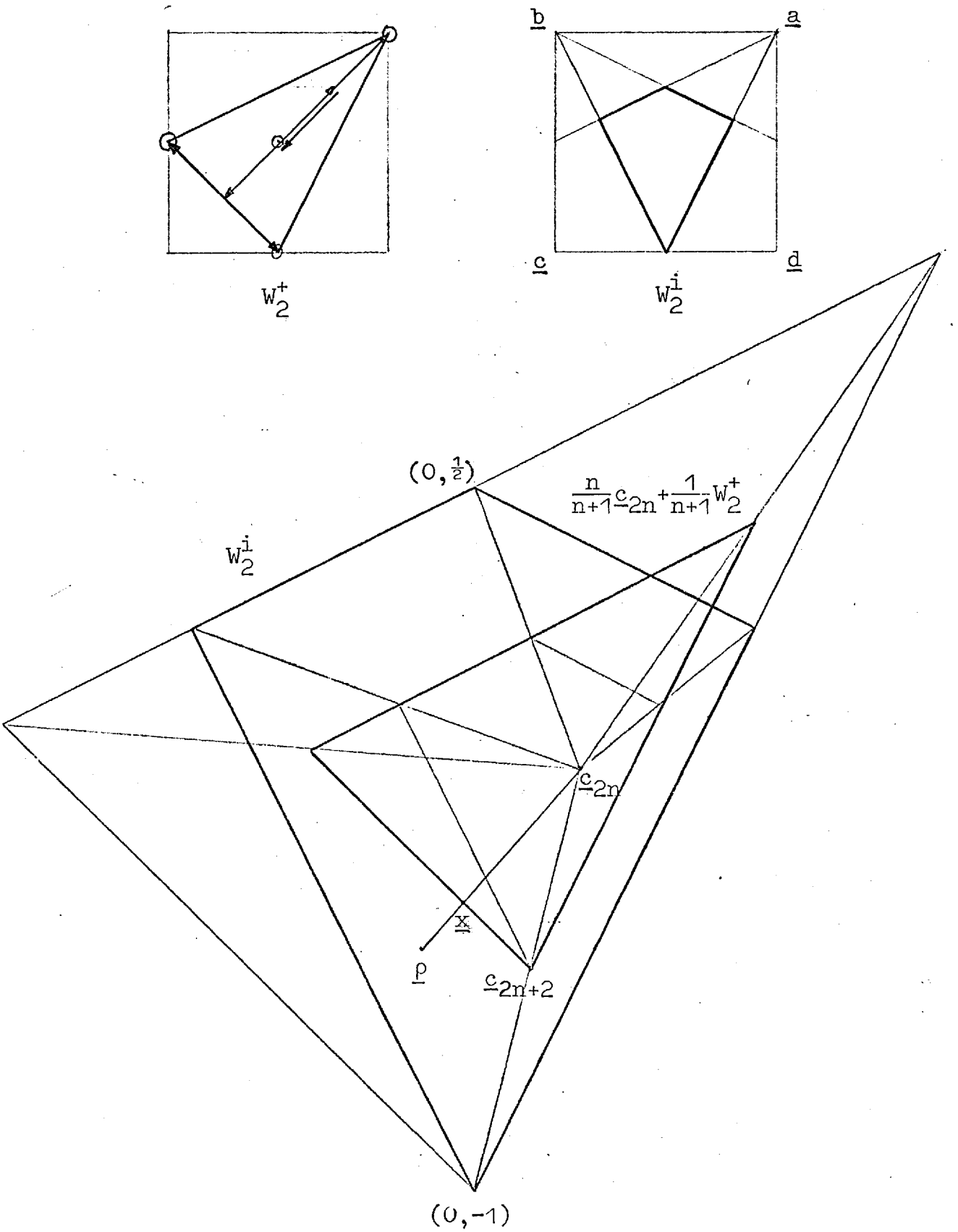


Figure 8

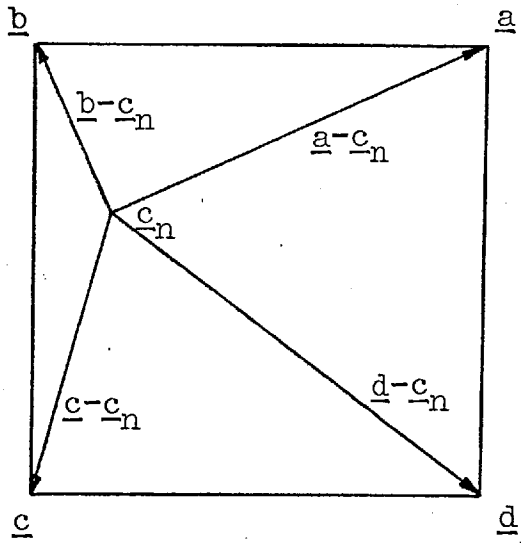


Figure 9

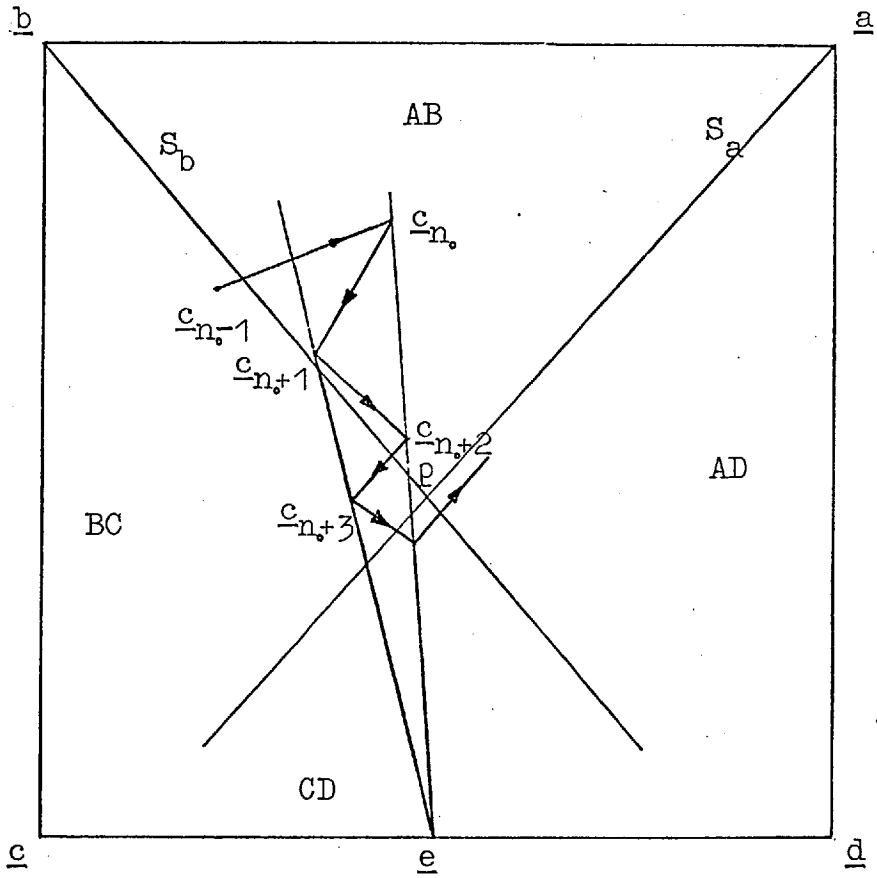


Figure 10

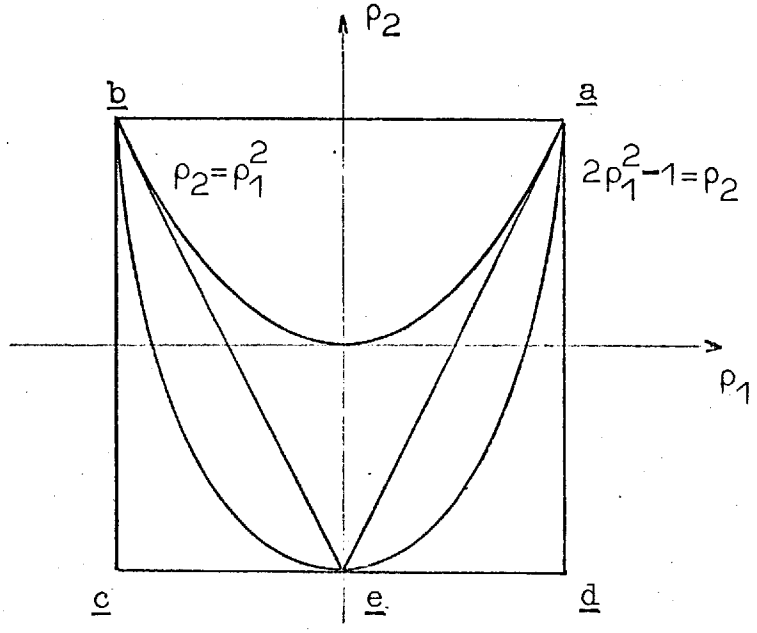
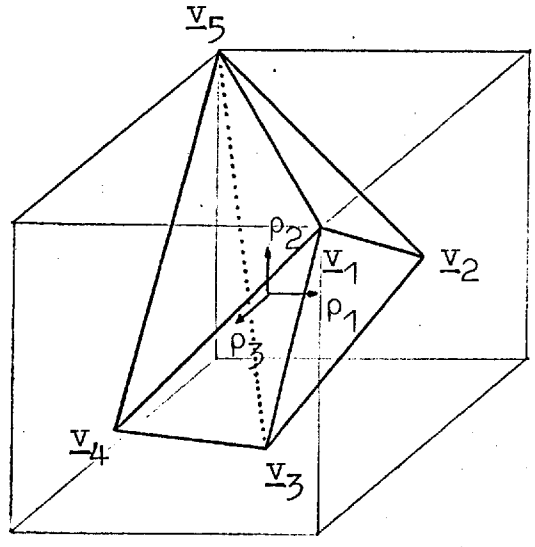


Figure 11



$\underline{v}_1 = (1,1,1)$

$\underline{v}_2 = (\frac{1}{3},-\frac{1}{3},-1)$

$\underline{v}_3 = (0,-1,0)$

$\underline{v}_4 = (-\frac{1}{3},-\frac{1}{3},1)$

$\underline{v}_5 = (-1,1,-1)$

Figure 12 : $\Pi_4 C$