ON GENERALIZED CONVOLUTION AND FOURIER TRANSFORMS

TOPICS ON GENERALIZED CONVOLUTION AND FOURIER TRANSFORMS: THEORY AND APPLICATIONS IN DIGITAL SIGNAL PROCESSING AND SYSTEM THEORY.

By

SORIN COHN-SFETCU, Dipl.Eng., M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

McMaster University

October 1976

DOCTOR OF PHILOSOPHY (1976)
(Electrical Engineering)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Topics on Generalized Convolution and Fourier Transforms: Theory and Applications in Digital Signal Processing and System Theory.

AUTHOR: Sorin Cohn-Sfetcu, Dipl. Eng. (Polytechnic Inst. of Bucharest)
M.Sc. (The University of Calgary)

SUPERVISOR: Professor R. Kitai

NUMBER OF PAGES: xix, 274

ABSTRACT

This thesis investigates some aspects of the theory of convolution and Fourier transforms on q-adic and multiplicative abelian groups, as well as their applications for solving various signal processing and system problems.

A brief introduction to the importance and basic theory of convolution and Fourier transforms on locally compact abelian groups is given, followed by four major sections:

1. Subsequent to a comprehensive introduction of generalized Walsh functions, Walsh-Fourier analysis and harmonic differentiation on q-adic groups, a presentation is made of the theory of q-adic translation invariant linear systems from the point of view of both input-output and state-space description. This is followed by an analysis of the structure of Walsh transforms, so that it becomes possible to point to (and critically review) those engineering problems for which Walsh functions are suited to bring an optimal solution, as well as those problems for which they may bring suboptimal but efficient solutions.

2. Signal processing in spaces of finite field-valued functions on finite abelian groups is investigated, emphasis being placed on the study of those linear operators whose eigenfunctions are the group characters. A harmonic differential calculus in finite fields is introduced.

3. A study of the concept of frequency is undertaken with the objective of generalizing it to function spaces other than that of complex-valued functions on the real line. A generalized concept of frequency is proposed. An analysis of the relationship between the concepts of sequency and frequency proves unfounded the claims that the former is a generalization of the latter.

4. The problem of analyzing signals formed of linear combinations of components having the same shape and location but different amplitude and widths parameters is investigated with the objective of providing a technique for its numerical solution. It is shown that this problem can be modelled as a convolution transform on a multiplicative abelian group. A brief introduction to the theory of Fourier transforms on multiplicative groups is presented, followed by the description of an efficient algorithm for performing the analysis. The problems pertaining to the practical implementation of this algorithm are discussed both in general terms and with reference to the analysis of multi-component exponential decays.

ACKNOWLEDGEMENTS

The author wishes to thank all the members of the Department of Electrical Engineering of McMaster University who have contributed in many ways to the completion of this work.

Most profound gratitude is due to Dr. R. Kitai for his supervision of the entire work and the valuable help in the writing and presentation of this thesis, for financially assisting my participation in the 1975 Walsh Symposium (Hatfield, England), as well as for the friendship he has extended to me since my coming to Hamilton.

Thanks are also due to Dr. R. de Buda, Dr. N. Sinha and D. J. Stewart who, as members of the supervisory committee, have offered encouragement and many helpful suggestions.

The author also wishes to thank:

- Dr. S.T. Nichols for his consideration and guidance during the year of research work in the Department of Electrical Engineering of The University of Calgary, and

- Dr. J.E. Gibbs for the fruitful correspondence carried with him during the past two years, as well as for the research work done together during the one-month visit to the National Physical Laboratory (Teddington, England), visit which he facilitated in no small measure.

Special mention deserves my friend, Dr. M.R. Smith, for the close collaboration and many heated discussions that yielded as many fruitful results.

The gracious work of Mrs. S. Ryan in typing this thesis is also acknowledged.

The completion of this work would not have been possible without the financial assistance of

- the Killam Foundation (in the form of an Isaac Killam Memorial Scholarship, 1973-1975),

- the National Research Council of Canada (in the form of a Postgraduate Scholarship, 1973-1976),

- McMaster University and The University of Calgary (in the form of part-time Teaching Assistantships, 1973-1976).

Finally, I would like to acknowledge the encouraging attitude of my colleagues at Bell-Northern Research, where I have been working for the past four months.

Are signals really bandlimited? They seem to be, and yet they seem not to be.

David Slepian,

"On bandwidth", Second Shannon Lecture, Proc. IEEE, Vol. 64, No. 3, pp. 292-300, March 1976.

TABLE OF CONTENTS

LIST OF FIGURES

INDEX OF SYMBOLS AND ABBREVIATIONS
(In order of appearance in the text)

SYMBOL, ABB.		PAGE
Z	the group of integers	6
$\|\cdot\|$	absolute value function	6
$\chi_\nu(\cdot), \chi(\nu,\cdot)$	character of index ν	6
ν, λ	indexes of characters	6
Γ	dual group of G	7
$\overline{(\cdot)}$	complex conjugate of $(\cdot)$	7
$\oplus$	Cartesian addition (product) of sets, also dyadic addition	7
$\{\cdot\}_{i=1}^{n}$	set of elements indexed from 1 to n	7
F	Fourier transform operator	8
$\widehat{(\cdot)}$	Fourier transform of $(\cdot)$	8
FT	Fourier transform	8
$<\cdot\ ,\ \cdot>$	inner product	9
Z_q	set (group) of smallest q integers	10
$\sim$	isomorphic to	10
DFT	discrete Fourier transform	11
R_+	set of non-negative real numbers	11
FFT	fast Fourier transform (algorithm)	12
EPR	electron paramagnetic resonance	12
WFT, WT	Walsh (-Fourier) transform	20
I	unit interval on the real line	25
P_q	set (group) of smallest q non-negative integers	28
$\mathrm{wal}(\lambda,\cdot)$	Walsh function of index λ	32

SYMBOL, ABB.		PAGE
$\mathrm{rad}_k(\cdot)$	k-th Rademacher function	33
$[\cdot]$	greatest integer function	35
W_q	q-th root of unity	35
$\dot{+}$	addition modulo q	41
S^n	set formed by Cartesian sum (product) of n copies of the set S	41
G_q	q-adic group	41
$(\underline{\cdot})$	n-type vector	42
$(\cdot)_i$	i-th element of the n-tuple $(\underline{\cdot})$	42
$(\underline{\cdot}, \underline{\cdot})$	pseudo inner product	44
$\dot{*}$	q-adic convolution	46
$\triangleq$	equal by definition	47
$\alpha(\cdot)$	Fine mapping	48
W	Walsh transform operator	55
A, A_N	matrix (of order N)	57
I	identity matrix	57
A^T	transpose of matrix A	57
D	harmonic differential operator	58
q-TIL	q-adic translation invariant linear (system)	65
TIL	translation ('time') invariant linear (system)	65
U	system input function space	65
u	input function	65
Y	system output function space	65
y	output function	65

SYMBOL, ABB.		PAGE
s	system relation	65
D_x	domain of function x	66
C_x	co-domain of function x	66
H	linear system operator	67
h	system impulse response	71
E	expected value operator	76
r_{vw}	(cross-) correlation function	77
ρ_{vw}	Walsh (cross-) spectral density function	77
e	error	80
$\sigma^2_{uy}(\cdot)$	coherence function	81
L_{λ_i}	diagonal matrix	87
wal(A,t)	Walsh matrix function	88
DWT	discrete Walsh transform	96
p	prime number	101
$\dot{(\cdot)}$	time (classical) derivative of $(\cdot)$	127
U	unitary transformation operator	131
R_{xz}	cross-correlation matrix	131
R_L	logical autocorrelation matrix	132
R_A	arithmetic autocorrelation matrix	133
R_Q	q-adic autocorrelation matrix	133
$T_{AL}(T_{AL})$	$R_A \rightarrow R_L$ transformation operator (matrix)	133
$T_{AQ}(T_{AQ})$	$R_A \rightarrow R_Q$ transformation operator (matrix)	133
P_W	Walsh power spectrum	135

SYMBOL, ABB.		PAGE
P_F	Fourier power spectrum	135
$T(h_T)$	matrix (impulse response function) of TIL system	136
$D(h_D)$	matrix (impulse response function) of dyadic TIL system	137
$Q(h_Q)$	matrix (impulse response function) of q-TIL system	139
$[\cdot]_{ij}$	ij-th element of matrix $[\cdot]$	139
H_{TQ}	transformation matrix T→Q	139
$i_A(\cdot)$	injective mapping of A→Z	146
J	integral domain	147
m	exponent of group G	147
ζ, ζ_N	primitive N-th root of 1 in K	148
$d(\cdot,\cdot)$	distance function	152
$tr_{K/F}$	trace function in field $K = F^n$	152
$(\cdot)'$	Newton-Leibnitz derivative of $(\cdot)$	160
ν	frequency	166
$\omega=2\pi\nu$	angular frequency	172
$\Phi=\omega t$	phase	172
A_f	modulus of function f	172
Φ_f	phase of function f	172
$\Psi_f=j\ \Phi_f$	imaginary phase of function f	172
$\lambda=j\omega$	imaginary angular frequency	172
Δ	slope (relative variation) operator	172
$n(\sigma,w)$	winding number of curve σ around point w	174

SYMBOL, ABB.		PAGE
ξ, ξ_N	primitive N-th root of unity in $\mathbb{C}$	178
Δ_N	slope operator in K[G] where order of G is N	179
$k(t,s)$	kernel function	189
J_a	Bessel function of order a	192
M	Mellin transform operator	192
$\dot{G}$	multiplicative abelian group	195
$\dot{R}_+$	multiplicative group of positive real numbers	196
$\tilde{(\cdot)}$	transform of function $(\cdot)$ under exponential change of variables	196
$\dot{T}$	multiplicative translation operator	200
$\dot{*}$	multiplicative convolution	201
$\dot{\chi}$	character of multiplicative abelian group	201
$\dot{F}$	multiplicative Fourier transform operator	202
m-	multiplicative -	202
$\hat{(\cdot)}$	multiplicative Fourier transform of $(\cdot)$	202

CHAPTER I

INTRODUCTION

1.1 AIM AND MOTIVATION

This dissertation concerns some aspects of the theory and applications of linear transformations on spaces of functions having a locally compact abelian group G as domain and a field K as co-domain. K may be either the field of complex numbers $\mathbb{C}$, real numbers R, or a finite Galois field $GF(q^n)$. The choice of research topic was motivated by two factors:

(i) the need for solving certain problems essential for the extraction of information from signals encountered in various experimental situations; in their turn, these problems can be grouped in two categories - (a) separation of signals from noise (sensitivity enhancement), and (b) separation of concurrent signals (resolution enhancement).

(ii) an intellectual curiosity with respect to understanding a variety of signal processing techniques, and the desire of extracting the common, essential mathematical structure laying at their foundation, so that these techniques may be optimally applied to solving engineering problems.

This introductory chapter serves a twofold purpose; in the first place, it attempts to present, in broad lines, the general mathematical framework within which the investigation has been conducted; and, in

the second place, it attempts to elucidate the personal motivation behind this study which, in turn, explains the line of thought adopted.

The pivotal idea of the dissertation is that of convolution and Fourier-type transforms which realize a mapping between function algebras with convolution product and function algebras with pointwise multiplication product. It would be difficult to underestimate the importance of the concept of convolution, especially when it is regarded from an engineering perspective, i.e. from the point of view of obtaining simple, mathematically manageable models of real world phenomena and relationships. Mikusinski [1] considered this concept fundamental enough to lay it at the foundation of operational calculus. Its pre-eminence is due to the fact that it appears as a natural consequence of the two most basic restrictions assumed in building such simple models: linearity and translation invariance.

Linearity implies usually a function space with a field (or at most a ring) as co-domain and the restriction to only those operators which are linear on this function space. In its turn, translation invariance implies a function domain which is a locally compact abelian (LCA) group [2]. A group is an ordered pair $(G,\dagger)$, where G is a set of elements which is closed under the single valued binary operation $\dagger$, which is associative, and relative to which G contains a neutral element e_G, and with each element x another element x^- (called its inverse) such that their composition produces the neutral element [3,4]. If the law of composition is commutative, the group is said to be abelian. The classic example of such a group is the group of real

numbers under usual arithmetic addition (+). Classical analysis, system theory and signal processing are almost exclusively concerned with the space $\mathcal{C}[R]$ of complex-valued functions on the group of additive real numbers. This thesis' concern lies, partly, somewhere else, with spaces of functions other than $\mathcal{C}[R]$.

The study of spaces other than $\mathcal{C}[R]$ is warranted, in the first place, by the fact that there are real-life signals and systems which are naturally modeled as functions and, respectively, relations between functions on abelian groups other than R, taking values in fields other than $\mathcal{C}$. In the second place, the structure implied by the axiomatization of an abelian group as domain permits very elegant mathematical treatments. In addition, this axiomatization has the advantage, from an engineering point of view, of leading to faster and more efficient use of digital techniques and hardware which, due to a continuous increase in speed and decrease in cost, size and power consumption, have come to play a major role in signal processing. The other major restriction (motivated by the same kind of considerations) is that the function space be a linear Banach space, which, wherever possible, will be also a Hilbert space, because, as emphasized by Lorch [5], this is, of all spaces, the one that presents the greatest orderliness and greatest plenitude of mathematical results. In the sequel, an attempt is made at introducing some of the basic concepts of harmonic analysis which will be later employed throughout the dissertation.

1.2 MATHEMATICAL BACKGROUND: CONVOLUTION AND FOURIER TRANSFORMS

Since this dissertation adopts an engineering point of view, being concerned with the digital application of transform techniques, it will be mostly interested, for obvious reasons, in finite groups, whose order, $\#(G)$, will be denoted by N. Nevertheless, whenever warranted, the mathematical treatment will be generalized to locally compact abelian groups since local compactness is the natural extension of finiteness [6]. The idea behind using LCA groups is that it is possible to construct on them a Haar measure [7] μ with the property of being translation invariant in the sense that $\mu(x+V) = \mu(V)$ for any $x \in G$ and any Borel set V in G. This measure is unique, up to a multiplicative positive constant, with the customary normalization that $\mu(G) = 1$ if G is compact, while, if G is discrete, every single element set in G is assigned a unit measure. With this in mind, function integration on G will always be considered with respect to the Haar measure, and will be denoted by $\int_G f(x)\,dx$; if the discussion is limited to discrete groups, the symbol $\Sigma_G f(x)$ will be used instead.

Consider a function space $K[G]$; the right and left translation operators T_+^y and T_+^y are defined by

$$T_+^y\{f(x)\} = f(x+y^-)$$

$$y \in G, \quad \forall f \in K[G]. \qquad (1.2.1)$$

$$T_+^y\{f(x)\} = f(y+x^-)$$

The set of all these operators $\{T^y\}_{y\in G}$ forms a group T with respect to multiplicative composition of operators, the neutral element being $T_+^{e_G}$.

The set of all right translation operators forms the right translation subgroup T_+. The set of all left translation operators together with the neutral translation form the left translation subset T_+, which has the interesting property that each element is its own inverse. Two other important theorems concern normed linear spaces of K-valued functions on G (such spaces will be denoted as $L^p[G]$):

(a) if G is LCA, the norm of $L^p[G]$, where $1 \leq p \leq \infty$, is translation invariant, and

(b) the mapping $y \to T^y\{f\}$ is a uniformly continuous mapping of, G into any $L^p[G]$.

In this context, the concept of convolution can be viewed as a generalization of the process of taking weighted linear combinations of a function's translates. Formally, generalized convolution is the mapping $* : K[G] \times K[G] \to K[G]$ defined by

$$(f*g)(x) = \int_G T^x f(y) g(y) dy. \tag{1.2.2}$$

(We shall attempt to streamline the notation whenever possible; for instance, we write Tf instead of $(T\{f\})(x)$, or $f*g$ instead of $(f*g)(x)$, etc.) The operation of convolution is commutative, associative and distributive with respect to K-addition in $L^1[G]$:

$$f*g = g*f, \quad \forall\ f, g \in L^1[G]; \tag{1.2.3a}$$

$$f*(ag) = (af)*g = a(f*g), \quad \forall\ f, g \in L^1[G], \quad \forall\ a \in K; \tag{1.2.3b}$$

$$f*(g+h) = f*g + f*h, \quad \forall\ f, g, h \in L^1[G]; \tag{1.2.3c}$$

$$f*(g*h) = (f*g)*h, \quad \forall\ f, g, h \in L^1[G]; \tag{1.2.3d}$$

such that $L^1[G]$ is a commutative Banach algebra if multiplication is defined by a convolution product. Moreover, if G is discrete, $L^1[G]$ has a unit denoted $\delta(x)$,

$$f*\delta = f, \quad \forall\ f \in L^1[G]. \tag{1.2.4}$$

For nondiscrete G, the role of unit is played by the Dirac distribution [8,9,10].

Convolution Banach algebras owe their role in signal processing and system theory to the fact that any linear translation invariant transformation can be represented by a convolution product [11,12]. Nevertheless, it is sometimes more advantageous to operate in pointwise multiplication algebras, and it is with respect to such situations that characters and Fourier transforms play a determinant role. It is, of course, possible to discuss classical Fourier analysis (on $\mathcal{C}[R]$ or $\mathcal{C}[\mathcal{T}]$, where $\mathcal{T}$, the circle group, is the quotient group of real numbers modulo the group of integers Z) without ever mentioning the word group character, but their introduction illuminates certain algebraic, and, hence, fundamental, structural features of Fourier transform theory.

Given the function space $\mathcal{C}[G]$, a character of a LCA group G is defined as a homomorphism into the group of unit absolute value complex numbers under the operation of multiplication:

$$|\chi_\nu(x)| = 1, \quad \forall\ x \in G;$$

$$\chi_\nu(x+y) = \chi_\nu(x) \cdot \chi_\nu(y), \quad \forall\ x, y \in G. \tag{1.2.5}$$

The set of all continuous characters of G form the dual group of G, denoted by $(\Gamma,@)$, with the important property [13] that every LCA group is (up to isomorphism) the dual group of its dual group. Moreover, if G is discrete, Γ is compact and vice-versa; if G is finite, Γ is also finite and isomorphic to it. Because of this duality, it is notationally advantageous to denote a character by $\chi(\nu,x)$ where $x \in G$ and $\nu \in \Gamma$. Some essential features of the characters follow easily from their definition:

$$\chi(\nu,x_1+x_2) = \chi(\nu,x_1)\ \chi(\nu,x_2) \quad \chi(\nu_1@\nu_2,x) = \chi(\nu_1,x)\ \chi(\nu_2,x); \tag{1.2.6a}$$

$$\chi(\nu,e_G) = \chi(e_\Gamma,x) = 1, \quad \forall\, x \in G, \quad \forall \nu \in \Gamma; \tag{1.2.6b}$$

$$\chi(\nu^-,x) = \chi(\nu,x^-) = 1/\chi(\nu,x) = \overline{\chi(\nu,x)}, \tag{1.2.6c}$$

where $\overline{(\cdot)}$ denotes complex conjugation. If G is compact, the integral of a non-constant character is equal to zero.

Two theorems [2] of consequence for this dissertation refer to the case when G is a sum of a set of groups $\{G_i\}$; i.e., when the set G is a cartesian product of the sets G_i, and the group operation is performed coordinatewise. In short, the theorems state that (1) if G is the direct sum of $\{G_i\}_{i=1}^n$, and $\{\Gamma_i\}$ is the set of dual groups, the dual group of G is the direct sum of $\{\Gamma_i\}$,

$$G = G_1 \oplus \ldots \oplus G_n \quad \rightarrow \quad \Gamma = \Gamma_1 \oplus \ldots \oplus \Gamma_n; \tag{1.2.7}$$

and (2) if G is the complete direct sum of a family of compact abelian groups $\{G_i\}$, then Γ is the direct sum of the corresponding dual groups

Γ_i, and

$$\chi(\nu,x) = \prod_i \chi(\nu_i,x_i) \qquad (1.2.8)$$

The Fourier transform (abbreviated FT) of any absolutely summable function $f \in L^1[G]$ is defined as

$$(F\{f(x)\})(\nu) = \hat{f}(\nu) = \int_G f(x)\chi(\nu,x^-)dx, \quad \nu \in \Gamma, \qquad (1.2.9)$$

and the mapping $f \to \hat{f}$ is a complex homomorphism of $L^1[G]$. Some of the most essential properties of the Fourier transform thus defined consist in the fact that [2]:

(a) the set $A[\Gamma]$ of all Fourier transforms $\hat{f}(\nu)$ is a dense subalgebra of the space $C_0[\Gamma]$ of continuous functions on Γ which vanish at infinity;

(b) $A[\Gamma]$ is invariant under translation and under multiplication by any character in the sense that

$$F\{T_+^y f(x)\} = \overline{\chi(\nu,y^-)}\cdot\hat{f}(\nu), \qquad (1.2.10)$$

$$F\{\chi(\mu,x)\cdot f(x)\} = \hat{f}(\nu\ominus\mu^-); \qquad (1.2.11)$$

(c) the Fourier transform of the convolution of two functions is equal to the pointwise multiplication of the FT's of the respective functions,

$$F\{f*g\} = \hat{f}\cdot\hat{g}; \qquad (1.2.12)$$

(d) the FT may be interpreted as a convolution,

$$\hat{f}(\nu) = (f*\chi_\nu)(e_G). \qquad (1.2.13)$$

Provided it satisfies certain conditions [2, sec. 1.5], a function $f \in L^1[G]$ can be recovered through an inverse FT defined as

$$f(x) = \int_\Gamma \hat{f}(\nu)\chi(\nu,x)d\nu, \quad \forall x \in G. \tag{1.2.14}$$

If the function space is restricted to $(L^1 \cap L^2)[G]$, the Plancherel theorem states that the FT defined by (1.2.7) and (1.2.14) is an isometry between $L^2[G]$ and $L^2[\Gamma]$. A corollary of this theorem is the Parseval formula

$$\int_G f(x)\ \overline{g(x)}\ dx = \int_\Gamma \hat{f}(\nu)\ \hat{g}(\nu)\ d\nu; \tag{1.2.15}$$

it may be interpreted as expressing the invariance of the inner product under a "complete orthonormal expansion" in the Hilbert space $L^2[G]$, in which the inner product has been defined as

$$\langle f\ ,\ g\rangle = \int_G f(x)\ \overline{g(x)}\ dx. \tag{1.2.16}$$

The classical examples of function spaces are $\mathcal{C}[R]$, $\mathcal{C}[\mathcal{T}]$, where $\mathcal{T}$ is alternatively viewed as the group of additive reals modulo 2π or the multiplicative group of complex numbers of unit absolute value, and $\mathcal{C}[Z]$. In all these cases, the characters appear to be expressed as exponential functions:

$$\chi(\nu,x) = e^{j\nu x} \begin{cases} G=R, \quad \Gamma=R, & \hat{f}(\nu) = \int_{-\infty}^{+\infty} f(x)e^{-j\nu x}dx, & (1.2.17a) \\ G=\mathcal{T}, \quad \Gamma=Z, & \hat{f}(n) = (1/2\pi)\int_{-\pi}^{\pi} f(e^{j\theta})e^{-jn\theta}d\theta, & (1.2.17b) \\ G=Z, \quad \Gamma=\mathcal{T}, & \hat{f}(e^{j\alpha}) = \sum_{n=-\infty}^{+\infty} f(n)e^{-jn\alpha}, & (1.2.17c) \end{cases}$$

$$G=Z_q, \quad \Gamma=Z_q, \quad \hat{f}(h) = \frac{1}{N}\sum_{k=0}^{q-1} f(h)e^{-j2\pi kn/q}, \qquad (1.2.17d)$$

where Z_q is the set of smallest q integers (a group with addition modulo q).

Rudin [6] considers the exponential function (also called the Euler function, especially if the argument is a complex number) as the most important function in mathematics; he does so for good measure because, besides being a character and hence defining a FT, the exponential function has many extremely interesting and intriguing properties:

(a) for any complex number z, $|z|<\infty$, exp(z) is a continuous non-zero function which can be represented by the absolutely uniformly convergent series

$$\exp(z) = \sum_{n=0}^{\infty} z^n/(n!) \qquad (1.2.18)$$

and for any $w \in \mathbb{C}-\{0\}$ there is a $z \in \mathbb{C}$ such that w=exp(z);

(b) there exists a positive transcendental number π such that

$$\exp(j\pi/2) = j, \text{and } e^z = 1 \text{ iff } (z/2j\pi) \text{ is an integer}, \qquad (1.2.19)$$

the exponential function being periodic of period $2j\pi$; (the number 2π also plays an important role in the normalization of the Haar measure on R as $G\sim\Gamma$.);

(c) it achieves an homomorphism of addition into multiplication,

$$\exp(z_1+z_2) = \exp(z_1)\exp(z_2); \qquad (1.2.20)$$

the mapping $x \to e^{jx}$ maps the real axis onto the unit circle in the complex plane, and the restriction e^x to the real axis is a monotonically in-

creasing positive function which realizes an isomorphism between the group of multiplicative reals R_+ and the group of additive real numbers;

(d) it is the eigenfunction of the differential operator,

$$(\exp(z))' = \exp(z). \qquad (1.2.21)$$

Any one of these properties may be used as a definition of the exponential function, starting from which it is possible to derive all the other properties. The intimate relationship between Fourier transform and differentiation is no accident; Fourier series were purposely invented in the eighteenth century as a tool for solving differential equations. The algebraic definition of the exponential as a character, related as it is to the concept of translation invariance, seems to be the more fundamental one [2,14]. A major and often recurrent theme of investigation throughout this dissertation consists in attempts at deriving similar properties for characters in function spaces other than $\mathcal{C}[G]$, where $G = R$, T or Z.

The engineering interest in FT on finite groups was kindled by the realization [15,16] that if the order of the group, N, is a highly composite number (initially it was $N=2^n$), then the corresponding discrete finite Fourier transform (DFT) can be computed much faster than was previously thought possible. The idea is that instead of computing

$$\hat{f}(i) = \sum_{k\in G} \chi(-i,k)\, f(k), \quad \chi(i,k) = \exp(j2\pi ik/N), \quad i \in \Gamma - G, \qquad (1.2.22)$$

directly, one can compute it in stages, over subgroups of G. Alternatively, the computation of a DFT can be viewed as a multiplication of the function vector (sequence of numbers) with the orthonormal matrix of

character values, which can be accomplished much faster if this matrix is represented as a product of sparse matrices. With certain precautions, the DFT may be used to approximate efficiently a continuous infinite Fourier transform [17,18]; and the speed of the fast Fourier transform algorithm (FFT), coupled with the decrease in cost and the increase in speed and storage capacity of digital hardware, led to a tremendous boost in the theoretical development and practical applications of digital signal processing [19,20,21].

1.3 PERSONAL INVOLVEMENT: PREVIOUS AND PRESENT RESEARCH

My personal interaction with this vast, rich and stimulating field of scientific investigation has been initiated and conditioned by my first research work environment - that of a magnetic resonance, more specifically, an electron paramagnetic resonance (EPR) laboratory. As is the case within any other measurement/experimental environment, the main task faced by an engineer is that of improving the performance of the instrumentation, i.e., its capability of extracting information about a certain phenomenon in the presence of sources of error, generatively called noise. This definition of performance encompasses two parameters describing the information extraction process in a restrictive sense. One of these parameters, called sensitivity, is a measure of the minimum phenomenon "intensity" for which information can be extracted with a non-zero probability; the other one is the resolution, which is a measure of the ability to distinguish between "adjacent" phenomena. There are situations when sensitivity can be improved only at the expense of resolu-

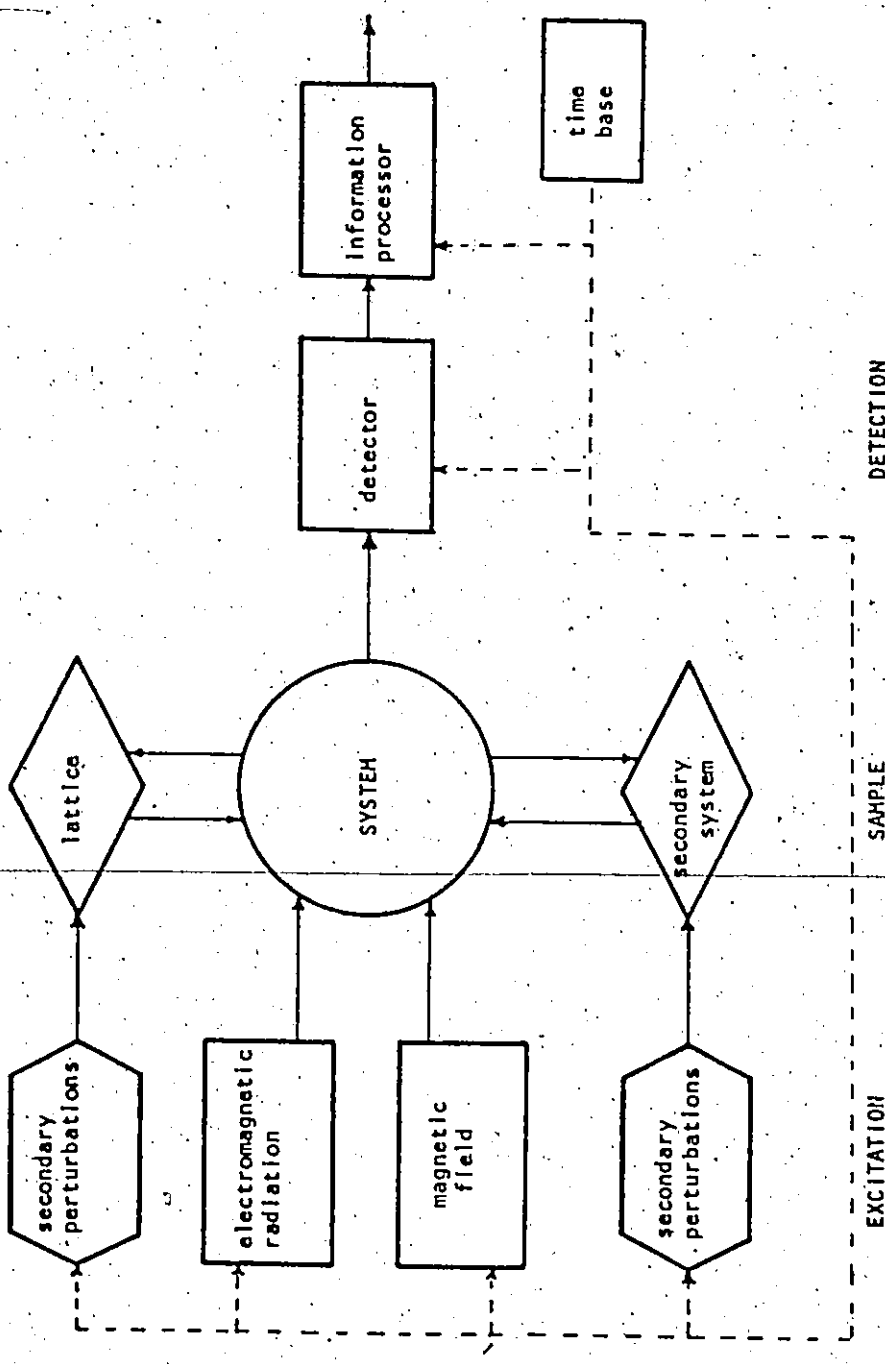


FIGURE 1.1 Block diagram of an EPR experiment

tion and vice-versa. The economic cost of extracting information is another very important parameter characterizing an experiment and, usually, a compromise must be achieved between performance and cost.

As a branch of magnetic resonance spectroscopy, EPR studies transitions between electronic energy levels whose energy separation is a function of an external magnetic field. An EPR experiment involves three basic units:

1) the system to be investigated;

2) the electromagnetic excitation; and

3) the detection system.

A block diagram of an experiment is given in Fig. 1.1, and Fig. 1.2 presents a schematic configuration of a superheterodyne EPR spectrometer. The complexity of EPR experiments can be understood only by considering both the variety of samples studied (which ranges from metals, semiconductors, dielectrics, gases, liquids and solids to complex biologic compounds) and that of the numbers and magnitude distribution of parameters involved: pseudostationary magnetic field 0 - 3T, with stability of 10^{-5} - 10^{-6} and modulation frequency of 10 - 10^{6}Hz; electromagnetic excitation field power of 10^{-12} - 10W at frequencies of 10^{6} - 10^{11}Hz (stabilized within 10^{-5} - 10^{-6}); sample temperature of 1 - 10^{3}K and pressure of 10^{-7} - 10^{5}atm; etc. Sensitivity and resolution requirements impose a time base which usually sweeps the pseudostationary magnetic field at a frequency of 10^{-4} - 10^{2}Hz, which means that the output signals have a very low "frequency content". Another characteristic of these signals is that they have known self-convolving shapes - usually Gaussian

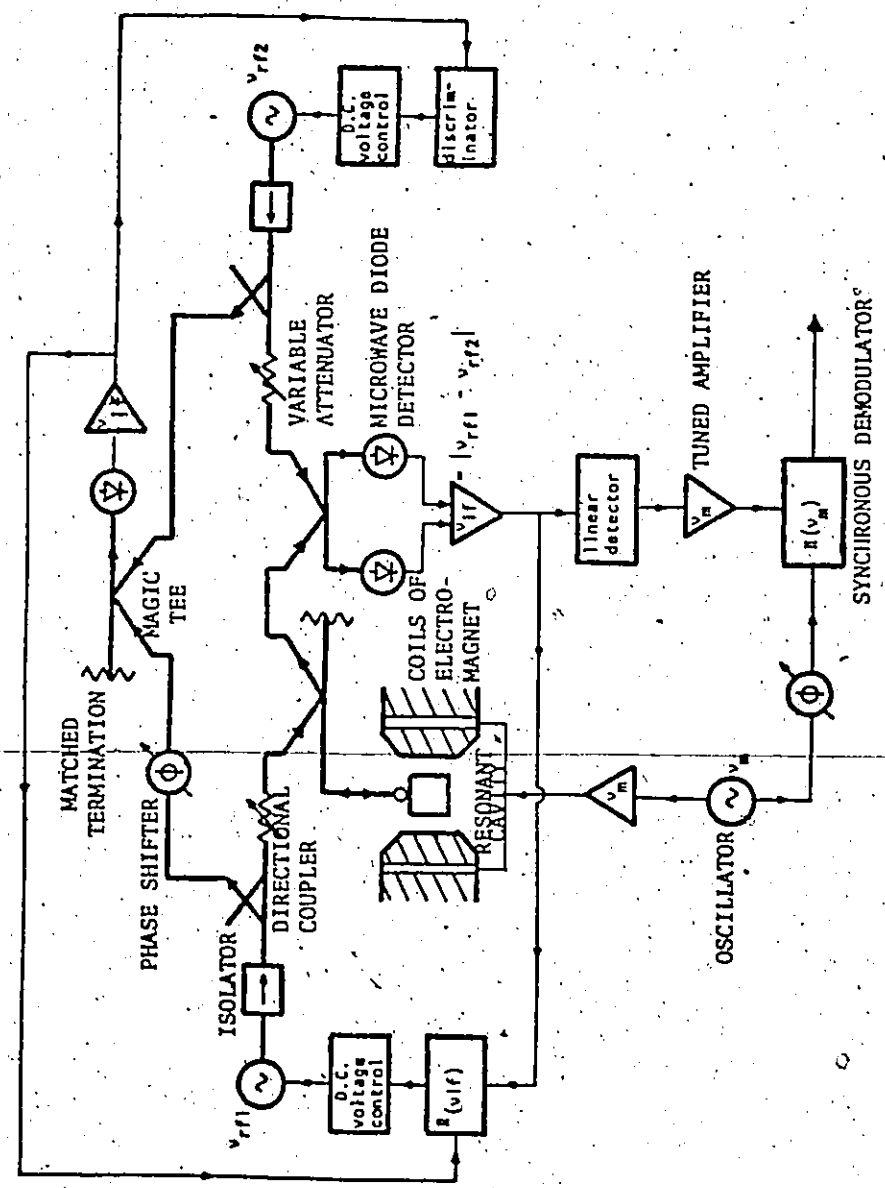


Figure 1.2. Block diagram of a superheterodyne spectrometer with frequency stabilization.

or Lorentzian bells - and they may be extremely complex, with up to 1000 individual peaks. Given this complexity, it is no wonder that it is practically impossible to determine a definite experimental configuration yielding maximum of information. In principle, the sensitivity may be improved by:

(a) increasing the signal via judicious use of various physical phenomena;

(b) decreasing the noise via careful design of the instrumentation and reduction of noise sources;

(c) separating the signal from noise, either by experimental design or by processing the signal at the output of the spectrometer.

Likewise, the resolution can be enhanced either by judicious design of the experiment and instrumentation, or by processing of the output signal.

The main idea behind my research work [22] was to employ digital signal processing methods in an attempt at solving some of the above-mentioned problems. To be more precise, the chosen tool was a digital Fourier analyzer, i.e. a general purpose computer and a FFT algorithm, which can serve a twofold purpose - that of measuring noise spectra in the low frequency region, and that of conveniently simulating certain linear signal processors. The choice of a digital Fourier analyzer was motivated by:

(a) the presence of easily repeatable signals at very low frequency regions where analogue techniques are costly and much less efficient;

(b) the complexity of instrumentation and the large variety of

parameters, requiring human intervention and extreme versatility which is best supplied by OFF-line processing with a general purpose computer;

(c) the low cost and availability of minicomputers and FFT algorithms. The actual instruments employed were a CDC 6400 and a PDP 11/20 computer for simulation and a HP 5450A digital Fourier analyzer for the noise measurements.

In order to achieve noise reduction, a systematic analysis of the microwave configuration and the noise sources has been made, and then the digital Fourier analyzer has been used to measure the spectral distribution of noise in order to determine experimentally the optimal instrumentation structure [23,24]. As a by-product, a new, in situ, method for measuring the performance of microwave detectors in the low frequency region has been developed [24]. The separation of signal from noise via processing of output signals can be viewed as a linear estimation problem [25,26] and, by making full use of the EPR signal peculiarities (e.g. the known self-convolving shape of the individual Gaussian or Lorentzian peaks), it has been shown [27,28] that DFT techniques could be very efficiently employed towards obtaining this objective. A linear filter which convolves its impulse response with the signal, can be simulated by a simple multiplication of its frequency transfer function with the FT of the signal, followed by an inverse FT. The decisive advantage of this procedure over ON-line digital processing is that it has a greater versatility, which is essential in the context of spectroscopic signals.

The problem of resolution enhancement has two major aspects. One is to separate signals of similar characteristics which, although concurrent and overlapping each other, have their energy (amplitude) peaks distributed in time. This problem may be modeled as an integral convolution equation

$$x(t) = \int_R u(\tau)\, k(t-\tau)\, d\tau, \quad t \in R, \tag{1.3.1}$$

where $x(t)$ is the experimental signal, $k(t)$ is the kernel shape function of the individual signal and $u(t)$ is the "spectrum" denoting the distribution of $k(t)$ signals in time. Again, a digital Fourier analyzer has been proven to be an efficient and valuable tool for solving such a problem [29].

The other aspect of resolution enhancement concerns the separation (analysis) of signals which occur simultaneously but have other parameters distributed. A poignant example is the case of signals having the same shape and time origin, but various width/duration characteristics, in which case the problem may be modeled as a multiplicative (sometimes called Mellin) convolution integral equation

$$x(t) = \int_{R_+} u(w)\, k(tw)\, dw, \quad t \in R_+, \tag{1.3.2}$$

where, again, $x(t)$ is the experimental signal, $k(t)$ is the kernel shape function and $u(w)$ is the "spectrum" denoting the distribution of signals' amplitudes versus signal widths. Although an elegant mathematical solution of (1.3.2) has been known for a long time [30], at the time of my M.Sc. thesis, no efficient numerical method had been devised. It was

then decided to dedicate at least part of the Ph.D. investigation towards elaborating a digital signal processing technique permitting the computation of u(w) [31]. It should be mentioned here that such problems of generalized spectral analysis of (multi-component) signals appear not only in magnetic resonance spectroscopy, but in numerous other fields of scientific investigation, e.g. electrical engineering, biology, chemistry, physics, radar, etc. ....

The other major interest concerns the theory and applications of orthogonal transforms other than the discrete Fourier transform of complex valued functions. At the time of my investigation into the use of DFT techniques for improving the performance of EPR spectrometers, a host of other discrete transforms began to be elaborated and discussed, predominantly the Walsh transform. Its net advantage consists in the fact that the Walsh functions form a complete set of functions which take only the values +1 or -1, such that the corresponding transform can be computed digitally very quickly and accurately. It was natural then to embark upon the study of the Walsh transform in the hope that it may prove beneficial to use it for solving the same type of problems as the ones mentioned above.

The investigations into, and the literature on, the subject of Walsh analysis can be grouped into two classes: one dealing with abstract mathematical theory, directed mainly at exploiting the similarity between the Walsh and Fourier transforms; the other one dealing with practical applications of the transform. Much of this latter work consisted, and still consists, in "rough" attempts at using the Walsh

Fourier transform (WFT) for the same problems in which DFT proved successful [32]. But, despite its much vaunted computational accuracy and speed, the WFT was proved experimentally to be much less successful in many cases. In a sense, this was a paradoxical situation, one of having a beautiful tool but without knowing precisely for which particular jobs it can be usefully employed. This is, in general, a rare occurrence in applied sciences where it is customary to be faced first with a problem, and only afterwards with the methods and tools for solving it.

It was decided then to direct investigative efforts not towards a particular application of the WFT, but towards obtaining a fundamental explanation of the experimental results reported in the literature, an explanation which would also point to those engineering problems which are optimally (or even suboptimally but advantageously) suited to be solved using WFT techniques. Three mathematical concepts play a fundamental role in this study into the inner structure of the WFT; they involve the relationship which convolution and group characters have with:

(a) harmonic differentiation,

(b) translation invariant systems (linear operators), and

(c) generalized frequency.

What differentiates the present study from others dealing with the same subject is that it is more comprehensive and general, not being limited to a particular species of Walsh functions and, more important, it is guided not by the desire to develop the theory for its own sake, but

for extracting its engineering significance. Concerning Walsh functions, the conclusion is simple and, in fact, has been mathematically evident all along: the WFT is nothing else but (isomorphically equivalent to) a multi-dimensional discrete Fourier transform. It follows that the only problems to which it may bring an optimal solution are those which may be modelled in a space of K-valued functions over a multi-dimensional domain (which, of course, has to be a LCA group [33]). There are, nevertheless, certain other problems in signal processing or system and control theory where the WFT might bring a solution which, although not optimal (in the sense of minimum error according to a given criterion of goodness, i.e. norm on the respective function space), may be cheaper and faster, and hence advantageous from a practical point of view.

A similar approach has been adopted with regard to signal processing and system theory in $K[G]$, where K is a finite field and G is a finite abelian group. The advantage of modelling signal filtering by convolutions in such a function space resides solely in the fact that under certain conditions it is possible to perform very fast and error free processing via Fourier transform techniques [34,35]. Again, the investigation has been directed towards obtaining a deeper understanding of convolution processors. A harmonic differential calculus is defined, thus permitting a "harmonic" state-space description of certain finite-valued signal processors [36]. It is hoped that this may facilitate advances towards the long range objective of freeing optimal digital processing from the constraints of referring it to some ideal optimum analogue model. Last, but not least, the meaning of Fourier coefficients

in finite fields is elucidated through the introduction of the concept of generalized frequency.

The word "frequency" is well entrenched in engineering vocabulary in connection with Fourier analysis and complex exponentials, and the appearance and widespread application in signal processing of discrete transforms other than the FT saw many an indiscriminate use of this word. At the same time, attempts have been made at defining a frequency concept general enough to encompass the ordering of all functions forming a complete set of orthogonal functions. This is how the concept of sequency [37], i.e. the ordering of Walsh functions according to their number of zero-crossings, came to be devised. The claim that it is a generalization of the concept of frequency played an important role (even if not always a positive one) in the development and proliferation of applications of the Walsh transform. Part of this dissertation is concerned with this claim and, more generally, with the problem of defining a concept of generalized frequency to be both consistent with the applied scientist's view of frequency as a measure of the speed of variation of a function, and general enough to permit a meaningful, unique ordering of any group of characters defining a Fourier transform. In order to attain this objective, it was felt necessary to determine first those features of the Euler functions which are essential to the understanding of the concept of frequency as used by engineers, so that they can be axiomatized in the definition of the concept of generalized frequency [38]. The analysis to be presented herein concludes that frequency should be viewed as (formally) the eigenvalue corresponding to the character as

eigenfunction of the harmonic differentiator, (locally) as the everywhere equal slope of the character's phase function, and (globally) as the "unit-time" winding number of the character's graph with respect to the origin of the K-field. It is also concluded that the sequency concept does not fulfill all of these conditions and that the claim that it represents a generalization of frequency is unfounded.

In one way or another, all the topics discussed in this dissertation are concerned with convolution and Fourier transforms on LCA groups. It was necessary to achieve a compromise between the number of topics and the depth of investigation into each of them. Chapter II is devoted to the presentation of Walsh functions and the essential features of Walsh analysis and dyadic (Gibbs) differentiation. Emphasis is placed on the isomorphism between the WFT and multi-dimensional discrete Fourier transforms. Chapter III is concerned with the theory of linear q-adic invariant systems, both from the point of view of input-output relationship and of harmonic state space description. The computational advantages of the WFT are analyzed in the first section of Chapter IV, while the other sections review some important applications of WFT techniques in signal processing and system theory, and point to those problems which are amenable to optimal solution via such treatment. Signal processing in finite fields is discussed in Chapter V. Chapter VI is devoted entirely to the concept of generalized frequency and a critical analysis of the concept of sequency. Chapter VII is concerned with the Mellin convolution; it presents a Fourier transform technique which leads to the solution of the problem of analyzing certain multicomponent

signals. It also led to an algorithm for a fast and accurate computation of Hankel transforms [33]. Besides general theory and some computer simulated examples, most of the chapter is devoted to the analysis of multicomponent exponential decays. Conclusions and ideas for further investigation are discussed in Chapter VIII.

CHAPTER II

WALSH FUNCTIONS AND THE WALSH-FOURIER TRANSFORM

2.1 INTRODUCTION

The interest in sets of orthogonal functions first arose in the context of mathematical physics, where they play a fundamental role in solving differential equations with boundary conditions (Sturm-Liouville problems); this is how the Fourier series made their first appearance, and this is how most of the other sets of orthogonal functions (Bessel, Lame or Mathieu, to name a few) came to be devised. But it was only at the beginning of this century that a comprehensive theory of orthogonal functions has been formalized. The first step was made by Hilbert (in his study of integral equations), who defined the important concept of a complete orthogonal system of functions. His work, continued by Riesz and Fisher, led to the foundation (Frechet and Schmidt) and development (mainly by Banach, Wiener and von Neumann) of functional analysis [40].

The basic idea behind the study of complete systems of orthogonal functions $\{\phi\}$ (on an interval I, in classical analysis) is that any function which is square summable on this interval can be approximated by means of an orthogonal projection on the subspace M_n spanned by the first n elements of $\{\phi\}$,

$$f \approx f_n = \sum_{i=1}^{n} \langle f, \phi_i \rangle \phi_i , \qquad (2.1.1)$$

with the essential property (Bessel inequality) that the mean square error can be made arbitrarily small by choosing n sufficiently large. Two other important advantages of function representation by such generalized Fourier series refer to invariance of the inner product, and the ease of extending the projection to a subspace M_{n+1} once the projection on M_n is known [11]. In addition, formulae for representing many functions with respect to various systems of orthogonal functions have been tabulated.

The study of non-continuous orthogonal functions made its debut with the work of Haar [41] who, at the beginning of this century, defined a complete system of orthonormal functions which can take only three values, and whose points of discontinuity are related to the dyadic rationals on the [0,1] interval (see Fig. 2.1).

$$\mathrm{har}_0(x) \equiv 1 \qquad \mathrm{har}_1(x) = \begin{cases} 1 & x \in [0,\tfrac{1}{2}] \\ -1 & x \in (\tfrac{1}{2},1] \end{cases}$$

$$\mathrm{har}_n^{(k)}(x) = \begin{cases} \sqrt{2}^{n-1} & x \in \left(\dfrac{k-1}{2^{n-1}}, \dfrac{2k-1}{2^n}\right) \\ -\sqrt{2}^{n-1} & x \in \left(\dfrac{2k-1}{2^n}, \dfrac{k}{2^{n-1}}\right) \\ 0 & x \in \left[0, \dfrac{k-1}{2^{n-1}}\right) \cup \left(\dfrac{k}{2^{n-1}}, 1\right] \end{cases} \qquad \begin{array}{l} n = 1,2, \ldots . \\ k = 1,2, \ldots, 2^{n-1} \end{array} \qquad (2.1.2)$$

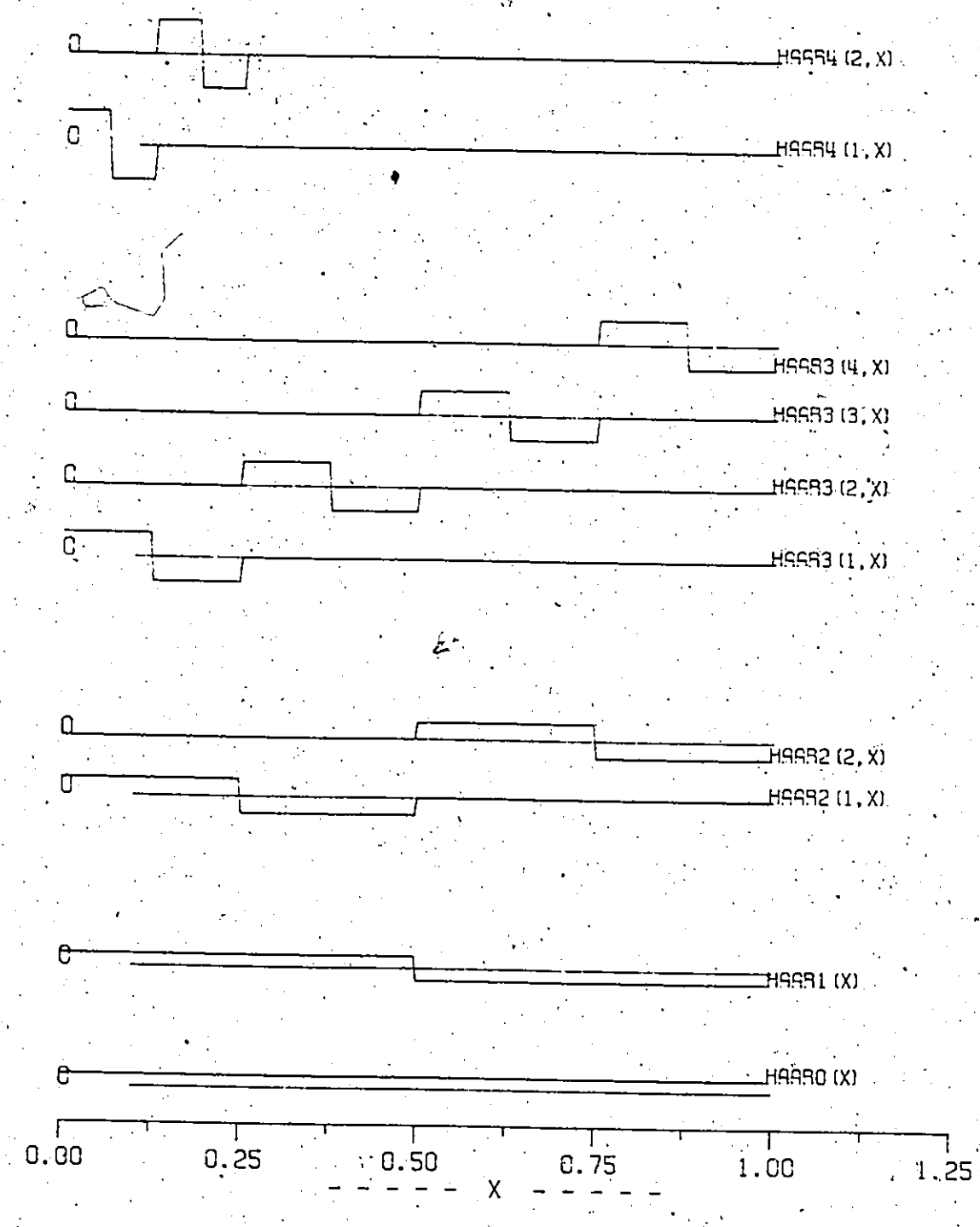


FIG.2.1 FIRST 10 HAAR FUNCTIONS ON UNIT INTERVAL

Given their obvious advantage for numeric commutation of Fourier coefficients, a very concentrated effort has been lately spent into devising and studying other such piecewise constant functions (Levy [42] calls them 'fonctions a paliers'), which are linearly independent, whose only points of discontinuity are the q-adic rationals, and which are, moreover, orthogonal and normal. A q-adic rational is any number x which can be brought to the form kq^{-n}, where q, k and n are integers. Any real number can be written in a radix q expansion

$$x = \sum_{j=-N_x}^{\infty} x_j q^{-j} \qquad \text{with } x_j \in Z_q \text{ or } P_q, \tag{2.1.3}$$

where it is customary to retain only the finite expansion for q-adic rationals (see [43] and [2] for a discussion of the relationship between q-adic rationals and the Cantor set). P_q denotes the set of q smallest non-negative integers. Levy also indicated the possibility of using a sequence of distinct integers as a radix for number representations in order to define the discontinuities of a yet more general form of piecewise constant orthogonal functions.

This chapter is concerned with the presentation of such a class of complete systems of orthogonal functions - the set of Walsh functions, first defined by Walsh in 1923 [44]. Besides having the properties mentioned above, the system originally defined by Walsh has the remarkable characteristics of (1) containing functions which are not only real-valued but simply bi-valued, and (2) being closed under the operation of pointwise multiplication, in the sense that the product of any

two Walsh functions is another Walsh function. In fact, as we shall see later, the Walsh functions, as any other set of orthogonal functions closed under multiplication, form a group [45,46]. The fact that their bi-valueness matches very well the nature of logic circuitry and digital computers explains the intensification of research into the theory and applications of Walsh functions after the mid-sixties.

There are many ways of defining the Walsh functions, as there are many ways of defining the exponential functions. One could use difference equations to obtain a recursive definition [44,47] or they can be defined as products of some other, simpler functions [48,49] or resulting from Hadamard matrices [50]. Also, they can be shown to be related to symmetry analysis [51,52] or, which is after all equivalent, they can be shown to be the characters of certain groups [53,54]. Alternatively, they may be defined as eigenfunctions of certain differential operators [55,56], or via polynomial considerations [57]; these latter definitions being just some other facets of their being the characters of a group. Again, the presentation may consider either the analog Walsh functions or only the discrete finite (in the sense of finite domain) functions, or both of them. From an engineering point of view, we are interested mainly in the discrete finite ones, which are implementable by digital computers and whose theory is mainly algebraic in nature, whereas mathematicians are attracted by the analog case, with all its wealth of analytic considerations. Historically, this latter case has been the one studied first (and exclusively until the late sixties), and the mathematicians' penchant on defining and analyzing the

Walsh functions as any other functions on an interval bears some blame for much of the unsuccessful engineering attempts at employing Walsh analysis. Since part of this disseration is concerned with analyzing these attempts, it was considered essential to present the Walsh functions not only through their fundamental definition as characters of a certain group, but also through the other definitions, even if only shortly. (At this point, it may be worthwhile mentioning that even in 1975 there were still people trying to find yet another "new" definition of Walsh functions [58].)

2.2 ANALYTIC DEFINITION OF WALSH FUNCTIONS

The initial definition of Walsh functions, due to Walsh [44] and, separately, to Kaczmarz [47], was confined to the interval [0,1] and was made via a recursive relation:

$$\phi_0(x) \equiv 1$$

$$\phi_1(x) = \begin{cases} 1 & x \in [0,1/2] \\ -1 & x \in (1/2,1] \end{cases}$$

$$\phi_2^{(1)}(x) = \begin{cases} 1 & x \in [0,1/4) \cup (3/4,1] \\ -1 & x \in [1/4,3/4] \end{cases}$$

$$\phi_2^{(2)}(x) = \begin{cases} 1 & x \in [0,1/4) \cup [1/2,3/4) \\ -1 & x \in [1/4,1/2) \cup (3/4,1] \end{cases}$$

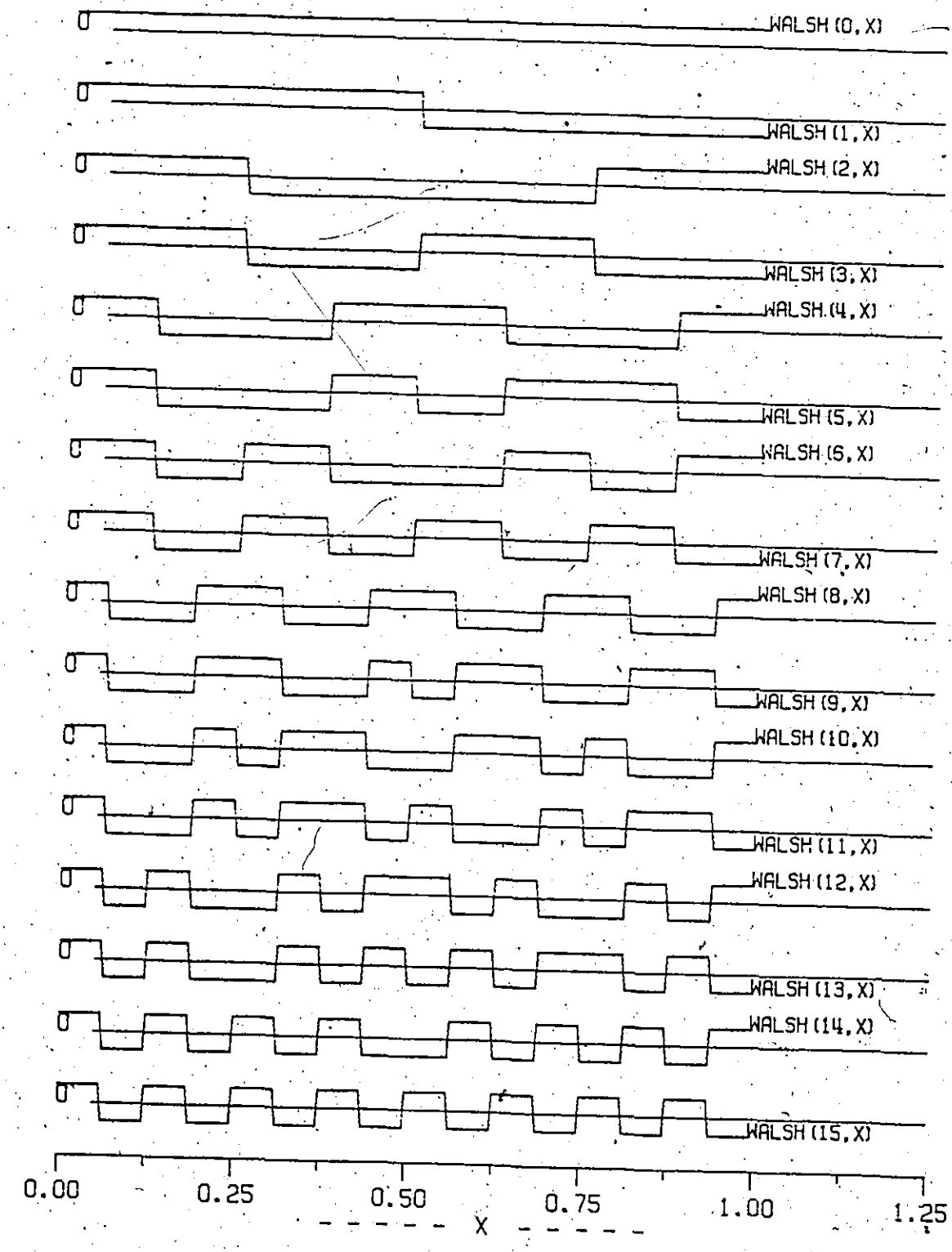


FIG. 2.2 FIRST 16 WALSH FUNCTIONS ON UNIT INTERVAL.
ARRANGED IN SEQUENCY ORDER

$$\phi_{n+1}^{(2k-1)}(x) = \begin{cases} \phi_n^{(k)}(2x) & x \in [0,1/2] \\ (-1)^{k-1}\phi_n^{(n)}(2x-1) & x \in (1/2,1] \end{cases}$$

$$\phi_{n+1}^{(2k)}(x) = \begin{cases} \phi_n^{(k)}(2x) & x \in [0,1/2) \\ (-1)^{k}\phi_n^{(k)}(2x-1) & x \in (1/2,1] \end{cases} \qquad (2.2.1)$$

$n = 1, 2, \ldots$

$k = 1, 2, \ldots, 2^{n-1}$

Starting from this definition, Walsh was able to prove the orthonormality and completeness of this system of functions and, hence, the possibility of using them for representing functions which are summable on the same interval I. The notation and equations used by Walsh may seem somewhat arbitrary until it is realized that a lexicographic ordering of the subscript (ordering by packets of functions) and superscript (ordering within the packet) leads to an ordering of the functions according to their number of discontinuities (which are called, sometimes, zero-crossings or sign-changes). Lately, it has become customary to index the Walsh functions by only one parameter; a tradition which will be followed here. In general, the functions will be noted by $wal(\lambda,x)$ where x is the free variable and λ is the indexing parameter. This parameter may belong to the set of positive integers (when the discussion is confined to a finite interval domain) (see Fig. 2.2) or to the set of real numbers (when discussion is generalized to the entire real line), or to a finite set of integers (when the discussion is confined to the discrete finite Walsh functions).

Sometimes, it is useful to use "integer letters" instead of λ when the index is obviously an integer. The importance (and the influence on the proliferation of engineering applications) of ordering the Walsh functions according to their number of discontinuities (thus resembling the ordering of the complete orthogonal system of sinusoids according to their number of zero crossings) will be analyzed at large in a later chapter. In the same context, it should be mentioned that Walsh was also the first to notice the alternately even and odd symmetry (with respect to the middle of the interval) of the functions he defined.

Walsh, who studied in detail the relationship between the system (2.2.1) and the Haar system (2.1.2) was not aware of the fact that his system of functions is the completion of the orthogonal system of Rademacher functions [59].

$$\mathrm{rad}_0(x) = 1, \quad \mathrm{rad}_k(x) = \mathrm{sgn}(\sin(2^{k-1}2\pi x)); \quad k = 1, 2, \ldots \tag{2.2.2}$$

It was Paley [48] who first realized this fact, although his completion of the Rademacher system led to a different ordering of the set of Walsh functions. To distinguish between the two orderings, we shall use, wherever necessary, the subscript, K, for Walsh-Kaczmarz ordering and the subscript, P, for Paley ordering. (See Fig. 2.3 for an illustration of three species (different orderings) of Walsh functions on the domain P_8.) Paley defined the Walsh functions as products of an arbitrary number of Rademacher functions,

$$\mathrm{wal}(k,x) = \prod_{k_i} \mathrm{rad}_{k_i}(x); \quad k = \sum_i 2^{k_i}. \tag{2.2.3}$$

$$W_K = \begin{bmatrix} + & + & + & + & + & + & + & + \\ + & + & + & + & - & - & - & - \\ + & + & - & - & - & - & + & + \\ + & + & - & - & + & + & - & - \\ + & - & - & + & + & - & - & + \\ + & - & - & + & - & + & + & - \\ + & - & + & - & - & + & - & + \\ + & - & + & - & + & - & + & - \end{bmatrix}$$

$$W_H = \begin{bmatrix} + & + & + & + & + & + & + & + \\ + & - & + & - & + & - & + & - \\ + & + & - & - & + & + & - & - \\ + & - & - & + & + & - & - & + \\ + & + & + & + & - & - & - & - \\ + & - & + & - & - & + & - & + \\ + & + & - & - & - & - & + & + \\ + & - & - & + & - & + & + & - \end{bmatrix}$$

$$W_P = \begin{bmatrix} + & + & + & + & + & + & + & + \\ + & + & + & + & - & - & - & - \\ + & + & - & - & + & + & - & - \\ + & + & - & - & - & - & + & + \\ + & - & + & - & + & - & + & - \\ + & - & + & - & - & + & - & + \\ + & - & - & + & + & - & - & + \\ + & - & - & + & - & + & + & - \end{bmatrix}$$

Figure 2.3. The Walsh-Kaczmarz, Walsh-Hadamard and Walsh-Paley matrices of order 8.

Later on, Toni [49] showed how to obtain the Walsh system in its original order directly from products of Rademacher functions (see also Yuen [60] for the relationship between the orderings). These functions, which are, as required, piecewise constant, real valued and with discontinuities occurring at the dyadic rationals, are called dyadic Walsh functions.

Levy [42] and Chrestenson [61] extended the discussion to systems of q-adic Walsh functions, i.e. complete orthonormal systems of piecewise constant functions which have their discontinuities at the q-adic rationals. This generalization from 2-adic to q-adic Walsh functions entails their loss of real-valuedness. Given $q \geq 2$, the root of unity, $w_q = \exp(2\pi i/q)$, and the expansions of (2.1.3) k and x, the Rademacher functions of order q can be defined as

$$\mathrm{rad}_k(x) = w_q^{[q^k x]}; \quad k = 0, 1, \ldots \tag{2.2.4}$$

where $[\cdot]$ represents the greatest integer function. Then, the q-adic Walsh functions are defined as

$$\mathrm{wal}(k,x) = \prod_i (\mathrm{rad}_{n_i}(x))^{k_i}; \quad k = \sum_i k_i q^{n_i}; \quad 0 \leq k_i \leq q;$$

$$n_i > n_{i+1}. \tag{2.2.5}$$

Alternatively, it can be written that

$$\mathrm{rad}_k(x) = w_q^{x_{k+1}}; \quad \mathrm{wal}(k,x) = w_q^z, \tag{2.2.6}$$

where

$$k = \sum_{i=-N}^{0} k_i q^{-i}$$

and

$$z = \sum_{i=0}^{N} k_{-i} x_{i+1},$$

so that it appears immediately that $\mathrm{rad}_k(x)$ has period q^{-k} and $\mathrm{wal}(k,x)$ has period 1. No distinction shall be made in notation between q-adic and dyadic Walsh functions because they have the same basic properties, with the exception of the fact that in the dyadic case the functions are (1) real-valued, with alternating even-odd symmetry instead of the usual complex-conjugate symmetry, (2) self-inverse, in the sense that the square of any 2-adic Walsh function is the unit Walsh function $\mathrm{wal}(0,x) \equiv 1$. Whenever possible, the presentation will be generalized to q-adic Walsh functions although, for obvious reasons, practically only the dyadic Walsh functions have been the subject of engineering applications.

Much of the mathematical work related to Walsh functions concentrated on studying the similarity between this system of functions and the system of sinusoids by establishing the analogues of the main results of the transformation theory of trigonometric series. In the main, these results indicate that:

- the Walsh functions form a complete system of orthonormal functions on the interval [0,1];

- this system is closed under multiplication;

$$\mathrm{wal}(k,x)\,\mathrm{wal}(i,x) = \mathrm{wal}(n,x); \quad \text{any } k, i \in \mathbb{P}, \quad k \oplus i = n \in \mathbb{P} \tag{2.2.7}$$

where $\oplus$ denotes dyadic addition, i.e. componentwise modulo 2 addition of the dyadic expansions of the respective numbers;

- for any summable function $f(x) \in L[0,1]$, the Walsh-Fourier series

$$S_\infty(x) = \sum_{k=0}^{\infty} c_k \mathrm{wal}(k,x); \quad c_k = \int_0^1 f(x)\, \overline{\mathrm{wal}(k,x)}\,dx \tag{2.2.8}$$

converges to f almost everywhere, the convergence being uniform in an interval of continuity or at a q-adic rational, with points of divergence located (if f is of bounded variation) at the q-adic irrational points of discontinuity; and, continuity of $f(x)$ is a sufficient condition for uniform Cesaro-summability of the Walsh-Fourier series;

- if $S_\infty(x)$ converges everywhere to an integrable function f, then it is the Walsh-Fourier series of f;

- there are difficulties in duplicating the results concerning the local-global duality of Fourier series and integrals, that is, the reflection of local features of f (such as smoothness) into global features of the transform (such as rapid decay at infinity); for instance, unlike in the trigonometric case, the only absolutely continuous functions on $[0,1]$ whose Walsh-Fourier coefficients are $0(1/k)$ are the constants. This problem of estimating orders of magnitude for Walsh coefficients is compounded by the difficulty of using classical differential calculus in relation to Walsh functions (it will be seen in Section 2.5 how this difficulty might be circumvented).

The proofs of these statements were rather cumbersome (being based on the Lebesgue integral on the unit interval), and the results concerning both continuity (and convergence) with respect to the usual

topology of the real line, as well as the closure of the Walsh system under multiplication appear to be somehow arbitrary. The situation has been radically changed since Fine, in 1946, had shown that the Walsh functions are the characters of a certain group, and Vilenkin [62] had established, in 1947, the general theory of Fourier transforms on topological groups. Although it was the group theory of analog Walsh functions on a finite interval or on the entire real line which was developed first, we shall start with the theory of finite discrete Walsh functions because it presents the greatest interest from the viewpoint of numerical applications, and because, being purely algebraic, it is not encumbered by any considerations of existence, continuity or convergence (of Walsh transforms). A more gradual, logical build-up of Walsh analysis, is thus enabled.

2.3 DISCRETE FINITE WALSH FUNCTIONS AND TRANSFORMS

By definition, the q-adic discrete finite Walsh functions are those elements of the space of complex-valued functions having a finite set of integers (let us say P_N) as domain, which are piecewise constant and which form a complete orthonormal basis closed under the operation of multiplication. It has been customary [63] to view them as the end result of the process of equidistant sampling the analog Walsh functions on a finite interval; indeed, this process can be laid on a proper mathematical foundation thanks to Kluvanec's abstract sampling theorem [65]. Unfortunately, this approach is unnecessarily cumbersome because it requires a preliminary mastering of the difficult analytic considera-

tions concerning Walsh analysis on an interval and makes, at best, very poor use of the powerful algebraic methods suitable for treating discrete Walsh functions; nor does it make clear the fundamental algebraic structure of the discrete Walsh transform.

Another approach [64] to discrete dyadic Walsh analysis is to introduce the dyadic Walsh functions directly through a (rather arbitrary) definition in terms of the binary representation of the indices,

$$\mathrm{wal}(k,i) = \prod_{r=0}^{n-1} (-1)^{k_r i_r} \quad \text{or} \quad \mathrm{wal}(k,i) = \prod_{r=0}^{n-1} (-1)^{(k_{n-r}+k_{n-r-1})i_r} \tag{2.3.1}$$

where $N = 2^n$, $i = 0, 1, \ldots, N-1$ $\quad k = 0, 1, \ldots, N-1$, and i_r and k_r are the binary bits of i and k; i.e.

$$i = \sum_{r=0}^{n-1} i_r 2^r.$$

Following this definition, the usual procedure is to provide "direct" proofs of their properties, viz.; their orthogonality,

$$\sum_{i=0}^{N-1} \mathrm{wal}(k,i)\mathrm{wal}(h,i) = N\delta_{kh} = \begin{cases} N\ ; & k = h \\ 0\ ; & k \neq h \end{cases} \tag{2.3.2}$$

their completeness,

$$\sum_{k=0}^{N-1} \mathrm{wal}(k,i)\mathrm{wal}(k,j) = N\delta_{i,j}, \tag{2.3.3}$$

their symmetry

$$\mathrm{wal}(k,i) = \mathrm{wal}(i,k) \tag{2.3.4}$$

and, finally, their product formula

$$\text{wal}(h,i)\ \text{wal}(k,i) = \text{wal}(h \oplus k, i); \quad h \oplus k = \sum_{r=0}^{N-1} 2^r (h_r + k_r) \bmod 2. \tag{2.3.5}$$

Then, the finite Walsh transform of a function (sequence) $f \in \mathcal{C}[\mathbb{P}_N]$ is defined as

$$\hat{f}(k) = \frac{1}{N} \sum_{i=0}^{N-1} f(i)\ \overline{\text{wal}(k,i)}\ ; \quad f(i) = \sum_{k=0}^{N-1} \hat{f}(k)\ \text{wal}(i,k) \tag{2.3.6}$$

after which follow (hard) proofs regarding the linearity (!) of this transform, its symmetry properties, its translation properties, etc..., finishing with the remark that the finite Walsh transform is not restricted to one dimension, the results being generalizable to any number of dimensions...

Since all results concerning linear transformations on finite dimensional normed linear spaces are independent of any considerations of continuity and so they are essentially algebraic, it is much more elegant and fundamental and, in the end, much more advantageous to introduce an algebraic definition of Walsh functions and the finite Walsh transform from the beginning. This can be done according to the general theory of finite discrete Fourier transforms [66], or in the context of unitary matrix theory. Because of the essential role played by the finite Walsh transform in the first part of this dissertation, its group theoretic definition will be the one presented in more detailed form.

As stated in the introduction, we consider the space of complex-valued functions whose domain is an abelian group; in this case, a finite abelian group G of order N. The finiteness of the group implies a discrete topology, with respect to which any complex-valued function on it is continuous [67]. In fact, the function space $\mathbb{C}[G]$ is an N-dimensional linear space, isomorphic to $\mathbb{C}^N$. Provided that the functions are restricted to be bounded (a reasonable engineering condition), this function space is trivially a Lebesgue L^p space which can be made into a (trivial) Hilbert space by introducing the inner product

$$\langle \cdot , \cdot \rangle : \mathbb{C}[G] \times \mathbb{C}[G] \to \mathbb{C}$$

defined as

$$\langle f_1, f_2 \rangle = \sum_{x \in G} f_1(x) \overline{f_2(x)}. \tag{2.3.7}$$

According to the structure theorem of group algebra [4], any finite abelian group can be represented as a direct sum (some call it product) of a finite number of cyclic groups. Our interest will be, in principal, directed to those groups which are (isomorphic to) a direct sum of n copies of the same cyclic group of order q. (The Levy generalization to the case of a sum of cyclic groups of various orders is conditioned only by notational difficulties and is of little interest in practice.) A cyclic group of order q is isomorphic to either of the groups $(\mathbb{P}_q, \dot{+})$ or $(\mathbb{Z}_q, \dot{+})$, where $\dot{+}$ denotes addition modulo q; for simplicity, we shall choose $\mathbb{P}_q$. Then, G is isomorphic to $\mathbb{P}_q^n$, in the sense that any element of G can be represented by an n-tuple

$\underline{x} = (x_0, \ldots, x_{n-1})$ and the group operation can be viewed as component-wise addition modulo q (also denoted as $\dot{+}$). From now on, we shall call this group the (finite) q-adic group $(G_q, \dot{+})$. As such, G_q is an n-dimensional module over $\mathbb{P}_q$ [4], where $\mathbb{P}_q$ is considered a ring, with addition and multiplication modulo q. If q is a prime number, $\mathbb{P}_q$ is a field and G_q is an n-dimensional vector space. Any element of G_q can then be represented as a linear combination of n linearly independent n-tuples. Such a set of n-tuples $\{\underline{\varepsilon}_i\}_{i=0}^{n-1}$ is called a basis, and the $\mathbb{P}_q$-scalars x_i which determine

$$\underline{x} = \sum_{i=0}^{n-1} x_i \underline{\varepsilon}_i \tag{2.3.8}$$

are called the coordinates of $\underline{x}$ with respect to the basis $\{\underline{\varepsilon}_i\}$. Addition of elements of G_q and their multiplication by $\mathbb{P}_q$-scalars can be performed coordinate-wise with respect to any such basis. Three examples of bases of n-tuples are:

$$\{\underline{e}_i\}_{i=0}^{n-1}; \quad e_{ik} = \delta_{i,k}; \quad k = 0, 1, \ldots, n-1 \tag{2.3.9a}$$

$$\{\underline{d}_i\}_{i=0}^{n-1}; \quad d_{ik} = \begin{cases} q-1; & k > i \\ 0; & k \leq i \end{cases}; \quad k = 0, 1, \ldots, n-1 \tag{2.3.9b}$$

$$\{\underline{h}_i\}_{i=0}^{n-1}; \quad h_{ik} = \delta_{i,n-k-1}; \quad k = 0, 1, \ldots, n-1 \tag{2.3.9c}$$

where it is evident that $\{\underline{h}_i\}$ is, apart from an inverse ordering, identical to $\{\underline{e}_i\}$. Figure 2.4 illustrates these bases for a 3- and 4-dimensional dyadic space. The influence which these bases have on the ordering of discrete Walsh functions will be discussed in Chapter VI. A

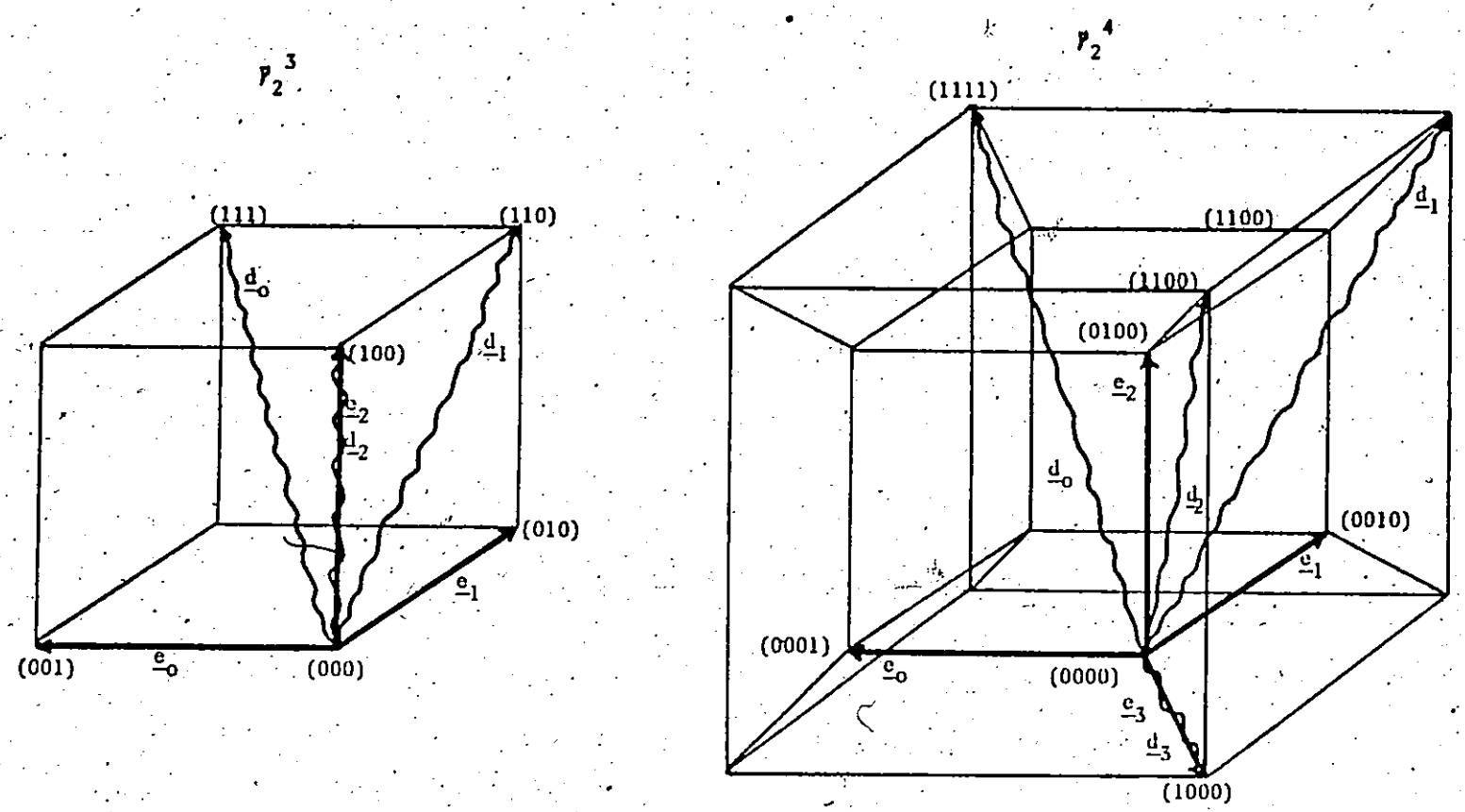


Figure 2.4. Illustration of the vector bases $\{\underline{d}_i\}$ and $\{\underline{e}_i\}$ in the Boolean spaces P_2^3 and P_2^4.

more detailed analysis of the G-module for $q=2$ can be found in [68]. For the moment, we shall limit ourselves to mentioning that the mapping

$$(\cdot\,,\cdot) : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{C}$$

defined by

$$(\underline{x}\,,\underline{y}) = \sum_{i=0}^{n-1} x_i y_i \qquad (2.3.10)$$

fulfills all the requirements for being an inner product on $\mathbb{F}_q$. The same symbol $(\cdot\,,\cdot)$ is used for the mapping defined by

$$(\underline{x}\,,\underline{y}) = \sum_{i=0}^{n-1}{}^{\circ} (x_i\, y_i) \bmod q$$

where $\sum^{\circ}$ implies summation modulo q; this mapping being a pseudo-inner product because $(\underline{x}\,,\underline{x}) = 0$ does not imply $\underline{x} = \underline{0}$.

The so-called sum-invariance lemma [69] $\sum_{x\in G_q} f(\underline{x}+\underline{y}) = \sum_{x\in G_q} f(\underline{x})$ appears as an immediate consequence of the fact that G is a group, and shows that by assigning to each element of G an equal measure we obtain a Haar (translation invariant) measure. Since G is both discrete and compact, we are faced with a contradiction between the desire of normalizing the Haar measure to be 1 for both the entire group and each point of it. The solution is either to assign each point a measure $1/N$ or to accept a total measure of N (the "problems of measure normalization" on both dual groups can be balanced by assigning a measure $N^{-1/2}$ to each point of both groups).

The function space $\mathcal{C}[G_q]$ appears as;

- a commutative Banach algebra $(\mathcal{C}[G_q],+,\cdot,1)$ if multiplication is defined pointwise and the unit is the function identical 1 everywhere; and

- a commutative Banach algebra $(\mathcal{C}[G_q],+,\overset{\circ}{*},\delta)$ if multiplication is defined as the convolution product

$$(f_1 \overset{\circ}{*} f_2)(\underline{x}) = \sum_{\underline{y}\in G_q} f_1(\underline{y})\ T^{\underline{x}}_{\leftarrow} f_2(\underline{y}) = \langle T^{\underline{x}}_{\leftarrow} f_2\ ,\ \overline{f}_1 \rangle, \tag{2.3.11}$$

and the role of unit is played by the function

$$\delta(\underline{x}) = \begin{cases} 1 & \text{if } \underline{x} = \underline{0} \\ 0 & \text{otherwise} \end{cases} \tag{2.3.12}$$

The characters of the group G are defined as the homomorphisms of G into the unit magnitude complex numbers (see Eqn. 1.2.5). They form the dual group Γ. Because G is finite, Γ is isomorphic to it, $\Gamma \sim G_q$. Since G_q is isomorphic to $\mathcal{P}_q^n$, $G_q \sim \mathcal{P}_q^n$, it follows (Eqn. 1.2.8) that a character of G_q can be represented by a product of characters of $\mathcal{P}_q$,

$$\chi(\underline{\lambda},\underline{x}) = \prod_{i=0}^{n-1} \chi(\lambda_i,x_i);\quad \underline{\lambda} = (\lambda_0, \ldots, \lambda_{n-1}) \in \Gamma;$$

$$\underline{x} = (x_0, \ldots, x_{n-1}) \in G. \tag{2.3.13}$$

A character of $\mathcal{P}_q$ can be represented as

$$\chi(\lambda_i,x_i) = \exp(j2\pi\lambda_i x_i/q) = w_q^{\lambda_i x_i}$$

(see Eqn. 1.2.17c)

such that

$$\chi(\underline{\lambda}, \underline{x}) = \prod_{i=0}^{n-1} w_q^{\lambda_i x_i} \triangleq w_q^{\sum_{i=0}^{n-1} \lambda_i x_i} = w_q^{\sum_{i=0}^{n-1} (\lambda_i x_i) \bmod q}$$

$$= w_q(\underline{\lambda}, \underline{x}) \qquad (2.3.14)$$

Note: Concerning Harmuth's assertion [88] that, unlike the sinusoid functions, the Walsh functions cannot be expressed in terms of a product between the index λ and the free variable x, it can be seen that a definition is possible in terms of inner products. The difference between sinusoids and Walsh functions is that in the latter case the domain is not one- but multi- dimensional (a fact which is stressed in Chapters IV and VI). In fact, in neither case is the word product quite proper because λ and x belong to different sets, such that the operation $(\underline{\lambda}, \underline{x})$ should be defined as a mapping: $\Gamma \times G_q \rightarrow \mathcal{C}$ (or $\mathcal{P}_q$) producing the "phase" (see Chapter VI) of the character.

The liberty taken for dispensing with this "hair splitting" is warranted by the isomorphism between G_q and Γ and by the force of tradition which, for example, finds acceptable the products of Eqn. 1.2.17. Another fact worth mentioning here refers to the fact that, since $\underline{\lambda}$ and $\underline{x}$ belong to different groups which are both isomorphic to $\mathcal{P}_q^n$ the "product" $(\underline{\lambda}, \underline{x})$ may be calculated without restricting the n-tuples $\underline{\lambda}$ and $\underline{x}$ to be represented with respect to the same basis; hence, the variety of species (i.e. different orderings of sets) of Walsh functions.

As any other characters, the functions defined above enjoy all the group (symmetry) properties described by Eqns. 1.2.6. In addition, as characters of a finite group (which is isomorphic to its dual),

Th. 2.4.a they form a complete "orthonormal" basis in the Hilbert space $\mathcal{C}[G_q]$. Indeed, the set of characters is complete because they form a set of N linearly independent functions (with respect to either of the spaces $\mathcal{C}[G_q]$ or $\mathcal{C}[\Gamma]$),

$$\sum_{\underline{x}\in G_q} \chi(\underline{\lambda},\underline{x}) = N\,\delta(\underline{\lambda}), \quad \sum_{\underline{\lambda}\in\Gamma} \chi(\underline{\lambda},\underline{x}) = N\,\delta(\underline{x}); \tag{2.3.15}$$

and the characters are mutually orthogonal because,

$$\langle \chi(\underline{\lambda}_1,\underline{x}) \,,\, \chi(\underline{\lambda}_2,\underline{x})\rangle = \langle \chi(\underline{\lambda}_1 \overset{\circ}{+} \underline{\lambda}_2,\underline{x}) \,,\, 1\rangle$$

$$= \sum_{\underline{x}\in G_q} \chi(\underline{\lambda}_1 \overset{\circ}{+} \underline{\lambda}_2,\underline{x}) = N\,\delta(\underline{\lambda}_1 \overset{\circ}{+} \underline{\lambda}_2). \tag{2.3.16}$$

Then,

Th. 2.4.b any function $f \in \mathcal{C}[G_q]$ can be represented as a linear combination of characters, the sequence of coefficients representing the Fourier transform $\hat{f}(\underline{\lambda})$.

$$f(\underline{x}) = \sum_{\underline{\lambda}\in\Gamma} \hat{f}(\underline{\lambda})\chi(\underline{\lambda},\underline{x}) = \langle \hat{f}(\underline{\lambda}) \,,\, \chi(\underline{\lambda},\underline{x})\rangle \triangleq F^{-1}\{\hat{f}(\underline{x})\}, \tag{2.3.17a}$$

$$\hat{f}(\underline{\lambda}) \triangleq F\{f(\underline{x})\} \triangleq \frac{1}{N} \sum_{\underline{x}\in G_q} f(\underline{x})\overline{\chi(\underline{\lambda},\underline{x})} = \langle f(\underline{x}),\, \chi(\underline{\lambda},\underline{x})\rangle \,. \tag{2.3.17b}$$

Immediate corollaries of these theorems indicate, as expected, the invariance of the inner product under this Fourier transform mapping (Eqn. 1.2.15) and the existence of the same translation invariance properties as those presented by Eqns. 1.2.12-1.2.13. Whence,

Th. 2.4.c the FT of Eqn. 2.3.17 realizes an isomorphic mapping between the Banach algebras $(\mathcal{C}[\mathbb{P}_q^n],+,\cdot,1)$ and $(\mathcal{C}[\mathbb{P}_q^n],+,\natural,\delta)$,

$$f \natural f_2 \underset{F^{-1}}{\overset{F}{\rightleftharpoons}} \hat{f}_1 \hat{f}_2 \qquad f_1 \cdot f_2 \underset{F}{\overset{F^{-1}}{\rightleftharpoons}} \hat{f}_1 \natural \hat{f}_2 \tag{2.3.18}$$

(where, for invariance of the inner product, we ignore the normalization factor 1/N).

The isomorphic bijective relationship between the characters $\chi(\underline{\lambda},\underline{x})$ (and their corresponding Fourier transform) on one side, and the discrete finite Walsh functions (Walsh transform) on the other side, appears evident once we realize that it is possible to devise a mapping (many, in fact) $\alpha : \mathbb{P}_q^n \to \mathbb{P}_{q^n}$, e.g.

$$\alpha(\underline{x}) \triangleq \sum x_i \, q^i \triangleq x \in \mathbb{P}_{q^n} \tag{2.3.19}$$

(with an obvious inverse α^{-1}) which performs an isomorphic correspondence between the group $(\mathbb{P}_q^n,\natural)$ and the set of first $N=q^n$ non-negative integers on which there is defined an addition $\natural$ (also!)

$$x \natural y = \alpha(\underline{x} \natural \underline{y}) = \alpha(\alpha^{-1}(x) \natural \alpha^{-1}(y)) \tag{2.3.20}$$

This mapping is by no means unique. Gibbs and his colleagues have called

it a contraction (expansion - for the inverse) mapping, and they provided various generalizations of this concept [70,71,72,73]. Most elegantly, the mapping α can be viewed as a norm on $\mathbb{P}_q^n$, associating with each $\underline{x}$ n-tuple a distinct length x. The fact that we want distinct norms for distinct n-tuples exludes the usual isotropic ("taxicab") norm $\sum_{k=0}^{n-1} |x_k|$, which assigns to each of the basis vectors $\underline{e}_i$ or $\underline{h}_i$ an equal (=1) norm. The contraction mappings α_1 and α_2,

$$\alpha_1(\underline{x}) = \sum_{i=0}^{n-1} q^i x_i$$

$$\alpha_2(\underline{x}) = \sum_{i=0}^{n-1} q^{n-i-1} x_i, \qquad (2.3.21)$$

which are seen to be anisotropic norms (assigning to each basis n-tuple $\underline{\varepsilon}_i$ a distinct length) realize a bijective correspondence between $\mathbb{P}_q^n$ and $\mathbb{P}_{q^n}$, and so they are the ones most often used. Then, if we replace $\underline{x}$ and $\underline{\lambda}$ as elements of $\mathbb{P}_q^n$ by x and λ as elements of $\mathbb{P}_{q^n}$, we obtain the entire theory of discrete finite q-adic Walsh functions. Equations 2.3.1-2.3.6 appear as a particular case for $q=2$; then, $w_q = -1$ and $\mathrm{wal}(\lambda,x)$ is real, bivalued and self-inverse.

What has been gained by adopting this definition of discrete Walsh functions and the corresponding Walsh transform? For one thing, generality is achieved through an embedding of Walsh analysis in the larger class of Fourier analysis on (finite) groups; secondly, the arbitrariness of definition 2.3.1 is avoided; and; very important, the intrinsically multi-dimensional character of the Walsh transform is

brought forward. This latter aspect is fundamental to the problems discussed in Sections 3.4, 4.2, 4.3 and 6.5 and will be considered anew there. Also, to be mentioned at this point, is the fact that the finiteness of G and Γ (apart from their multi-dimensionality) precludes any possibility of a natural ordering of their elements, with the consequence that there is no unique ordering of the set of Walsh functions. There are as many orderings as there are permutations [66,74]. However, no particular ordering has any bearing upon the fundamental properties discussed above. There are, nevertheless, extraneous features like, e.g. computational advantages [75,76] which may distinguish certain species of Walsh functions.

2.3 THE GROUP THEORETICAL DEFINITION OF ANALOG WALSH FUNCTIONS

The introduction by Fine [53,54] (for the dyadic case) and Chrestenson [61] (for the q-adic case) of a group theoretic approach to Walsh analysis has freed it from the stranglehold imposed by the consideration of an interval as the domain of definition for Walsh functions. Instead, as it will appear presently, it is more natural to consider that, by definition,

Def. 2.4.a the Walsh functions are the characters (hence, their continuity) of the q-adic group G_q represented by the set of all sequences $\underline{x} = (0, \ldots, 0, x_{-N_x}, \ldots, x_0, x_1, \ldots)$, $x_i \in P_q$, with termwise addition modulo q.

The "infinite" dimensionality of G_q requires very careful treatment. The translates of the neighbourhood of the identity V_0^r (defined as the set of all sequences $\underline{x}$ with $x_i=0$ for $i < r$) provide an open basis and, hence, a topology for G_q. Then, G_q may be shown [53] to be a totally disconnected LCA topological group. Thus, the Fourier transform on the q-adic group G_q (see Eqns. 1.2.5-1.2.17). The equivalence between these characters (and their corresponding Fourier transform) on one side, and the Walsh functions (and the Walsh transform) on the other side may be established in two stages.

First, it is necessary to establish the almost everywhere (a.e.) one-to-one correspondence between the set of sequences $\underline{x}$ and the set of non-negative reals R_+. This correspondence is achieved via the (contraction) Fine mapping [53,71] $\alpha: G_q \to R_+$

$$\alpha(\underline{x}) = x = \sum_{i=-N_x}^{\infty} x_i \cdot q^{-i} \tag{2.4.1}$$

and its inverse $\alpha^{-1}: R_+ \to G_q$, which is restricted to selecting only the right ended (zero-stationary to the right) sequences for q-adic rationals. Moreover, this mapping is an isomorphism between the group G_q and the set R_+ to which we attach an operation of addition (also denoted by $\dotplus$) defined as

$$x \dotplus y = \alpha(\underline{x} \dotplus \underline{y}) = \alpha(\alpha^{-1}(x) \dotplus \alpha^{-1}(y)). \tag{2.4.2}$$

To achieve an isomorphic a.e. bijective mapping between G_q and the group of all real numbers, it is sufficient, for $q > 2$, to require that

$x_i \in Z_q$ instead of F_q. Alternatively, this correspondence can be realized even for $q=2$ and/or $x_i \in F_q$ at the expense of adding a sign term $s_x \in F_2$ to every sequence $\underset{\sim}{x}$ such that $\underset{\sim}{x} = (s_x; \ldots, x_{-1}, x_0, \ldots)$ the operations between sign terms being carried modulo 2. The mapping $\alpha : G_q \to R_+$ defines a norm and, hence, a metric (continuous) on G_q [77, 78] which, in its turn, induces a topology on G_q equivalent to the original. Indeed, $d(\underset{\sim}{x},\underset{\sim}{y}) = \alpha(\underset{\sim}{x} \ominus \underset{\sim}{y})$ has all the characteristics of a distance function. The existence of the mappings α and α^{-1} leads immediately to the invariance of the Lebesgue integral on R_+ (or R, or any subgroup of it, as for example, the interval $[0,1] = I$) with respect to the q-adic translation operator T^y where $T^y f(x) = f(x \ominus y)$,

$$\int f(x \ominus y)\, dx = \int f(x)\, dx. \tag{2.4.3}$$

Because of this correspondence $G_q \to R$, it is not always necessary to carry on distinguishing between real numbers x and sequences $\underset{\sim}{x}$; in these cases, the underbar will not be used. Notwithstanding this remarkable correspondence, it is evident that continuity of functions on G_q does not imply continuity on R (with its usual topology), fact which explains why the Walsh functions which are, by definition, continuous with respect to G_q (and are infinitely harmonic differentiable [79,80]) seem to be anything but continuous or differentiable when viewed as functions of a real variable.

The second stage for showing the equivalence between the characters on G_q and the Walsh functions entails proving the fact that

Th. 2.4.a the characters of the subgroup $(I,\dot{+})$ are identical with the Walsh functions (of Sec. 2.2, Eqns. 2.2.1-2.2.6).

The proof runs as follows:

(1) The subgroup $(I,\dot{+})$ is a-isomorphic to the subgroup G_q^* containing all the sequences $\underline{x}$ with $x_i=0$ for $i \leq 0$; hence, it is the complete direct sum of a countably infinite set of copies of P_q. Since P_q is a compact abelian group, so is $(I,+)$.

(2) The dual group of a complete direct sum of compact abelian groups $\{G_\kappa\}$ is the direct sum of the corresponding dual groups Γ_κ, so that a character is representable by a finite product of characters of G_κ,

$$\chi(\lambda,x) = \prod_\kappa \chi(\lambda_\kappa,x_\kappa) \qquad (2.4.4)$$

(3) Since G_κ is isomorphic to P_q, its characters can be expressed as $w_q^{\lambda_\kappa x_\kappa}$, which appear as Rademacher functions if considered with respect to the entire group G_q^*, i.e. the interval I.

The characters thus obtained form the denumerable set of Walsh functions on the unit interval; regarded as functions on the real line, they are the periodic Walsh functions of period 1. Their index spans the set of (non-negative) integers P or, alternatively, the set of q-adic rationals (less than 1) [70,79]. As the characters of a (sub-) group, they form a group under multiplication. More generally, Fine [46, 81] showed that

(1) any orthonormal system of functions (on a measurable space) which is simultaneously a group under multiplication can be essentially derived from characters of compact abelian groups; and

(2) if G is any countable abelian group, there exists on the unit interval I an orthonormal group (of functions) G' isomorphic to G, and if G is infinite, then G' is complete in $L^2[I]$.

Since the denumerable set of Walsh functions on the interval I appears as its subgroup (of functions of period 1), the group of characters on $G_q \sim (\mathbb{R}_+, \dot{+})$ is called, by extrapolation, the group of generalized Walsh functions [54]; they shall be also denoted as $\mathrm{wal}(\lambda,x)$. Two important equations relating them to the periodic Walsh functions are

$$\mathrm{wal}(mq^{-n},x) = \mathrm{wal}(mq^{-n},x+kq^{-n}); \quad k,m,n \in \mathbb{P}, \tag{2.4.5a}$$

and

$$\mathrm{wal}(\lambda,x) = \mathrm{wal}([\lambda],x)\ \mathrm{wal}(\lambda,[x]). \tag{2.4.5b}$$

In his disseration [82], Pichler provided a different but interesting definition of generalized dyadic Walsh functions by making full use of the even/odd partition of the denumerable set of Walsh functions. In the name of similarity with the cosine and sine functions, he called the even and odd Walsh functions cal and sal functions respectively, and then extended them periodically and aperiodically (the term aperiodic extension is, indeed, a tour de force - it is associated with the use of a counterpart of the α^{-1} mapping which relates a dyadic rational to its infinite dyadic representation). In what concerns the Walsh trans-

form, Pichler (and, after him) Harmuth partitions it into two real parts according to the even (cal) and odd (sal) symmetry. This procedure becomes less appealing for $q > 2$, when Walsh functions are complex q-valued, and it will not be followed here. A polynomial counterpart of Pichler's definition was given by Butin [57].

An interesting expression relating Walsh functions to a "product" of the index and the free variable has been proven by Fine for the dyadic case. It can easily be generalized to any q-adic case for which q is a prime number,

$$\mathrm{wal}(\lambda,x) = \mathrm{wal}(1,\lambda^{\circ}x) = w_q^{\,z} \;;\; z = \sum_i \lambda_i \, x_{1-i}, \tag{2.4.6}$$

because then the set of sequences $\underline{x}$ is not only a group G_q under addition $\overset{\circ}{+}$, but also a field F_q if multiplication $^{\circ}$ is defined polynomially, as if $\underline{x}$ were a formal power series $\underline{x} = \sum_i x_i \, \lambda^i$. The mapping α realizes an isomorphism between F_q and $(\mathbb{R}_+, \overset{\circ}{+}, ^{\circ})$. The Note to Eqn. 2.3.14 explains why the word "product" has been put in quotation marks above.

Following the precepts of harmonic analysis summarized in the introduction, the Walsh-Fourier transform (WFT) of a L^p-summable function is defined as

$$\underline{\text{Def. 2.4.b}} \quad W\{f\} \triangleq \overset{\circ}{f}(\lambda) \triangleq \int f(x) \, \overline{\mathrm{wal}(\lambda,x)} \, dx, \tag{2.4.7a}$$

with an inverse transform computable (when it exists) as

$$\mathcal{W}^{-1}\{\hat{f}\} \triangleq f(x) = \int \hat{f}(\lambda)\ \mathrm{wal}(\lambda,x)\ d\lambda, \qquad (2.4.7b)$$

where the domain of integration is either G_q or R_+ or R. If the x-domain is the unit interval, λ spans the set of integers, and we are faced with the Walsh series of Eqn. 2.2.8. As any Fourier transform on a group, the WFT is characterized by certain translation invariance properties which are essential in our study of linear systems and numeric applications of convolution transforms.

Th. 2.4.c $\mathcal{W}\{T^y f(x)\} = \mathrm{wal}(\lambda,y)\ \mathcal{W}\{f(x)\} = \mathrm{wal}(\lambda,y)\ \hat{f}(\lambda).$ (2.4.8a)

The proof is trivially based on Eqns. 2.4.3 and the multiplication rule of Walsh functions as characters. Similarly, one can prove

Th. 2.4.d $\mathcal{W}\{\mathrm{wal}(y,x)\ f(x)\} = \hat{f}(\lambda \oplus y).$ (2.4.8b)

As for the counterpart of the classical convolution theorem, if we define a q-adic convolution product $\circledast : \mathbb{C}[G_q] \times \mathbb{C}[G_q] \to \mathbb{C}[G_q]$ as

Def. 2.4.c $(f \circledast g) \triangleq \int f(y)\ g(x \ominus y)\ dy = \int f(y)\ T_+^x g(y)\ dy,$ (2.4.9)

it can be immediately seen that this operation is associative and commutative and that the Walsh-Fourier transform is an isomorphism between q-adic convolution and pointwise multiplication Banach algebras,

Th. 2.4.d $\mathcal{W}\{f \circledast g\} = \mathcal{W}\{f\}\ \mathcal{W}\{g\}.$ (2.4.10)

The proof, which makes use of Th. 2.4.c, of Fubini's theorem and of Pontryagin's duality is simple. In addition, as already mentioned, if the function space has a domain which is measurable and a subgroup of $R(,\oplus)$,

Th. 2.4.e the corresponding Walsh functions are integrable and form a (complete) orthonormal basis for the Hilbert space of complex-valued functions on that domain.

The denumerable set of periodic Walsh functions can be embedded in a much larger class of orthogonal functions. This becomes more evident when they are presented as the rows (columns) of certain orthogonal matrices [83,84,86,86,87] obtainable through Kronecker products. Regarding the dyadic discrete Walsh functions, Harmuth [88] has shown that they can be related to Hadamard matrices whose order is a power of 2. A Hadamard matrix of order N is an $N \times N$ matrix H_N with ± 1 elements, which obeys an orthogonality equation of the form $H_N \cdot H_N^T = N \cdot I_N$, where H^T denotes the transpose of H and I is the identity matrix. More interestingly, Crittenden [89] and, recently, Larsen and Madych [50] have shown how to obtain the analog Walsh functions from countable Kronecker-like products of Hadamard matrices. Of course, not all the sets of orthogonal functions obtained likewise are Walsh-type functions because not all of them are closed under multiplication. Although it may be possible to prove most (if not the entire) Walsh theory starting from a matrix definition, it would be a rather contortuous way of revealing the symmetry and translation properties of Walsh functions, whereas they are so readily and fundamentally evident once a group theoretic definition is adopted. Also, although it is possible to define the Walsh functions as eigenvectors of certain matrices [90,72] the essential fact that these matrices represent linear operators resembling the Newton-Leibniz differential operator is not so evident as when a

group definition is forwarded. The presentation of Walsh functions as eigenfunctions of certain linear operators forms the subject of the next section.

2.5 WALSH FUNCTIONS AND HARMONIC DIFFERENTIATION

J.E. Gibbs, in 1967, had the ingenious idea of emulating the classical approach of defining systems of orthogonal functions from differential equations with boundary conditions by constructing a linear operator with properties similar to those of a Newton-Leibniz differentiator so that the (discrete dyadic) Walsh functions appear as its eigenfunctions. The theory of such operators (called "logical", "dyadic" or "Gibbs" differentiators in connection with dyadic Walsh analysis) was later developed to apply to any space of complex-valued functions on a compact (or even a locally compact) abelian group [73]. Thus the theory of harmonic differentiation in which, by definition,

Def. 2.5.a a harmonic differential operator D on $\mathcal{C}[G]$ is any linear operator on $\mathcal{C}[G]$ whose set of eigenfunctions is the set of characters of G. Since there are many such operators in each space, uniqueness is assured by relating the eigenvalue corresponding to each character to the character's index ν,

$$D\{\chi(\nu,x)\} = \sigma(\nu)\,\chi(\nu,x) \tag{2.5.1}$$

The mapping $\sigma(\nu)$ depends on the structure (as groups) of the domains G and Γ [71,72,73,38]. Since, for the moment, we are not interested in the structure of D but in its formal relationship with Walsh functions

on "unidimensional" domains like $\mathbb{F}_{q^n}$, $[0,1)$ or R_+, $\sigma(\nu)$ will be considered as the natural injection of ν in the complex number field, so that we may write

$$\mathcal{D}\chi(\nu,x) = \nu\,\chi(\nu,x). \tag{2.5.2}$$

$\mathcal{D}$ has many of the properties of the classical differentiator, including the following

- a function has zero harmonic derivative iff it is constant;

- the relationship harmonic derivative - Fourier transform on group is the same as in classical analysis

$$(\mathcal{D}f)\hat{} = \nu\,\hat{f} \quad \text{and} \quad \mathcal{D}f = F^{-1}\{\nu f\}; \tag{2.5.3}$$

- the harmonic derivative of the convolutory product of two functions is equal to the convolution of one function with the derivative of the other one,

$$\mathcal{D}\{f * g\} = (\mathcal{D}f) * g = f * (\mathcal{D}f); \tag{2.5.4}$$

- the extension to higher order harmonic derivatives can be made either by induction (for integer orders) or via the definition of derivatives of the harmonic "delta" function (i.e. the unit of the convolution function algebra),

$$\mathcal{D}^t\delta \triangleq F^{-1}\{\nu^t\} \qquad \mathcal{D}^t f = f * \mathcal{D}^t\delta \tag{2.5.5}$$

so that the harmonic differentiator appears as a convolutional operator [91,92,72];

the harmonic derivative may be given a geometric interpretation, and may be computed as a linear combination of some of the function's values [73,93].

Since, as we have seen, the Walsh functions are characters of a certain LCA group, it follows that they are the eigenfunctions of the harmonic differentiator on the same group. This is rather trivial, because it was the operator which has been defined according to this relationship. Nevertheless, the Walsh functions can be properly defined as the eigenfunctions of the harmonic differentiator provided we follow Gibbs' original path and construct the operator from considerations other than Def. 2.5.a, for instance, from considerations of function variation along directions in the multi-dimensional q-adic group. (See [70] for a very suggestive comparison between the role played in physics by differential equations with boundary conditions and the role played by harmonic "polychotomous" differential equations in information sciences.) Many other systems of orthogonal functions can be obtained as solutions of harmonic differential equations. An example is the system of discrete complex Walsh functions [94] which appear from a second order equation. Also, by generalizing the idea of a logical (dychotomous) differential equation, it is possible to construct the orthogonal system of Takayasu Ito functions [95,96], of which the dyadic Walsh system is but a subclass [70].

Alternatively, one can dispense with any heuristic-geometric interpretation of harmonic differentiation and start with an ad-hoc

definition; this is the path followed by Butzer and Wagner who (subsequent to an idea of Pichler [91]) have extended, from the viewpoint of analysis and approximation theory, the concept of dyadic differentiation to functions on either the unit interval or the entire non-negative real line [56,80,97,98]. They start from an infinite series

$$\frac{1}{4}\sum_{k=-m_1}^{m_2}\frac{f(\cdot)+f(\cdot\dotplus 2^{-k})}{2^{-k}}; \qquad \begin{matrix} m_1\, m_2 \to \infty \text{ for } f\in L_1[R_+] \\ m_1=-1,\ m_2\to\infty \text{ for } f\in L_1[0,1] \end{matrix} \tag{2.5.6}$$

and if this series converges pointwise, they define its limit as the pointwise dyadic derivative of f; whereas if the series converges to a function ϕ with respect to the norm of the function space, they call this function ϕ the strong dyadid derivative of f (much in the sense of a Frechet derivative [99]). Then they are able to prove that the operator thus defined obeys all the formal properties of a harmonic differentiator and that the periodic (aperiodic) Walsh functions are the eigensolutions of the first order dyadic differential equation.

$$\mathcal{D}f = \nu f \quad \text{with} \quad \lim_{x\to 0+} f(x) = 1 \tag{2.5.7}$$

Much of their interest has been directed towards problems concerning the existence of dyadic derivatives, the approximation of Walsh series coefficients by means of dyadic derivatives, and the solution of dyadic wave equations.

We shall not go into any of these details because our main interest concerns only the possibility of defining the Walsh functions as the eigenfunctions of a linear operator having some of the formal properties

of the classical differentiator. These will suffice to formalize a harmonic state-space analysis of certain linear systems in Chapters III and V. For the moment, we shall prove two very important theorems concerning the q-adic harmonic differentiator.

Th. 2.5.a The q-adic harmonic differentiator commutes with the q-adic translation operator,

$$(T^{\gamma}\mathcal{D}f)(x) = (\mathcal{D}T^{\gamma}f)(x). \tag{2.5.8}$$

The theorem can be proved simply by using Eqns. 2.5.3 and 2.4.20 to compute the (Walsh) Fourier transforms of both sides of Eqn. 2.5.8, which turn out to be identical.

Th. 2.5.b The q-adic harmonic differentiator is not, in general, a Ritt-Kolchin operator [100], i.e. it does not obey the product rule

$$\mathcal{D}\{fg\} = f\cdot(\mathcal{D}g) + (\mathcal{D}f)\,g \tag{2.5.9}$$

This fact was noticed by Butzer and Wagner [98], who were able to give a counter example, and, independently, by Cohn-Sfetcu [101], who showed that of all harmonic differentiators on complex-valued functions only the classical one obeys the product rule.

Ritt [102] and Kolchin [103] generalized the concept of derivative by taking the product rule of the Newton-Leibnitz operator as fundamental property. In the context of their differential algebra, a derivative operator on a ring R (field K) is, by definition, any operator D satisfying

$$D(a + b) = Da + Db$$

$$D(a \cdot b) = (Da)\, b + a\, (Db); \qquad \forall\, a,b \in R(K). \tag{2.5,10}$$

Of course, as for the harmonic differentiator, not all the properties of the classic differential operator are characteristic to the Ritt-Kolchin operators.

The proof of the theorem makes use of the fact that the Haar integral (and hence the Fourier transform and the inner product as well) is a linear operator with respect to (arithmetic) addition in the field of complex numbers but it is not linear with respect to q-adic addition. Again, the proof is made in the "transform domain". For classical differentiators and Fourier analysis, one has that (written formally and disregarding the $2\pi j$ factor)

$$\begin{aligned} F\{D(fg)\} = \nu(\hat{f} * \hat{g}) &= \int \nu\, \hat{f}(\eta)\, g(\nu-\eta)\, d\eta \\ &= \int (\nu-\eta+\eta)\, f(\eta)\, g(\nu-\eta)\, d\eta \\ &= \int \eta\, f(\eta)\, g(\nu-\eta)\, d\eta + \int f(\eta)\, (\nu-\eta)\, g(\nu-\eta)\, d\eta \\ &= (Df)\hat{} * \tilde{g} + \hat{f} * (Dg)\hat{} \\ &= F\{gDf + fDg\}, \end{aligned}$$

while for q-adic differentiators and Walsh-Fourier transforms one has that

$$\begin{aligned} W\{D\{fg\}\} = \nu(\hat{f} \overset{\circ}{*} \hat{g}) &= \int \nu\, f(\eta)\, g(\nu \overset{\circ}{-} \eta)\, d\eta \\ &= \int (\nu \overset{\circ}{-} \eta + \eta)\, f(\eta)\, g(\nu \overset{\circ}{-} \eta)\, d\eta \\ &\neq (Df)\hat{} \overset{\circ}{*} \hat{g} + \hat{f} \overset{\circ}{*} (Dg)\hat{}\,. \end{aligned}$$

The fact that the q-adic harmonic differentiator is not a Ritt-Kolchin operator helps to explain some of its pecularities [104,105,106].

To put it in different words, Th. 2.5.b is a corollary of the fact that of all LCA group-domains for $\mathbb{C}[G]$, R is the only one which is also the additive group of the field of complex numbers which constitutes the co-domain of the function space. Of course, for function spaces, K[G], in which the co-domain is a field other than $\mathbb{C}$, there is a different (i.e. not the N-L differentiator) operator which is both a harmonic and a Ritt-Kolchin differentiator. This observation [101] prompted Gibbs [107] to define differentiations for dyadic functions, thus recapturing the Boolean differential operators of Thayse and Davio [108].

Many researchers [108,109] have also drawn attention to the fact that, while the classical differentiator is concerned with local behaviour of functions, the dyadic differentiator considers their global behaviour. This is correct only as long as we regard the functions' domain as a uni-dimensional space, disregarding the fact that the Walsh functions are characters of the dyadic group, with its multi-dimensional structure and topology [110,22]. This problem, as well as the relationship between harmonic differentiators and local variation ("slope") differentiators is discussed in more detail in Chapter VI.

CHAPTER III

q-ADIC TRANSLATION INVARIANT LINEAR SYSTEMS

3.1 INTRODUCTION; TRANSLATION INVARIANT LINEAR SYSTEMS

This chapter is concerned with one aspect of the application of Walsh analysis to the development of system theory; namely, the theory of q-adic translation invariant linear (q-TIL) systems. These are the systems for which the WT plays the role which the FT plays with respect to "time"-invariant linear (TIL) systems. The chapter is preoccupied mostly with an investigation of the formal and (very important) conceptual similarity between the theory of TIL systems and that of q-TIL systems. The study of such systems (another "instance of the growing algebraic presence in systems engineering" [112]) is motivated both by purely theoretical reasons and by the need for providing a system theoretical framework to applications of Walsh analysis in communications and control. Also, the existence of natural phenomena which could be best modelled as q-TIL systems cannot be excluded a priori [113,114,115].

The study of relationships between phenomena (called signals) is difficult if not impossible unless these relationships are approximated by mathematical models, systems, possessing special properties. In its most abstract form [116,117,118], a system $\tilde{S}$ can be defined as a triplet $\{U, Y, \Delta\}$ where U and Y are sets and Δ is a binary (input-output) relation in $U \times Y$; the domain of the relationship is the input set

$U_s \subset U$, and the co-domain is the output set $Y_s \subset Y$. In general, U and Y are sets of mappings having domains D_u and D_y respectively, and co-domains C_u and C_y respectively. Defined likewise, the system is much too abstract, providing a model which is hardly tractable at all. The situation can be alleviated by imposing certain structures on the input and output sets, as well as on the relation δ itself.

This can be done first by structuring both the domains and co-domains of U and Y. Dynamic systems require that the domains D_u and D_y be ordered sets (called time sets). In practice, it is customary to consider the same time set for both input and output sets (a counter-example is a sampling device), the normal choice being either R (continuous systems) or Z (discrete systems) with their usual order relation $\leq$. The nature of the order relation plays a fundamental role in connection with the concept of causality [118,119]. The structures of the co-domains C_u and C_y may, in their turn, vary from that of a group [120] to that of a vector space or even an algebra [121,117,119,112]. Most of the classical applications consider the field of complex numbers $\mathbb{C}$ as the co-domain for both the input and output sets of single input - single output systems (a counter-example is a quantizer). For k-multiple input - m-multiple output systems the co-domains C_u and C_y are the finite dimensional vector spaces $\mathbb{C}^k$ and $\mathbb{C}^m$ respectively. It is also customary to restrict the input and output sets to be spaces of p-summable functions, and, moreover, to give them the structure of Banach or, even better, Hilbert spaces.

The second stage in the specification of a system involves the structuring of the relation δ itself. The usual first step in this direction is to make δ a morphism, that is to make it a map $H : U \to Y$ compatible with the structure of input and output sets. In the classical context of linear (usually complex-valued function) spaces U and Y this translates as the requirement that the system be linear, i.e.

$$H\{\alpha_1 u_1 + \alpha_2 u_2\} = \alpha_1 H\{u_1\} + \alpha_2 H\{u_2\} = \alpha_1 y_1 + \alpha_2 y_2 \qquad (3.1.1)$$

where $\alpha_1, \alpha_2 \in$ field K (usually $\mathbb{C}$), $u_1, u_2 \in U$ (usually $\mathbb{C}[R]$), and $y_1, y_2 \in Y$ (usually $\mathbb{C}[R]$).

The next step is to specialize the investigation to the linear systems possessing certain symmetries; for instance, systems described by linear operators which commute with certain other operators. The classical example is the class of linear "time"-invariant systems whose operators commute with the (arithmetic) translation operator T. The name time-invariant reflects the fact that in this case the transformation realized by the system is essentially independent of the "time" at which the excitation is applied,

$$Hu(t-\tau) = HT^{\tau}u(t) = T^{\tau}Hu(t) = T^{\tau}y(t) = y(t-\tau) \qquad (3.1.2)$$

The word "time" has been put in quotation marks to underline the fact that it is used generically; there are systems which are space-invariant but obey the same formal relationships as the TIL systems. Translation

invariance preconditions a group structure for the functions' domain; otherwise, the translation operators would not form a group, and the function space would not be closed under the operation of translation. In classical system theory, the role of D_u and D_y is played, as already mentioned, by either of the groups $(R,+)$ or $(Z,+)$, or their subsets. This chapter is concerned with a similar class of systems - the systems for which $D_u = D_y$ = a q-adic group (generically denoted G_q), and the operators are linear and commute with the q-adic translation operator. These systems are said to possess q-adic symmetry and are called q-adic translation invariant systems.

The study of q-TIL systems evolved almost exclusively in connection with the dyadic Walsh transform and, hence, was confined to the particular case of dyadic invariant systems. Much of the development of their theory can be traced to the work of Pichler [91,92,74,122] although initial steps had been taken earlier by Nailor [123] and LaBarre [124]. Lately, theoretical contributions were made by Yuen [125], Cheng and Liu [126,127], Gethoffer [128,129], Le Dinh et al [130,131], Morettin [132,133], Pearl [134], Hook [135] and Cohn-Sfetcu [135,106] amongst many others. Harmuth [88,136] played an essential role with regard to the realization, dissemination and applications of dyadic systems. In general, the theory of dyadic systems has been developed separately for discrete and continuous ones, although the formalism is identical [106]. q-adic systems were also discussed by Seviora in his

study of generalized filtering [138] viewed as a multiplicative shaping of (generalized Fourier) spectra of functions on LCA groups; but he was less preoccupied with the theory of the system per se as with the possibility of its realization by sequential circuits.

All these contributions may be classified as homogeneous applications of Walsh analysis to linear system theory because, as will be shown shortly, the eigenfunctions of a q-TIL system are just the Walsh functions. The heterogeneous applications of Walsh analysis to system theory refer to studies of non-q-TIL systems and will be discussed in Chapter IV. Of this chapter, Section 3.2 is devoted to the input/output (I/O) description of q-TIL systems, Section 3.3 to their identification, and Section 3.4 is concerned with a "harmonic" state-space description of such systems.

3.2 I/O DESCRIPTION OF q-TIL SYSTEMS

For simplicity and clarity of exposition, the first part of this section will be concerned only with single input - single output linear systems. The extension to multiple input - multiple output systems is formally immediate and involves merely notational considerations. As in the preceding chapter, the "medium" will be the space of summable complex-valued functions on a q-adic abelian group (which, for generality, may take any of the forms discussed in Chapter II). On this linear space an inner product, hence a norm, is defined in the usual way (Eqn. 1.2.16). The Hilbert space thus obtained will be denoted

generically as $\mathcal{L}[G_q]$, and will be considered as both input and output set for the systems to be discussed. (The condition of a Banach input/output space would suffice for the forthcoming discussion; nevertheless, the choice of restricting the space to be a Hilbert one is warranted by the fact that it has a richer structure which nevertheless suffices to cover all practical applications; also, it permits a notational unification of the presentation by replacing either of the symbols of summation or integration with that of the inner product. $\mathcal{L}[G_q]$ is also a Banach algebra if multiplication is defined as q-adic convolution. If G_q is discrete, the algebra has a unit δ, if G_q is not discrete, a unit (also δ) may be added in a manner similar to that in [119].

From now on, a linear system will be considered as the triplet $(\mathcal{L}[G_q]\ ;\ \mathcal{L}[G_q],\ H)$ where H is a linear operator on $\mathcal{L}[G_q]$. The discussion will be limited to bounded linear operators, whose theory is very well established [138,99] and which are a better reflection of natural relationships. The above definition covers only the zero-state linearity in the sense of [139]. To make it more general and not limit linearity to the zero-state condition would have meant to consider input/output spaces containing functions defined only on subsets (not subgroups) of G_q, a fact which would lead to unnecessary complications for defining translation invariance (since the functions' domain is no longer a group). Besides, as we shall see, the concept of state as well as that of causality are not so readily definable in our case, and

the generalization to non-zero state linearity is not warranted by any practical implementations.

From the basic Riesz representation theorem and its extension to distributions and measures [139,99], the action of the operator H on an input function $u(t)$ can be represented formally as

$$Hu(t) = \langle h(t,\cdot) , u(\cdot) \rangle \quad . \tag{3.2.1}$$

Def. 3.2a A q-adic translation invariant system is a system whose operator is linear and commutes with any element of the group of q-adic translation operators,

$$H\, T^{\tau} u(t) = H\, u(t \dot{-} \tau) = T^{\tau} H u(t) = T^{\tau} y(t) = y(t \dot{-} \tau). \tag{3.2.2}$$

The following two theorems are the fundamental results of the I/O theory of q-TIL systems.

Th. 3.2a A q-TIL system is uniquely represented by a q-adic convolution operator, i.e. by a linear operator H whose kernel $h(t,\xi)$ depends only on $(t \dot{-} \xi)$,

$$Hu = \langle h(t \dot{-} \cdot) , u(\cdot) \rangle = h \overset{\bullet}{*} u \tag{3.2.3}$$

The kernel h is called the impulse response of the system because

$$H\, \delta = h \overset{\bullet}{*} \delta = h. \tag{3.2.4}$$

Th. 3.2b A q-TIL system is uniquely represented by a pointwise multiplication function relating the Walsh transforms of the input and output

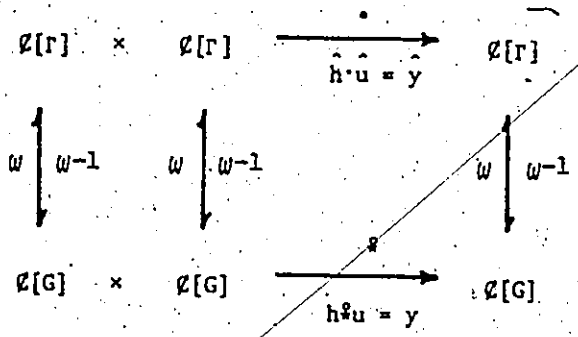


Figure 3.1. Diagram of Walsh transform mapping of q-TIL systems.

signals; this function is the WT of the impulse response and is called the transfer function of the system.

$$y = Hu \leftrightarrow \hat{y} = \hat{h} \cdot \hat{u} \tag{3.2.5}$$

An immediate corollary of these theorems states that the q-adic Walsh functions are the eigenfunctions of a q-TIL system,

$$H\ \mathrm{wal}(\lambda,t) = \hat{h}(\lambda) \cdot \mathrm{wal}(\lambda,t). \tag{3.2.6}$$

This corollary, together with Th. 3.2b represents a variant of the theorem stating that the Walsh (Fourier) transform is an isomorphic mapping between a Banach function convolution algebra and a Banach pointwise multiplication algebra. Fig. 3.1 illustrates diagramatically these relationships. For multiple input - multiple output systems u and y are vector functions, while h and $\hat{h}$ appear as matrix functions, but otherwise the formal relationships are the same as above. The composition (series or parallel) of q-TIL systems obeys the same laws as those for time-invariant linear systems.

The two theorems above appear as particularizations of theorems 3.8.3 and 3.8.4 in [2] which state that given any space $L^1[G]$ where G is a LCA group,

- for any bounded linear operator ψ on $L^1[G]$ commuting with a translation operator there is a function ϕ on the dual group Γ such that $(\psi f)\hat{}\,(\nu) = \phi(\nu) \cdot \hat{f}(\nu)$ for any $f \in L^1[G]$ and, conversely,

if ψ obeys the above equation, then it commutes with any translation operator;

- the bounded linear operators on $L^1[G]$ which commute with all translation operators are precisely the transformations of the form $\psi f = f * \mu$, where μ is a regular measure on G [99]. A more abstract theory of translation invariant operators on L^p spaces is presented in [140].

The validity of these theorems for any LCA groups shows that the formal parallelism between the I/O description of time-invariant and q-adic translation invariant linear systems is total because, in both cases, one deals with bounded linear transformations and Fourier transforms on LCA groups without any reference to the particular structure of the groups involved. Hence, the entire I/O theory of q-TIL systems may be developed along the same lines as for time-invariant linear systems. Unfortunately, this fact was not fully recognized in the engineering community from the beginning, so that there were many contributions each treating only particular aspects of this formal parallelism. As mentioned before, Harmuth [88] and Pichler [91,122] have treated mainly continuous dyadic systems, while Cheng and Liu [127] and Pearl [134] and others paid attention to discrete ones; also, Gethoffer [129] who developed a theory of transformations between various dyadic systems. Le Dinh et al [131] were concerned with the response of discrete dyadic systems to "bandwidth" limited signals, i.e. signals having WT's whose

support is a proper subset of the dual group Γ.

Of more practical interest is the problem of identifying q-TIL systems [132,135] because:

(1) although no real-life q-adic system has been yet discovered, there is need for identifying man-made dyadic systems, and

(2) q-adic system identification may play a useful role as an intermediate stage in a more economical procedure for identifying time-invariant systems [141,142,143]. As an example, the next section considers the identification of dyadic invariant systems with the help of random signals; the generalization to q-adic systems is formal and immediate. The identification of linear systems (time-invariant or not) via estimation of intermediate dyadic models is discussed in Section 4.5.

3.3 DYADIC STATIONARY RANDOM PROCESSES AND THE IDENTIFICATION OF 2-TIL SYSTEMS

A prerequisite to the I/O identification of dyadic invariant systems is the theory of dyadic stationary random processes and their interaction with dyadic systems. Dyadic stationary random processes were discussed by Weiser [144] and, more recently, by Gibbs [79], Pearl [109], Maqusi [145], and Morettin [132,133]. Their work should be dissociated from the work of Gibbs and Pichler [146], Robinson [147], Pearl [134] and Ahmed et al [143] who, amongst others, used Walsh analysis in an effort to obtain a more (computationally) advantageous representation of ergodic (arithmetic) stationary random processes.

Their work, which may be classified as a "heterogeneous" approach to random processes theory, is paralleled (with greater depth) by the work of Sawicki [148] with regard to the Haar transform, and will be analyzed in the next chapter. For the moment, attention will be directed to a "homogeneous" approach in which both signals and linear systems have dyadic symmetry. Since the discussion is limited to the dyadic case, it is sufficient to consider real input and output function spaces. The dyadic addition operator (identical to that of subtraction) is symbolized by $\oplus$, while the operation of dyadic convolution is symbolized by θ.

Def. 3.3a A random process [149] has dyadic symmetry (i.e. it is strictly dyadic stationary) if it is characterized by a probability structure whose (k,ℓ)-th element is strictly a function of $k \oplus \ell$. This expresses the fact that the probability of transition between events k and ℓ is invariant under a dyadic translation of the events [109].

A random process (denoted liberally as $v(t)$) is said to be weakly dyadic stationary if its first and second order statistics are independent of dyadic shifts. This translates into the requirement that the expected value

$$\eta_v(t) \triangleq E\{v(t)\} \tag{3.3.1}$$

is a constant (assumed henceforth, with no great loss of generality, to be zero), and that the autocorrelation function depends solely on the dyadic difference of its arguments [133],

$$r_{vv}(t_1,t_2) \triangleq E\{v(t_1)\, v(t_2)\} = r_{vv}(t_1 \oplus t_2). \tag{3.3.2}$$

This definition can be extended to multi-dimensional stochastic processes. For two random processes defined over the same probability space, the condition of joint weakly dyadic stationarity requires that

$$r_{vz}(t_1,\ t \oplus \tau) \triangleq E\{v(t_1)\ z(t_1 \oplus \tau)\} = r_{vz}(\tau). \tag{3.3.3}$$

As expected, the properties of dyadic stationary random processes are similar to those of (arithmetic) stationary random processes. For instance, the relations

$$|r_{vv}(\tau)| \leq r_{vv}(0) \tag{3.3.4a}$$

$$r_{vz}(\tau) = r_{zv}(\tau) \tag{3.3.4b}$$

$$r_{vz}^{\ 2}(\tau) \leq r_{vv}(0)\ r_{zz}(0) \tag{3.3.4c}$$

are also valid. The notions of correlation, orthogonality and independence apply as well to dyadic stationary random processes. Likewise, one can define spectral density functions.

The Walsh transform of the autocorrelation function $r_{vv}(\tau)$ of a dyadic stationary random process is called the Walsh power spectral density function

$$\rho_{vv}(\lambda) \triangleq \mathcal{W}\, r_{vv}(\tau) = \hat{r}_{vv} \tag{3.3.5}$$

It is evident that

$$E\{v^2(t)\} = r_{vv}(0) = \langle \rho_{vv}\ ,\ 1 \rangle \tag{3.3.6}$$

Similarly, the Walsh cross-spectral density function is defined as

$$\rho_{vz}(\lambda) \triangleq W\, r_{vz}(\tau), \tag{3.3.7}$$

with the corollary that

$$\rho_{vv}(\lambda) \geq 0, \tag{3.3.8a}$$

$$\rho_{vz}^{2}(\lambda) \leq \rho_{vv}(\lambda)\, \rho_{zz}(\lambda). \tag{3.3.8b}$$

These results can be summarized by stating that the Walsh transform constitutes a Karhunen-Loeve expansion for dyadic stationary random processes.

On the basis of the fundamental theorem regarding the commutativity of the expected value operator and any linear operator, it can be shown that the degree of dyadic stationary characterizing the input process to a linear system is reflected in the output process. Moreover, the following theorem is valid for dyadic convolution systems and weakly dyadic stationary processes.

Th. 3.3a

$$r_{uy}(t) = h(t) \oplus r_{uu}(t) \leftrightarrow \rho_{uy}(\lambda) = \hat{h}(\lambda)\, \rho_{uu}(\lambda) \tag{3.3.9a}$$

$$r_{yy}(t) = h(t) \oplus r_{uy}(t) \leftrightarrow \rho_{yy}(\lambda) = \hat{h}^{2}(\lambda)\, \rho_{uu}(\lambda) \tag{3.3.9b}$$

Only the first equation will be proved; the others follow immediately.

$$
\begin{aligned}
r_{uy}(\tau) &= E\{u(t \oplus \tau)\ y(t)\} \\
&= E\{u(t \oplus \tau)\ \langle T^{t} h(\sigma)\ ,\ x(\sigma)\rangle\} \\
&= \langle T^{\tau} h(\sigma)\ ,\ E\{u(t \oplus \tau)\ u(\tau)\}\rangle \\
&= \langle h(\sigma \oplus t)\ ,\ r_{uu}(t \oplus \tau \oplus \sigma)\rangle \\
&= \langle T^{\tau} h(\sigma \oplus t \oplus \tau)\ ,\ r_{uu}(\sigma \oplus t \oplus \tau)\rangle \\
&= h(\tau) \oplus r_{uu}(\tau)
\end{aligned}
\tag{3.3.10}
$$

where use was made of the definition of dyadic invariance for both systems and random processes, and of the fact that $\tau \oplus \tau = 0$.

A white noise process has a correlation function which is non-zero only for equality of the arguments and is related to the unit function in the convolution Banach algebra,

$$r_{ww}(t_1, t_2) = P_o\ \delta(t_1, t_2), \tag{3.3.11}$$

so that the process is also dyadic stationary and has a uniform Walsh power spectral density function

$$\rho_{ww}(\lambda) = P_o. \tag{3.3.12}$$

Accordingly, any dyadic stationary process can be viewed as the response of a dyadic invariant system to a white noise excitation.

As for the identification of single input - single output linear dyadic systems, the problem can be formulated as follows: Given two signals $u(t)$ and $y(t)$, find the linear dyadic invariant system which can best describe the relationship $u(t) \rightarrow y(t)$. If $u(t)$ and $y(t)$

are deterministic signals with definable Walsh transforms, then Eqn. 3.2.5 readily provides an answer since

$$\hat{h}(\lambda) = \hat{y}(\lambda)/\hat{u}(\lambda). \tag{3.3.13}$$

But, if $u(t)$ and $y(t)$ are realizations of random processes, then it is necessary to define a criterion with regard to which one can look for a best approximation of a linear system relating any realizations of the random processes $u(t)$ and $y(t)$. In practical situations when an unknown physical system has to be identified, the use of random signals has the advantage of not impairing the normal operation of the system (plant).

The criterion most widely used is the minimum mean squared error criterion. Then, the problem can be formulated as follows: Given two random signals $u(t)$ and $y(t)$ related by an ideal mapping H_o, find the linear mapping system H such that the mean squared error e between $y(t)$ and the output $z(t)$,

$$z(t) = Hu(t) \tag{3.3.14}$$

of the system H is minimum

$$e = E\{|y(t) - z(t)|^2\}. \tag{3.3.15}$$

The solution is similar to that provided by the Wiener-Kolmogorov theory [11]. Indeed, if the two processes u and y are weakly dyadic stationary, then the optimum linear system is a dyadic invariant system with a transfer function

$$\hat{h}(\lambda) = \rho_{uy}(\lambda) \,/\, \rho_{uu}(\lambda) \tag{3.3.16}$$

the expression for the minimum mean squared error being

$$e_{min} = \langle \rho_{yy}(\lambda) \,,\, 1 - \gamma^2_{uy}(\lambda) \rangle \,, \tag{3.3.17}$$

where $\gamma^2_{uy}(\lambda)$ is the coherence function

$$\gamma^2_{uy}(\lambda) = \frac{\rho_{uy}{}^2(\lambda)}{\rho_{uu}(\lambda) \,,\, \rho_{yy}(\lambda)} \tag{3.3.18}$$

Proof:

$$e = E\{|y(t)-z(t)|^2\} = E\{y^2(t)\}-2E\{y(t)\, z(t)\} + E\{z^2(t)\}, \tag{3.3.19}$$

$$z(t) = \langle h(t,\cdot) \,,\, u(\cdot) \rangle \,, \tag{3.3.20}$$

$$e = \langle h(t,s) \,,\, \langle h(t,\sigma) \,,\, r_{uu}(s,\sigma) \rangle \rangle - 2 \langle h(t,s) \,,\, r_{uy}(s,t) \rangle + r_{yy}(t,t). \tag{3.3.21}$$

The stationary points with respect to $h(t,s)$ of the functional e are given by the equation

$$\langle h(t,r) \,,\, r_{uu}(s,r) \rangle = r_{uy}(s,t). \tag{3.3.22}$$

If the processes are weakly dyadic stationary, then Eqn. 3.3.22 becomes Eqn. 3.3.9a, and the solution is immediately obtained as Eqn. 3.3.16 with the mean squared error given by Eqn. 3.3.17.

If $\gamma^2_{uy}(\lambda) = 1$ for all λ, then there is full coherence between the input and output processes and the mean squared error is null. Otherwise, the error is greater than zero. This may signify either

that (a) the mapping is not linear dyadic invariant, or (b) extraneous noise is present in the measurement, or (c) $y(t)$ is an output due to other inputs besides the signal $u(t)$. Hence, the coherence function γ^2_{uy} may be interpreted theoretically as the fractional portion of the mean squared value at the output which is contributed by the input $u(t)$ at the wave index λ; it plays an essential role as a measure of the goodness of estimation of linear system transfer functions [150,135].

The problem of identifying multiple input - multiple output dyadic invariant linear systems can be formulated identically as for the case of single input - single output systems, with the exception that now we are dealing with, for generality, a k-dimensional input signal and an m-dimensional output signal. Due to the linear restriction of the problem, it is evident that the identification of such a system can be reduced to m k-input - 1 output system identification problems. The identification algorithm [135] closely resembles that for time invariant systems [150] and is presented in Appendix A together with a computer implementation of a numerical example.

There are methods for identifying linear systems which do not make extensive use of fast Fourier transforms [151,152,153]. These are not discussed here because they are not of practical interest in the context of q-TIL systems, which are exclusively realized through the intermediary of fast transform algorithms (their only proven raison d'etre) and multiplicative shaping of signals' spectra.

3.4 HARMONIC STATE-SPACE DESCRIPTION OF q-TIL SYSTEMS

The preceding sections have explored the I/O description of q-TIL systems and showed that it resembles very closely the theory of linear time-invariant systems. In fact, as long as the theory is concerned only with the fact that the functions' domains are LCA groups, the parallelism between the two classes of systems is total. But if the structures of the respective LCA groups enter into consideration, this parallelism breaks down and the resemblance ends. The most pertinent example of such an occurrence is the problem of causality and state-space description of q-TIL systems. The continuous or discrete "time" sets usually considered in the theory of linear systems are R and Z, and both of them are torsion free groups [137] such that it is possible to induce on them an order relation $<$ which is translation invariant

$$x < y \rightarrow x + u < y + u \quad \forall\, x,y,u \in R \text{ or } Z, \tag{3.4.1}$$

so that R and Z are ordered LCA groups and the concepts of causality and state may be consequently defined [119,129,154]. Unlike these two groups, a q-adic group is not an ordered group because it is not torsion free, because all of its elements are of finite (q) order [3]. It is true that an order may be induced on the set G_q (by a Fine mapping, for instance) but this order relation is not translation invariant. Hence, there is no meaningful possibility of introducing the (intertwined) concepts of causality and state of a system [116].

Two attitudes with respect to the concept of state and its causal meaning are manifested in the literature on dyadic systems:

- one is of ignoring "pur et simple" the problem [91,105,127];

- the other one is to circumvent the problem by amalgamating dyadic and time invariance characteristics [92]. Pichler's definition of a linear dyadic invariant system [Def. 4.1 in [92]) as a quadruple (U,Y,X,η) where U, X, Y are Banach q-adic convolution algebras and η is a map $\eta: U \times X \to Y$ defined by $y = \eta(u,x) = c^{*}x + d^{*}u$ is open to question. This is so because there is no mention of state evolution (without which the concept of state has no sense), and the state is set up to be independent of the input and to play a role indistinguishable from that of a second input. Also debatable is the definition of "dyadic invariant time systems" (Def. 6.1 and Eqn. 26 in [92]); at best, they can be viewed as a particular example of a 2^n discrete time invariant system with 2^n inputs and 2^n outputs. It is rather difficult to understand the meaning of their spectral representation (Eqn. 27 in [92]) via a Walsh transform, because it merely achieves a "scrambling" of inputs, outputs and "states" as if dyadic translations could be performed on the "group" of inputs (or states, or outputs).

It is nevertheless true that harmonic differential calculus supplies a tool for developing a formalism similar to the state-space theory of linear time invariant systems. The adjective "harmonic" will be used to qualify this "state"-space description of q-TIL systems in which the Newton-Leibnitz differentiator, or the unit increment operator

is replaced by a harmonic differentiator. It is Pichler's merit for being the first to recognize the role which the dyadic differentiator may play with regard to a subclass of dyadic invariant systems. Later, Hook [105] had the brilliant idea of defining a matrix function wal(A,t) in a manner similar to the classic definition of the matrix function exp(At) so that he could initiate a theory of discrete, finite (in the sense of finite G_q) dyadic systems. His work has been continued and generalized by Cohn-Sfetcu [106] who showed that the formal resemblance between the state-space theories of TIL and q-TIL systems is not all-pervasive and does not translate into a conceptual resemblance, as is the case with the I/O description. The main results of the "harmonic" state-space theory of q-TIL systems are summarized in the following paragraphs, starting first with a formal generalization of Pichler's theory of dyadic differential systems [91]. In all following discussions, the group G_q and its dual Γ_q will be viewed as the images of Fine mappings into the real number field.

Def. 3.4a A q-adic differential system is the linear system whose input and output signals are related through a linear q-adic differential equation

$$A(\mathcal{D})y = B(\mathcal{D})u, \tag{3.4.2}$$

where $\mathcal{D}$ is the q-adic differentiator and $A(\cdot)$ and $B(\cdot)$ are polynomials

$$A(\cdot) = \sum_{i=0}^{k} a_i \bullet (\cdot)^i \qquad B(\cdot) = \sum_{i=0}^{m} b_i \bullet (\cdot)^i \tag{3.4.3}$$

As in classical calculus, the general solution y can be found as the sum of the solution y_{zi} (zero input response) of the homogeneous equation $A(\mathcal{D})y = 0$ with boundary conditions $\{\mathcal{D}y(t_o)\}_{i=0}^{k}$, and any particular solution y_{zs} (zero state response) of the Equation 3.4.2 with zero boundary conditions,

$$y = y_{zi} + y_{zs}. \tag{3.4.4}$$

In order to find y_{zi} one looks for the roots of the characteristic polynomial $A(\lambda)$. If the roots $\{\lambda_i\}_{i=0}^{k}$ are distinct and belong to Γ_q (necessarily viewed as its real number representation), then the homogeneous solution is

$$y_{zi} = \sum_{i=0}^{k} c_i \mathrm{wal}(\lambda_i, t), \tag{3.4.5}$$

where the constants c_i depend on the boundary conditions (let us assume $t_o=0$, so that they are the "initial" conditions) and can be found by solving the algebraic system of equations

$$\sum_{i=0}^{n} c_i \lambda^n = (\mathcal{D}^n y)(0) \; ; \quad n = 0, 1, \ldots, k. \tag{3.4.6}$$

The particular solution can be found very simply by applying the WT to Eqn. 3.4.2, from which it follows that

$$Wy_{zs} = \frac{B(\lambda)}{A(\lambda)} Wu \;\rightarrow\; \hat{y}_{zs} = \hat{h}\cdot\hat{u} \;\rightarrow\; y = h * u \;\rightarrow\; h = W^{-1}\frac{B(\lambda)}{A(\lambda)} \tag{3.4.7}$$

Pichler defines the vector $X(0) = (c_1 \dots c_k)$ as the initial state vector, and the vector function $\Psi(t) = (wal(\lambda_1,t) \dots wal(\lambda_k,t))$ as a state transition vector, so that the total response can be expressed as

$$y(t) = \Psi(t)\, X(0) + h(t) \overset{\bullet}{} u(t). \tag{3.4.8}$$

This formal similarity between dyanmic systems and harmonic differential systems can be elevated to a more general form by defining a system as the usual quintuple of input U, output Y, and state X spaces, coupled with maps $U \times X \to X$ and $U \times X \to Y$ expressed by harmonic differential operators:

Def. 3.4b

$$\begin{aligned} (\mathcal{D}X)(t) &= E(X(t),U(t),t) \\ Y(t) &= G(X(t),U(t),t), \end{aligned} \tag{3.4.9}$$

where U, X, Y represent the input vector, the "state" vector and the output vector respectively, and $E(\cdot)$ and $G(\cdot)$ are vectorial functions. t belongs to G_q in its real number representation. For linear systems, these "dynamical" equations take the form

$$\begin{aligned} \mathcal{D}X(t) &= A(t)\cdot X(t) + B(t)\cdot U(t) \\ Y(t) &= C(t)\cdot X(t) + D(t)\cdot U(t), \end{aligned} \tag{3.4.10}$$

where $A(t)$, $B(t)$, $C(t)$ and $D(t)$ are matrix functions of the variable t.

If they are constant matrices, the system is said to be q-adic translation invariant, being characterized by the equations

$$\begin{aligned} \mathcal{D}X(t) &= A\cdot X(t) + B\cdot U(t) \\ Y(t) &= C\cdot X(t) + D\cdot U(t) \end{aligned} \qquad (3.4.11)$$

The solutions of these q-adic differential equations are based on the proof that $wal(A,t)$ and $wal(A,t\ominus t_0)$ are the fundamental and state-transition matrices respectively. The proof, due to Hook [105] for the particular case $G_q = P_{q^n}$ and later generalized by Cohn-Sfetcu [106], starts from the definition of the matrix function $wal(A,t)$.

Given the k by k square matrix A with distinct eigenvalues $\{\lambda_i\}_{i=1}^{k}$ all in the dual group, Γ_q, there exists a modal matrix Q which can diagonalize A through a similarity transformation,

$$L_{\lambda_i} = Q\cdot A\cdot Q^{-1} \qquad A = Q^{-1}\cdot L_{\lambda_i}\cdot Q \qquad (3.4.12)$$

where L_{λ_i} is diagonal with λ_i as non-zero elements. By similarity with the classical definition of

$$\exp(At) = Q^{-1}\cdot\exp(L_{\lambda_i}t)\cdot Q = Q^{-1}\cdot L_{\exp(\lambda_i t)}\cdot Q \qquad (3.4.13)$$

the matrix function $wal(A,t)$ is defined as

<u>Def. 3.4c</u> $wal(A,t) = Q^{-1}\cdot wal(L_{\lambda_i},t)\cdot Q = Q^{-1}\cdot L_{wal(\lambda_i,t)}\cdot Q.$ (3.4.13b)

It is simple to prove that the set of matrices $\{wal(A,t)\}_{t\in G_q}$ with the operator of matrix multiplication is isomorphic to the group G_q; indeed,

$$wal(A,0) = wal(\underline{0},t) = I;\ \underline{0} = \text{null matrix},\ I = \text{identity matrix} \tag{3.4.14a}$$

$$wal(A,t_1)\ .\ wal(A,t_2) = wal(A,\ t_1 \dot{+} t_2). \tag{3.4.14b}$$

Similarly, it is relatively easy to prove that

Th. 3.4a the matrix function $wal(A,t)$ is the eigensolution of the first order q-adic differential matrix equation

$$\mathcal{D}f(A,t) = A\cdot f(A,t) \tag{3.4.15}$$

the derivative of a matrix being defined as usually, $\mathcal{D}C = \mathcal{D}[c_{ij}] = [\mathcal{D}c_{ij}]$. The proof runs as follows:

$$\begin{aligned} wal(A,t) &= \mathcal{D}(Q^{-1}\ L_{wal(\lambda_i,t)}Q) = Q^{-1}\cdot L_{\mathcal{D}wal(\lambda_i,t)}\cdot Q \\ &= Q^{-1}\cdot L_{\lambda_i wal(\lambda_i,t)}\cdot Q \\ &= Q^{-1}\cdot L_{\lambda_i}Q\cdot Q^{-1}\cdot L_{wal(\lambda_i,t)}\cdot Q \\ &= A\cdot wal(A,t). \end{aligned} \tag{3.4.16}$$

Alternatively, one may use Eqn. 3.4.15 as definition of $wal(A,t)$ and then prove Eqns. 3.4.13 and 3.4.14.

If A has all the eigenvalues in Γ_q, but they are not necessarily distinct, a similarity transform can bring A at most into a

Jordan canonical form

$$A = Q^{-1} \cdot L_{J_{\lambda_i}} \cdot Q = Q^{-1} \cdot \begin{bmatrix} J_{\lambda_1} & & \\ & & \\ & & J_{\lambda_i} \end{bmatrix} Q, \tag{3.4.17}$$

where J_{λ_i} is a $n_i \times n_i$ Jordan block

$$J_{\lambda_i} = \begin{bmatrix} \lambda_i & 1 & \dots & \\ . & \lambda_i & 1 & \dots. \\ & & \dots\dots\dots & \\ & & \dots\dots\dots & \lambda_i \end{bmatrix} \tag{3.4.18}$$

such that $\sum_{i=1}^{i} n_i = k$. Then, the Walsh matrix function may be defined as

$$wal(A,t) = Q^{-1} \cdot L_{wal(J_{\lambda_i},t)} \cdot Q, \tag{3.4.19}$$

where

$$wal(J_{\lambda_i},t) = \begin{bmatrix} wal(\lambda_i,t) & \dots \\ . & \dots \\ . & \dots \end{bmatrix}, \tag{3.4.20}$$

so that Eqns. 3.4.14 and 3.4.15 continue to be valid.

A different problem arises when A has some eigenvalues not in Γ_q (obviously, G_q and Γ_q are not algebraically complete), but the difficulty may be circumvented by realizing that the q-adic differential equation $\mathcal{D}f(t) = \lambda_i f(t)$ has no non-trivial solution if λ_i does not belong to Γ_q, in which case it can be written symbolically that $wal(\lambda_i,t) = 0$. Consequently, the matrix $L_{wal(\lambda_i,t)}$ has zero values in place of $wal(\lambda_i,t)$.

In general, the matrix A may appear as a combination of the three cases mentioned above (distinct eigenvalues all in Γ_q, multiple eigenvalues in Γ_q, or distinct eigenvalues not in Γ_q) and the matrix function wal(A,t) is defined accordingly. Essential is the fact that the properties expressed by Eqns. 3.4.14 and 3.4.15 are always valid.

We are now in a position to solve Eqn. 3.4.11. First, it can be verified that the homogeneous equation

$$\mathcal{D}X(t) = A \cdot X(t); \qquad X(t_o) = X_o \tag{3.4.21}$$

admits as (zero input) solution the vector

$$X_h(t) = \text{wal}(A, t \circ t_0) \cdot X_o. \tag{3.4.22}$$

As before, the particular solution can be found by applying the Walsh transform to both sides of Eqn. 3.4.11,

$$\lambda I \cdot \hat{X}(\lambda) = A \cdot \hat{X}(\lambda) + B \cdot \hat{U}(\lambda) \tag{3.4.23}$$

from which it results that

$$X(\lambda) = (I - A)^{-1} \cdot B \cdot U(\lambda). \tag{3.4.24}$$

A condition for the validity of this solution is that U(t) satisfies certain orthogonality conditions [104,105]. Referred to a scalar equation $\mathcal{D}x = a \cdot x + u$, these conditions require that either $a \notin G_q$ (in the case of Fine mapping), or if $a \in G_q$, $\hat{u}(a) = 0$.

Let us denote the inverse Walsh transform of $(\lambda I - A)^{-1}$ by

$\Phi(A,t)$. Then, the (formal) solution of the "dynamic" equations 3.4.11 is

$$X(t) = wal(A,t\stackrel{\circ}{-}t_0)\cdot X_o + \Phi(a,t)\cdot B \stackrel{\circ}{*} U(t) \qquad (3.4.25a)$$

$$Y(t) = C\cdot wal(A,t\stackrel{\circ}{-}t_0)\cdot X_o + [C\cdot\Phi(A,t)\cdot B + D\delta(t)] \stackrel{\circ}{*} U(t). \qquad (3.4.25b)$$

The input-output relationship ("zero-state" response) can be described by a matrix impulse response function H defined as

$$H(t) = C\cdot\Phi(A,t)\cdot B + D\cdot\delta(t) \qquad (3.4.26)$$

or, alternatively, by a matrix transfer function

$$\hat{H}(\lambda) = C\cdot(\lambda I - A)^{-1}\cdot B + D. \qquad (3.4.27)$$

The theory of q-TIL systems may be continued by developing the same pattern of relationships as for linear time-invariant systems; for instance, equations describing canonical decompositions of "dynamical" equations, the tandem, parallel or feedback composition of systems, or irreducible realizations of rational transfer function matrices are formally the same for both classes of linear systems [106].

The results obtained so far represent a q-adic image of those obtained for TIL systems [12], but there are some aspects which make the similarity to be less than perfect (and motivate a Cronos-like behaviour). One such aspect refers to the fact that the excitation vector $U(t)$ must satisfy certain orthogonality conditions with respect to the modes of the system. Another concerns the fact that for q-TIL systems the

number of columns of $wal(A,t)$ which are linearly independent solutions of $\mathcal{D}X(t) = A \cdot X(t)$ is not equal to the dimension of A but to the number of distinct eigenvalues of A in Γ_q. More significantly yet, there is no simple relationship between $\Phi(A,t)$ and $wal(A,t)$ as was the case for TIL systems, where one does not use the Fourier but the Laplace transform with regard to which $\exp(At)$ and $(sI-A)^{-1}$ form a transform pair. Even if a Laplace transform operator L were defined along the lines advocated by Mackey [155,156] (i.e. by defining generalized characters formed as the product of a genuine character taking values on the unit circle with another "character" taking values on the positive real axis) we would still not be able to write

$$L\{\mathcal{D}\{f\}(t)\} = \lambda L\{f(t)\} - f(0) \tag{3.4.28}$$

because $\mathcal{D}$ is not (a) a Ritt-Kolchin differentiator, and (b) a local but a global operator, in the sense that the q-adic derivative of f cannot be defined unless f is known for values of its argument spread all over the domain.

This is another facet of the impossibility of defining causality due to the fact that G_q is not an ordered group. (Kalman [121] has also remarked, in a different context, upon the antithetic character of causality and the concept of differential equations.)

It is this impossibility of defining causality and the concept of state which constitute the greatest impediment to the development of

a harmonic state space description of q-TIL systems. The q-adic differential calculus is indeed a fascinating (and almost virgin) subject for mathematical research. It provides a tool for developing a theory similar to the state-space description of TIL systems, but the similarity is only formal, lacking conceptual mening such that it is not conducive to engineering applications. This raises a barrier which, although only "psychological", has proven to be a rather formidable one.

3.5 CONCLUDING REMARKS

This chapter has presented a unitary theory of q-adic translation invariant linear systems with respect to both external (input/output) and internal (harmonic state-space) representations. The I/O theory is formally and conceptually similar to the I/O theory of (zero-state) linear time invariant systems, both of them being subsumed by the theory of translation invariant bounded linear operators on the space of complex-valued function having a LCA group as domain. It is also possible to develop a harmonic state-space formalism which is similar, up to a certain point, with the state-space theory of TIL systems; but this formal similarity between the two theoretical bodies does not reflect a conceptual similarity. Indeed, this similarity between the theory of TIL and q-TIL systems is meaningful only as long as the structure of the underlying group domain does not enter into consideration. This is an essential conclusion, which was not realized until now because the subject has been treated mostly in an engineering context, where applications dealt

exclusively with off-line implementations of q-TIL systems via fast Walsh transforms, and where previous theoretical contributions concentrated upon treating the functions' domain as a subset of the reals or the integers without realizing that, fundamentally, it is a group which can not be ordered.

A somewhat similar analysis can be made for systems with function spaces having finite fields as co-domains; this subject is treated in Section 5.3, following the introduction of harmonic differentiation in Galois fields. The next chapter is concerned with engineering applications of Walsh analysis and q-TIL systems theory in complex-valued function spaces.

CHAPTER IV

ON THE ENGINEERING USE OF WALSH FUNCTIONS AND THE WALSH TRANSFORM

4.1 INTRODUCTION

Simultaneous with the proliferation of digital computers and the development of more sophisticated computer languages there came a tremendous effort at devising, developing and applying numerical methods to solving a multitude of engineering problems. An important part of this effort has been directed towards the use of various discrete finite transforms in Communications, Signal Processing, and Automatic Control [19, 20,157]. The attractiveness and extensive use in engineering practice of orthogonal transforms for representing data can be explained by the fact that, in many cases, it is more advantageous to manipulate the data in the transform domain. "Advantageous manipulation of data" is understood either in the sense of more economical and safer representations (and transmission/reception) of data, or in the sense of permitting simpler relationships between, and classifications of, various sets of data. The first aspect refers to data coding and compression, while the second refers to pattern recognition, signal processing and system theory.

Of the many transforms proposed and applied, the two predominant ones have been the discrete Fourier transform (DFT) and the discrete Walsh transform (DWT); the first one because of its affinity to real-world problems viewed through the prism of linear (arithmetic) translation

invariant systems, the second because of its affinity with numerical computations on digital binary computers. The computational advantage associated with the Walsh transform have induced many applied scientists to forego the better suitability of the discrete Fourier (or other) transforms in solving certain problems, and to replace these transforms by the Walsh or other (Haar, for example) computationally attractive ones. In many such cases, the results have been rather disappointing; a fact which led to several attempts at discussing the "goodness" of the Walsh transform. Unfortunately, many of the explanations offered in the engineering literature remain at the surface of the matter; in the sense that they are limited to a (often empirical) comparison of the usefulness of various transforms for certain particular problems. The author is not aware of any attempt to present the fundamental reason for the frequent unsuitability of Walsh analysis or, more significantly, to indicate the problems for which the Walsh transform leads to a theoretically optimal solution.

As mentioned in the introductory chapter, it was one of the major goals of this dissertation to provide such a comprehensive analysis. In the main, this is the subject of Chapter IV: a constructively critical assessment of the engineering usefulness of Walsh functions and the Walsh transform. Chapters II and III have made a mathematical presentation of the Walsh transform and q-TIL systems, respectively, precisely in order to lay a theoretical foundation for the forthcoming discussion. From the multitude of attempted (reported) and possible applications of

Walsh analysis to engineering problems, only those clearly related to signal processing or system theory have been selected for discussion.

First of all, attention will be directed towards the computational advantages of the discrete Walsh transform because these were the driving force behind the use of Walsh functions in the majority of cases. This is done in Section 4.2. Section 4.3 is devoted to analyzing and presenting (in fact, reviewing) the true character of the Walsh transform as a multi-dimensional discrete Fourier transform. This fundamental fact has already been mentioned in Chapter II, but it is in Section 4.3 that conclusions are drawn with respect to the practical applications of Walsh functions. Some of the problems which are suited to an optimal solution with the aid of Walsh analysis are presented in Section 4.4, together with a succint review of the literature on this subject. The following section is devoted to an analysis of other (by no means all) problems in signal processing, system theory and automatic control for which the Walsh transform may be useful in the sense that, although it does not lead to a theoretically optimum solution, it provides the best (or at least a good) one from a practical point of view. In other words, it provides a solution characterized by a felicitous compromise between mathematical goodness and real cost- and time-efficient implementability. This matter of choosing a compromise between abstract goodness and implementability is another aspect of the primordial engineering problem of coexistence, correlation and balance between mathematical models (systems) and the real world which we can know, measure or manipulate

only in a finite, discrete and "rational number" sense [158]. This latter fact explains why in the following chapters attention will be mainly directed towards function spaces on finite, discrete domains and their corresponding finite discrete transforms.

4.2 COMPUTATIONAL ADVANTAGES OF WALSH TRANSFORMS

First and foremost, the Walsh functions form a complete orthonormal set in the Hilbert space of complex-valued functions on a compact abelian group, which may (but need not) be seen as a finite interval on the real line. As such, a Walsh function representation enjoys all the advantages of signal representations with respect to a complete orthonormal set of functions:

- the (finite) representation is optimum in a least square sense;
- the representation converges, the error decreasing with increase in the number of basis functions considered;
- the inner product is invariant under the representation (hence, the energy is too);
- the coefficients of the representation are independent of each other, being dependent only on the corresponding basis function, such that the improvement of the representation by considering a larger number of terms does not require the recomputation of all coefficients but only of the new ones. The coefficients are computed as the inner products between the data function and the corresponding basis functions.

For data which can be described by explicit analytic functions, it is usually possible to obtain analytic formulae for the coefficients of certain representations as, for example, the Fourier, Lagrange, Chebyshev, etc. representations. But, in practice, the data is gathered experimentally and is contaminated by noise, such that there is no question of closed form formulae for the representation coefficients; they have to be computed numerically. Hence, the advantage of complete orthogonal sets of piecewise constant functions like the Haar or q-adic Walsh functions.

Within the class of q-adic Walsh functions, the dyadic ones enjoy the additional advantage of being real-valued. Accordingly, there is a marked increase in computational efficiency because the analysis need not be concerned with complex-valued functions, such that there is no need to store the imaginary components of the (transformed) data vector, and computations are performed (faster) in real and not complex arithmetic. More than that, dyadic Walsh functions take only values of +1 or -1, such that there is no need to perform any multiplications for computing the coefficients.

There is yet another advantage associated with the use of dyadic or 4-adic Walsh functions; it refers to the fact that these are the only q-adic Walsh functions which can be computed by rational numbers - the only numbers which we can measure or represent exactly [158]. This fact was first noticed by Rader [159] who proved two important theorems,

which were later generalized and completed by Cohn-Sfetcu [160]. Besides their theoretical significance, these theorems have the merit of drawing attention to an often overlooked source of errors in the computation of discrete transforms.

Th. 4.2a Either the real or imaginary part of $w_q = \exp(j2\pi/q)$ is an irrational number, with the exception of the trivial cases $q = 1$, 2 or 4.

Rader has used reductio ad absurdum to prove this theorem (only) for $q = p =$ prime. The proof is based on the fact that conditions $(w_q)^q = 1$ and $w_q = a/c + jb/c$, where a, b and c are integers not all divisible by the same number, are irreconcilable. As for the case, $q = 2^n p$, we offer the following proof:

- if the prime p is greater than 2, then $w_q = (w_p)^{1/2^n}$, and since w_p is not a rational number, neither is w_q (because the integer root of an irrational number is also irrational);
- if $p = 1$ and q is greater than 4, then $q = 8 \cdot 2^m$, and since w_8 is irrational, so is $w_q = (w_8)^{1/2^m}$.

The theorem concerns the discrete Fourier transform as well, because a DFT over N points is nothing else but a "1-dimensional" N-adic Walsh transform.

The second theorem is more general because it refers to z-transforms [161].

Th. 4.2b The z-transform of a sequence of length N, where $N \in \mathbb{P}$, cannot be evaluated exactly in a rational representation with less than N digits for values of z on the unit circle, with the exception of the trivial case when $z = \pm 1$ or $\pm j$.

The proof is based on the derivation of an integer number equation whose solution represents all possible arithmetic bases and all possible initial values of z on the unit circle. This problem can be viewed as one of obtaining the arithmetic radix r such that the complex number z can be represented by at most k digits.

$$z = r^{-k}(\pm a \pm jb) \quad \text{where} \quad \begin{array}{l} r \in \{2,3,\ldots\} = \mathbb{P} - \{0,1\} \\ k \in \{1,2,\ldots\} = \mathbb{P} - \{0\} \\ a,b \in \{0,1,\ldots,r^k-1\} = \mathbb{P}_{r^k} \end{array} \tag{4.2.1}$$

Since z has unit modulus, Eqn. 4.2.1 can be rewritten as

$$|z|^2 = 1 = r^{-2k}(a^2+b^2). \tag{4.2.2}$$

or as

$$a^2 + b^2 = (r^k)^2. \tag{4.2.3}$$

Hence, our problem reduces to the famous one of finding all the Pythagorean triples. For each possible r, the solution corresponding to $k = 1$ is called the fundamental solution. The simplest and most immediate fundamental solutions are the following:

any r, z_0 can be either ± 1 or $\pm j$,

$r = 5$, z_0 can be either $(\pm 3/5 \pm j4/5)$ or $(\pm 4/5 \pm j3/5)$

$r = 10$, z_0 can be either $(\pm 6/10 \pm j8/10)$ or $(\pm 8/10 \pm j6/10)$

$r = 13$, z_0 can be either $(\pm 5/13 \pm j12/13)$ or $(\pm 12/13 \pm j5/13)$.

But, since for representing the m-th power (m being an integer) of z_0 one needs k to be at least equal to m, with the exception of the trivial case, it is impossible to evaluate the z-transform of a sequence of length N using a finite arithmetic representation with at most k digits unless $k \geq N$, or the sequence $Z = 1, z_0^1, \ldots, z_0^{N-1}$ is cyclic with a period less or equal to k. In its turn, the requirement that the sequence Z is cyclic implies that z_0 must be a rational root of unity, which, as proven in Th. 4.2a, is clearly impossible with the exception of the already excluded trivial case. This concludes the proof of Th. 4.2b.

The major computational advantage which characterizes the finite discrete Walsh transform, and which it shares with some of the discrete Fourier transforms, consists in the possibility of performing it via a fast algorithm. This is an inherent property of any q-adic Walsh transform, whereas for a N-point DFT to have it, it is necessary that N be a highly composite number. The idea of fast transform algorithms originated with Good [16] and Cooley and Tukey [15]. Such algorithms can be, and have been, devised and analyzed from two points of view

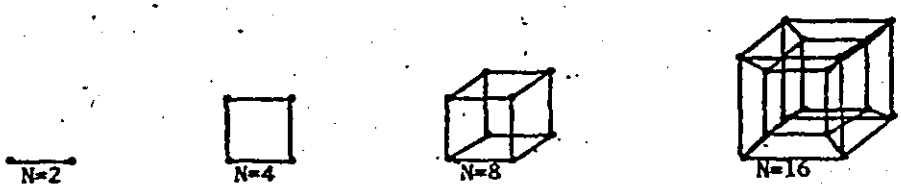
(which, of course, are ultimately identical in the case of discrete Walsh-Fourier transforms). One point of view regards the transformation as a multiplication of the data vector by a highly redundant orthogonal matrix expressible either as a Kronecker product of some elementary orthogonal matrices of much smaller dimensions, or as a (usual matrix) product of very sparse matrices, such that fewer operations are required to implement the transformation [162,83,84,85]. If the matrix is of rank $N = q^n$, the matrix multiplication can be performed in only about nq instead of N^2 operations. The other point of view considers the transformation as a Fourier transform on a finite abelian group which is (or can be regarded) as a direct sum of groups of smaller order. Then, for $N = q^n$, instead of performing a N-point Fourier transform, it is sufficient to perform n q-point Fourier transforms followed by a recombination/re-ordering of the coefficients [63,66]. Fast transform algorithms have been extensively and exhaustively discussed in the literature [164,165] so that there is no need to present them anew.

The first fast Walsh-Hadamard transform algorithms were described by Greene [166], Vandivere and Carrick [167], Whelchel and Guinn [168] and Shanks [169]. Other papers followed [170,171,172,173], most of them being preoccupied with the problem of efficient ordering/re-ordering of the transform coefficients. This is a problem that is characteristic (but not exclusively) to finite transforms; it can be viewed as a consequence either of the invariance of orthogonality to (matrix) row or column shifts, or of the fact that finite abelian groups do not admit a

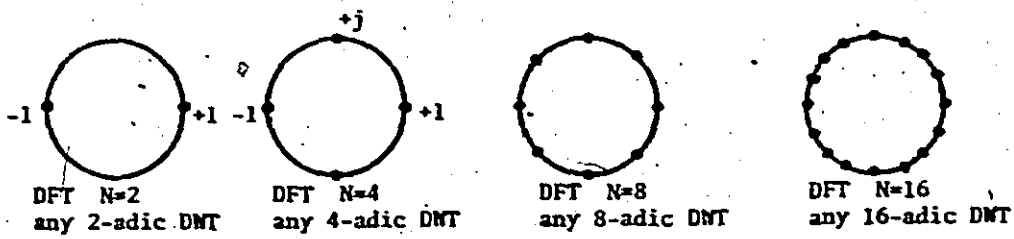
unique order. Whereas it may be computationally more advantageous to obtain the coefficients in a certain order, another order may be preferable from the point of view of data interpretation [83,75]. The significance of transform coefficients and the concept of generalized frequency are analyzed in Chapter VI. As for re-ordering algorithms, Nicholson [66] has presented a comprehensive analysis of all of them. Computational errors related to finite Fourier transform algorithms have been analyzed in [174,19,175].

It would be erroneous to conclude from the preceding discussion that the q-adic, and especially the dyadic, Walsh transforms are the best discrete transforms; they are so but only from a computational point of view (the Haar transform can be implemented even faster, but it is not a generalized Fourier transform). In truth, the resolution (finesse) with which 2-adic and 4-adic Walsh transforms can represent a large variety of functions (signals) is poor and quite inadequate, so that a DFT is preferable even if it cannot be evaluated by rational numbers or if the transform takes longer to be computed. This concept of resolution of representation may be best understood if it is referred to the variation and the total number of possible values which the characters take in the complex plane. In a q-adic Walsh transform, the characters take q values (Andrews and Caspari [162] call them quantum levels) equidistant/equiangularly distributed on the unit circle. From this point of view, the essential difference between what is usually called a Walsh transform and a discrete Fourier transform consists in the fact that in

DFT domains $N = 2^n$



DWT domains $N = 2^n$



Character values on the unit circle.

Figure 4.1. Domains and character values for discrete Walsh and Fourier transforms.

the first case the resolution is independent of the number of points in the domain (length of data vector), whereas in the second case it is directly dependent of this number. This is the explanation of the asymptotic convergence/divergence of DFT/DWT towards the Karhunen-Loeve transform in the case of stationary signals [176]. Figure 4.1 illustrates this situation for N = 2, 4, 8 and 16. If the characters corresponding to discrete finite transforms are viewed as sampled, step-like approximations of the characters on the unit interval, it is evident that an increase in the number of sampling points (i.e. a decrease of the sampling interval) has no effect on the "quantization" of Walsh functions (which are step-like anyway), whereas it leads to a much better approximation of continuous functions like the sinusoids or complex exponentials. In algebraic terms, this phenomenon can be even better explained by the fact that a N-point DFT is a 1-dimensional Fourier transform on an abelian group of order N such that an increase of N increases the order of the group and improves the resolution; whereas a N-point ($N = 2^n$) dyadic Walsh transform is a n-dimensional DFT on an abelian group of order 2, such that an increase of N has no bearing upon the order of the group, or the resolution, but only upon the dimensionality of the space. This, then, is the fundamental explanation of the difference between discrete Fourier and Walsh transforms; it was mentioned in Section 2.3, and it will be analyzed at length in the next section.

4.3 THE WALSH TRANSFORM AS A MULTI-DIMENSIONAL DISCRETE FOURIER TRANSFORM

Ever since it was realized that a 2^n-point Walsh transform admits a faster implementation than a 2^n = N-point DFT, there have been numerous attempts at supplanting the latter with the first, and then studying the efficiency of the "new" transform [168]. More often than not, these studies were done in an empirical way, simply by considering certain particular signals and comparing the efficiency with which either the DWT or the DFT could represent them. Since the signals chosen were usually stationary random processes or very smooth continuous functions, the conclusions were, of course, unfavourable to the Walsh transform. This was the attitude of Blachman [32] in his discussion of "Sinusoids versus Walsh functions". Not only does he develop a comparison of the goodness of WT and FT in representing frequency-bandlimited signals (an unnecessary exercise since the result is predictable at the outset), but he also attaches an inordinate importance to hardware synchronization problems, ignoring the essential use of Walsh transforms in a digital environment, on already sampled functions, or in the software implementation, where no question of synchronization arises. Besides, nowhere does he explain (nor do his opponents in [177]) the fundamental reasons for the different behaviour of Walsh and Fourier transforms.

It seems that only Pearl attempted to introduce a quantitative, mathematical comparison between the two transforms; he studied the Walsh processing of random signals [134] and introduced the concepts of

basis restricted transformations and performance measures for spectral representations, both from the point of view of filtering and of coding [178,179]. These concepts (to be discussed in Section 4.5) were later partly incorporated in the concept of generalized Wiener filtering [180]. Significantly, Pearl showed that, independent of the performance criterion - be it rate distortion or Hilbert-Schmidt decorrelation measure, the distance in performance between discrete Fourier and Walsh transforms does not become negligible as $N \to \infty$, although either one of them may tend asymptotically to the ideal, Karhunen-Loeve, transform depending on the nature (arithmetic or dyadic stationary) of the random process. But he did not explain why it is so, although he may have been aware of the multi-dimensional character of the DWT [109], and tried to look for communication problems which could be described by dyadically symmetric models.

As mentioned in Section 2.3, the multi-dimensional character of the discrete Walsh transform was known to mathematicians ever since the publication of Fine's and Morgenthaler's papers. It was also known to Lechner who used it in combinational logic [181], to Green and Posner who used it in coding theory [182], to Gibbs who defined logic differential calculus [55], to Yuen [60,75] who was mostly preoccupied with generating Walsh functions and comparisons between different orderings to Koenig and Zolesio [183] and, possibly to Pichler [74]. But, this essential feature of the Walsh transform was not known, or was simply ignored in the fields of signal processing, communications, or system

λ \ x in $\{\underline{e}_i\}$		$0\underline{e}_o$, $0\underline{e}_1$	$1\underline{e}_o$, $0\underline{e}_1$	$0\underline{e}_o$, $1\underline{e}_1$	$1\underline{e}_o$, $1\underline{e}_1$
λ in $\{\underline{e}_i\}$	x	0	1	2	3
$0\underline{e}_o$ $0\underline{e}_1$	0	+1	+1	+1	+1
$1\underline{e}_o$ $0\underline{e}_1$	1	+1	-1	+1	-1
$0\underline{e}_o$ $1\underline{e}_1$	2	+1	+1	-1	-1
$1\underline{e}_o$ $1\underline{e}_1$	3	+1	-1	-1	+1

$\underline{e}_o=(0,1)$ $\underline{e}_1=(1,0)$

$\underline{d}_o=(1,1)$ $\underline{d}_1=(1,0)$

λ \ x in $\{\underline{e}_i\}$		$0\underline{e}_o$, $0\underline{e}_1$	$1\underline{e}_o$, $0\underline{e}_1$	$0\underline{e}_o$, $1\underline{e}_1$	$1\underline{e}_o$, $1\underline{e}_1$
λ in $\{\underline{d}_i\}$	x	0	1	2	3
$0\underline{d}_o$ $0\underline{d}_1$	0	+1	+1	+1	+1
$1\underline{d}_o$ $0\underline{d}_1$	1	+1	+1	-1	-1
$0\underline{d}_o$ $1\underline{d}_1$	2	+1	-1	-1	+1
$1\underline{d}_o$ $1\underline{d}_1$	3	+1	-1	+1	-1

Figure 4.2. Walsh function values for different bases of the domains.

theory and automatic control, where the majority of engineering usages of Walsh transforms took place, and where the Walsh transform was considered as any other one-dimensional discrete transform [88,187].

Our attention will be concentrated mainly on dyadic Walsh analysis because it occupies a preponderant position and has been almost exclusively used in applications. The reason for this is the fact that the functions' domain $\mathbb{F}_2^n$ is a self-inverse abelian group with each element being its own inverse (being of order 2) such that the operations of addition and subtraction are indistinguishable (denoted by $\oplus$); hence, its other particularities: (1) the characters are real, rational-valued, (2) the alternately even and odd symmetry of characters replaces the usual complex-conjugate symmetry, and (3) the harmonic differential and slope operators are proportional (see Chapter VI).

The n-dimensional vector space $\mathbb{F}_2^n$ is homogeneous in the sense that its local structure and properties are the same at each point. Once an origin and a set of n linearly independent vectors are chosen, any point (vector from the origin) of the space can be uniquely represented by an n-tuple of co-ordinates along each direction. There are many possible sets $\{\underline{\varepsilon}_i\}_{i=0}^{n-1}$ of basis vectors ($2^n!$ if their permutations are also considered). Of these bases, the ones most often used are

$$\{\underline{e}_i\};\quad e_{ik} = \delta_{ik} \tag{4.3.1a}$$

i.e. the set of edges (of a unit hypercube) concurrent at origin, and

$$\{\underline{d}_i\};\quad d_{ik} = \begin{cases} 0 ; & k < i \\ 1 ; & k \geq i \end{cases} \tag{4.3.1b}$$

i.e. the set of maximal diagonals of order n-i of the unit hypercube with vertex at origin. (By "diagonal of order i" it is meant a diagonal of an i-dimensional sub-hypercube, "maximal" being meant in the sense of the α_1-norm of Eqn. 2.3.21. - see Fig. 2.4). The importance of these bases stems from the fact that they are inductive [184,79], and are used to define the W-P and W-K Walsh functions respectively. A displacement in the space P_2^n can be described by the usual laws of vector addition, i.e. co-ordinate-wise addition modulo 2.

For finite transforms P_2^n represents both the group G and its dual Γ, and the value of characters at points $\underline{x}$ and $\underline{\lambda}$ depend on the inner product of the two vectors (see Eqn. 2.3.14)

$$wal(\underline{\lambda},\underline{x}) = (-1)^{(\underline{\lambda},\ \underline{x})} \qquad (4.3.2)$$

but $\underline{\lambda}$ and $\underline{x}$ need not be represented with respect to the same basis. It is customary to use the basis $\{\underline{e}_i\}$ for $\underline{x}$; if $\{\underline{e}_i\}$ is used for $\underline{\lambda}$ as well, then one gets the W-P functions; but if $\{\underline{d}_i\}$ is used, one gets the W-K functions. This is illustrated in Fig. 4.2.

It has been shown in Section 2.3 that injective norms (for example α_1 and α_2 norms from Eqn. 2.3.21) may be attached such that a one-to-one correspondence is set up between this multi-dimensional space and a subset of the one-dimensional space of real numbers. The existence of such an injective map makes possible the description and analysis of Walsh functions and WFT theory in terms of functions defined on a one-dimensional domain. This being so, it might be argued that

there were no reasons to bother with the spatial, multi-dimensional view. It is, indeed, perfectly possible to dispense with this concept, as do most of the researchers in the field; multi-dimensionality is, however, an essential feature of the group operation (addition, displacement, translation) in the domain of the functions, and there are problems that cannot be adequately treated unless the full spatial figure is used. Among such problems is that of determining those tasks for which the DWT might be optimal (Section 4.4), that of explaining the unpromising results of many attempted applications of Walsh analysis (Section 4.5), and that of finding a consistent and unique criterion for ordering the dyadic Walsh functions so that it is also applicable to q-adic ones (Section 6.5).

4.4 OPTIMAL APPLICATIONS OF WALSH ANALYSIS

This section is intended as a succint review of those engineering problems for which the Walsh transform may bring an optimal solution. Due to the multi-dimensional character of the Walsh transform, these are the problems which can be modelled in spaces of complex-valued functions having as domain a multi-dimensional discrete space. On the surface, this is equivalent to requiring that the signals to be processed or the systems to be analyzed have q-adic symmetry; this was the point of view adopted by Pichler [122] and by Pearl [134] when they discussed optimum filters for dyadic stationary signals, but without conditioning the existence of dyadic signals or systems to a multi-dimensional structure of the function space. In other papers, they seem to be aware of this connection, but they do not state it clearly in unequivocal terms. For

instance, Pichler [74] indicates the (optimal) usage of WT in the field of switching functions, but sets it on a par with the non-optimal usage in the field of image retrieval. Pearl appears to be closest to this interpretation of the nature of Walsh transforms in his study of spectral analyses of probability distribution functions [109], where he does search for an "environment where the dyadic sum is the natural process of combining variables", although he continues to compare the DFT and DWT in $N=2^n$ points as if they were alike.

What was, and still is, the most difficult obstacle to the proliferation of optimal solutions via Walsh transforms is neither the lack of understanding of Walsh analysis, nor our "penchant" on time-invariance. The obstacle is more fundamental because it resides in (1) the non-existence (at least according to present-day knowledge) of natural phenomena which could be best modelled on n-dimensional q-adic spaces and (2) an immutably linear flow of time as we know it [88]. The most noteworthy attempts at finding such phenomena were those of Rosenbloom [114], with respect to magnetism and systems of spins, and of Harmuth [115], with regard to elementary particles. To my knowledge, their pioneering work has not been followed by others, although there seems to be a new interest in the Ising model of magnetism [a private communication from Dr. J.E. Gibbs]. Nota bene: all this work represents a look out for the phenomena to be described by our model and not vice-versa, as it is the custom in applied sciences!

The situation is totally different in the realm of man-made technical world, because people have had since long an affinity for binary logic [185] and bistable devices (the later affinity being forced by technological means). The first example occurs in the field of transmission of binary signals over noisy channels, where the Reed-Mueller codes [186,187] represent nothing but a coding of a binary n-dimensional vector signal by the 2^n-tuple of the corresponding Walsh-Hadamard character [188]. The real computational advantage of Reed-Mueller codes became apparent only after the developing of fast transform algorithms [163,166], which permitted real-time implementations of coding/decoding operations of signals at least up to 6 digits long even in 1968 [182]. It was then that a (32,6)-biorthogonal code (represented by the concatenation of two 2^5-Hadamard matrices of opposite sign) has been successfully used for the Mariner '69 High Rate System (16.2Kbits/second, with bit rate error of less than 5×10^{-3}).

A parallel development took place in relation with Boolean logic and switching functions, where Ninomya [189] continued Mueller's work in studying the analytic properties of certain expansions of functions without being aware of their relationship to Fourier analysis. The main interest of the early research into Fourier expansions of switching functions was directed towards studying functional equivalence. Starting in 1963 [190,191], this work has been developed and set on a firm theoretical basis by Lechner, who produced a comprehensive study on the harmonic analysis of Boolean functions [181]. Hurst [192] and Edwards [193] have

extended the theory to threshold logic, whereas the MBLE-Louvain group have related it to the Boolean differential calculus [108,194] and have extended the theory to multiple-valued logic [195,196,197,198].

Lechner [181] realized the strong relationship between Fourier transforms on Boolean vector spaces and Walsh transforms, but considered them to be disjoint from the point of view of applications: the signal processing ones for WT, and the FT on Boolean vector spaces for switching functions. He uses the latter transforms with two objectives in mind:

- one is related to combinatorial applications, specifically to prime implicant extraction;
- the other is related to prototype equivalence relations (under restricted affine group transformations) and their use for the synthesis of LSI circuits via encoded logic.

His main tools for realizing the first objective are the convolution theorem, the quotient group character theorem, and the Poisson summation theorem. The convolution theorem has been discussed throughout this dissertation. The quotient group character theorem [2] concerns a subgroup H of a LCA group G and its annihilator A; i.e., the set of all $\underline{\nu} \in \Gamma$ such that $\chi(\underline{\nu},\underline{x}) = 1$ for all $\underline{x} \in H$. It can be shown that A and the quotient group Γ/A are the dual groups of G/H and H respectively, and a function f is concentrated on H if and only if its transform $\hat{f}$ is constant on the cosets of A. This is the most important tool for detecting prime implicants. The Poisson summation theorem refers to the relationship between the summation of a function

on a subgroup H and the summation of its transform on the annihilator of H; for $f \in P_2[P_2^n]$ and $H = P_2^k$, this relationship reads

$$\sum_{\underline{x} \in P^k} f(\underline{x}) = 2^{k-n} \sum_{\underline{\nu} \in \Lambda} f(\underline{\nu}). \tag{4.4.1}$$

Lechner could not find a FT method for detecting the minimal covering selection of prime implicants of a logic function, but he was able to provide an algorithm which can detect simultaneously all the implicants of a fully defined function (and of its complement), the algorithm having the advantages of (1) being independent of the size of the vector $\underline{x}$ for which $f = 1$, and (2) detecting the implicants in decreasing order such that the search can be stopped immediately after the function has been covered (this is not necessarily a least cost cover). Of course, the complexity of the algorithm depends on the dimension of the function space and on the size of the subgroups on which the function is concentrated, such that a computationally optimal algorithm should combine this approach with the classical ones of Quine-McCluskey and Morreale-Necula.

The other problem to which Lechner provides a harmonic analysis solution is that of prototype equivalence classification of Boolean functions and their synthesis via encoded logic, a problem which is of great value for LSI circuits. The basic idea behind this approach is elegant; it purports to classify all Boolean functions of n variables so that any function can be implemented via simple transformations of an equivalent prototype function. Golomb [199] considered equivalent

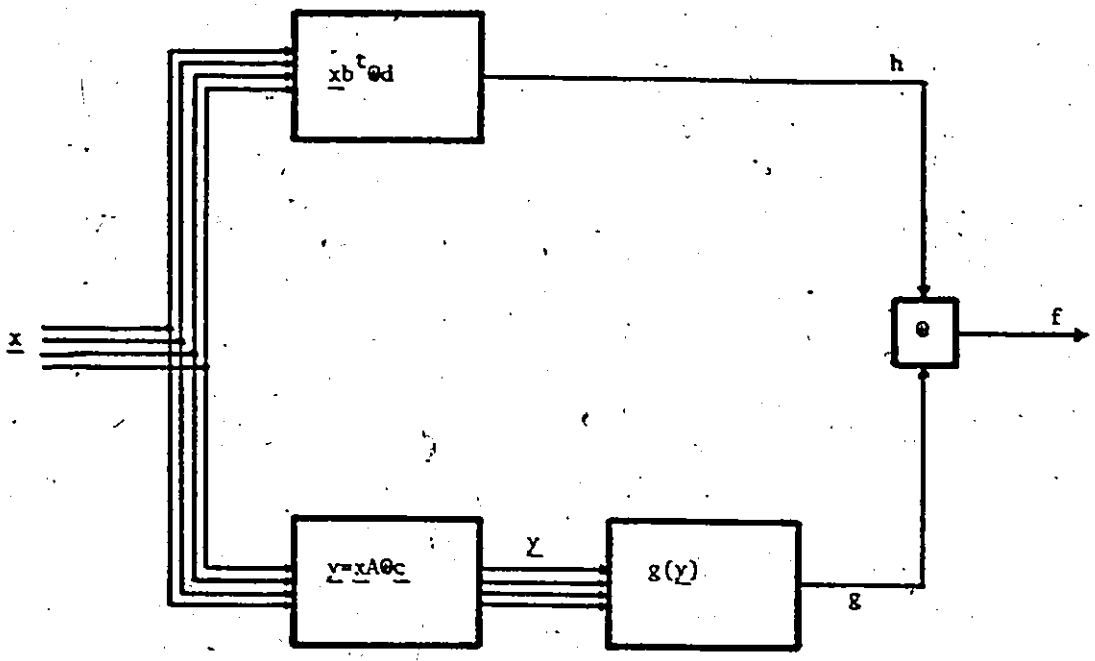


Figure 4.3. Block diagram of encoded-logic realizations of Boolean functions.

all functions which are characterized by a complementation/permutation of their variables. Lechner generalized this equivalence to one under any restricted affine group (RAG) transformations. Two functions f, $g \in \mathbb{F}_2[\mathbb{F}_2^n]$ are said to be related by a RAG transformation if [181]

$$f(\underline{x}) = g(\underline{x}A \oplus \underline{c}) \oplus \underline{x}\underline{b}^t \oplus d \tag{4.4.2}$$

where A is a non-singular $n \times n$ matrix over $\mathbb{F}_2$ (as the Galois field GF(2)), $\underline{b}$ and $\underline{c}$ are arbitrary binary n-tuples, $d \in \mathbb{F}_2$, and $\oplus$ denotes addition in both $\mathbb{F}_2$ and $\mathbb{F}_2^n$. A RAG transformation is composed of a domain encoding transformation $\underline{y} = \underline{x}A \oplus \underline{c}$ and a range encoding transformation $f = g \oplus \underline{x}\underline{b}^t \oplus d$. Fourier analysis may be applied to relate the transforms of these Boolean functions. The relationships are more elegant and symmetric if they are referred to the functions $f_N, g_N : \mathbb{F}_2^n \to \{-1,1\}$ where $f_N(\underline{x}) = 1/2 - f(\underline{x})$, such that $\hat{f}_N(\underline{\nu}) = 2^{n-1}\,\delta(\underline{\nu}) - \hat{f}(\underline{\nu})$. Lechner's fundamental invariance theorem states that, given the relationship 4.4.2 between f and g, their transforms are related as follows

$$f_N(\underline{x}) = (-1)^{\underline{x}\underline{b}^t \oplus d}\, g_N(\underline{x}A \oplus \underline{c}) \tag{4.4.3a}$$

if and only if

$$\hat{g}_N(\underline{\nu}) = (-1)^{\underline{\nu}\underline{c}^t \oplus d}\, \hat{f}_N(\underline{\nu}A^t \oplus \underline{b}) \tag{4.4.3b}$$

The advantage of considering equivalence under RAG transformations consists in the fact that the number of classes it defines grows much more slowly with the number of variables than in the case of other

equivalence relations. For instance, it can be shown [181] that there are 8, 48, and approximately 10^5 RAG classes for $n = 4$, 5, and 6, whereas the corresponding numbers of equivalence classes under the full linear group transformations are 92, 2744, and approximately 10^9 respectively. It would then be economical to devise LSI circuits only for prototype functions of each class, the other functions of the same class being realized by the respective domain and range encodings (see Fig. 4.3)) whose design and implementation is relatively much easier. The problem is to find (1) to which equivalence class does a function belong, (2) to design a cost-efficient implementation of a prototype function (this step implies the search for a simple prototype), and (3) to find the required transformation relating the desired function to its class prototype. Lechner has shown how to use the fundamental invariance theorem, and harmonic analysis in general, for solving this problem in an efficient way. The sheer size of equivalence classes makes practically impossible an exhaustive search for the optimal prototype if $n > 4$, but some heuristic criteria based on Fourier coefficients may lead to good selections.

A similar approach has been adopted by Edwards [193] with respect to threshold logic. Dertouzos [200] had shown that the Chow parameters used for specifying a threshold function are the spectral coefficients corresponding to the Rademacher functions. Edwards' interest is not manifested towards the study of threshold functions per se, but towards the synthesis of binary functions using threshold logic. With this

objective in mind, he studied the significance of spectral coefficients and equivalence relations under certain transformations (which turn out to be Lechner's RAG transformations) in order to come up with the same idea of encoded logic synthesis. He has the merit of (1) recognizing that threshold functions could be efficiently used for implementing a large number of prototypes (for instance 7 out of 8 for $n = 4$), and (2) designing such implementations using optimized versions of the universal threshold gate of Hurst [201].

4.5 EFFICIENT ENGINEERING APPLICATIONS OF WALSH ANALYSIS

The problems which admit a theoretically optimal solution via the use of Walsh analysis are, as we have just seen, precisely those described by models in function spaces having q-adic translational symmetry; i.e. having multi-dimensional domains. With the exception of the field of switching functions, such cases are a rare occurrence in applied sciences and engineering practice, where phenomena are dependent on time, space, etc., in their usual euclidean/archimedean acceptance as variables on the real line, plane, etc.... True then, Walsh analysis cannot provide a theoretically optimal solution, but it can procure an engineeringly optimal one in the sense of compromising between theoretical goodness and efficiency (cost - and time-wise) of implementation. All these cases are more or less related to the problem of approximating functions with the help of Walsh series and polynomials.

Approximation theory based on the Stone-Weierstrass theorem [67] is not very helpful in the case of Walsh functions because they are not continuous with respect to the usual topology of the real line (or plane) [77,97]. But interest in the use of Walsh polynomials for approximations has been strong ever since they were discovered, most of the theoretical investigations being directed towards uncovering the parallelism between trigonometric and Walsh series, with particular accent on the problems of summability, continuity and order of coefficients. Some of the basic results are mentioned in Chapter II, the conclusions being that, evidently, Walsh functions are not particularly suitable for approximating smooth, continuous functions. But what about efficiency of computation?

The first attempt at considering efficiency of computation with respect to Walsh polynomial approximations is due to Polyak and Schreider [202]. Their study was somehow premature because it preceded the introduction of fast transform algorithms, a fact which explains their negative conclusion regarding the efficiency of computing Walsh coefficients. But they show that Walsh polynomials may be useful for data compression, in the sense that less storage may be needed to characterize data (of course, not perfectly, but sufficiently well) in a Walsh representation than either in a simple tabular form or in an algebraic polynomial representation. (An example of their correct conclusion is the successful use of Hadamard transform for the compression of particle count data [203].)

The work of Polyak and Schreider has been continued by C.K. Yuen [204,76,111]. His is a more systematic study because he sets the problem in a more general form: given a function f, find the index set M and the coefficients c_m such that $\sum_{m\in M} c_m \mathrm{wal}(m,t)$ is a good approximation to $f(t)$ over a certain finite interval. The particular scheme discussed by Polyak and Schreider involved the selection of c_m as the Walsh coefficients $\hat{f}(m) = \langle f(\cdot), \mathrm{wal}(m,\cdot)\rangle$, and the restriction of M to containing only those indices for which $|f(m)|$ is greater than a certain threshold (whence the name "threshold sampling" scheme [205]), the criterion of goodness being either convergence in the mean or uniform convergence. Concerning this procedure, Yuen shows that the absolute error bound obtained by Polyak and Schreider is overly pessimistic. He proposes a different scheme in which the truncation is not done according to magnitude of coefficients but according to their rank and degree. This scheme is no longer optimal in the least mean square error sense, but has the (somewhat doubtful) advantage of having an error bound that is easy to compute due to its relationship with the derivative of the signal. Yuen also shows the absence of Gibbs' oscillatory phenomena in relation to finite Walsh approximation. More interesting is his work on mini-max Walsh series. A first stage considers a fixed set M, but improves the approximation by an iterative search for the best coefficients c_m. The second stage attempts an optimization of the index set M in the sense of finding the smallest set for a fixed maximum error. But, again, the conclusion is that Walsh polynomials are advantageous

only for data compression and "small machines low-precision function evaluation by means of special Walsh series hardware" [76].

Application-oriented studies concerning the relationship between Fourier and Walsh series coefficients have been carried on by Harmuth and de Buda [206], and by Siemens and Kitai [207,208,209]. The conversion between the (∞-dimensional) vectors of Fourier and Walsh coefficients of periodic waveforms is described by a non-singular square ∞-dimensional matrix multiplication which has relatively few non-zero elements. Later on, Kitai [210] showed that it is possible to truncate the vector of Walsh coefficients to 2^n terms and still obtain the correct Fourier coefficients of a frequency bandlimited waveform provided the (now finitely dimensional) conversion matrix is multiplied by a simple diagonal compensation matrix which tends to the identity matrix as $n \to \infty$. A practical use of this fact has been made in the construction of a waveform synthesizer [211] which uses only 8 of the first 32 Walsh functions.

The conversion theorem between Walsh and Fourier coefficients has also been used in the design of a frequency response analyzer [212]. Systems may be identified by exciting them with a sinusoid (at various frequencies), but instead of performing a direct Fourier analysis of the response, it is possible to perform a Walsh analysis followed by a conversion of the Walsh coefficients into Fourier coefficients. Experimentally, it was found [212] that a hardware (16 coefficients, 8 bit words) implementation using standard TTL devices performs satisfactorily in providing real-time frequency-response identification up to 60kHz;

a performance which was considered to be commercially attractive.

Identification of systems via intermediate domain processing have been studied even before by de Marsily, Emsellem and Poitrinal [213,214] who were unaware until recently [215] of the fact that they were using the Walsh transform. Faced with the problem of identifying the impulse response of hydrological (basin) systems from input u (rainfall on the entire basin) - output y (run-off at the outlet) data whose measurement had been contaminated by noise, they model the problem as one of linear deconvolution, $y(t) = \int_0^\infty u(t-\tau)\, h(\tau) d\tau$. Restricted to sampled, finite data, the problem appears in matrix form as:

$$\begin{bmatrix} y_0 \\ y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} \triangleq \underline{y} = U\ \underline{h} \triangleq \begin{bmatrix} u_0 & \cdots & u_{-n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ u_m & \cdots & u_{m-n} \end{bmatrix} \begin{bmatrix} h_0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ h_m \end{bmatrix} \qquad (4.5.1)$$

where, given the output data vector $\underline{y}$ and the input data matrix U, it is necessary to find a best approximation for the impulse response vector $\underline{h}$. For $n = m$, one is faced with a simple matrix inversion problem. The solution can be efficiently computed by using the FTT algorithm, but it is unsatisfactory due to its noise-induced instability (see also Chapter VII). The situation is less critical for $n < m$, where there is plenty of redundancy to reduce the effect of noise. Emsellem and de Marsily had the ingenious idea of looking for a least mean square error

optimal response $\underline{h}$ via its representation in a basis of orthogonal functions. For the sake of computational efficiency, they chose a system of step-like orthogonal functions which turned out to be the Walsh one. The interest of their approach consists in the fact that the representation is not computed at once, but iteratively, by successive projections on Walsh functions of higher index, so that constraints (e.g., non-negativity and smoothness) can be imposed on the solution after each step of the projection process. In a sense, this procedure is similar to the mini-max representation of Yuen, but with an additional dimension imposed by the constraints. The method has been proved to be very successful in practice, and de Marsily and Poitrinal have indicated its potentiality for studying non-stationary problems as well.

In fact, this had been already done, independently, in [216], where Walsh series expansions were used for identifying the dynamic characteristics of weakly oscillatory systems. The approach is somehow different nonetheless, because (1) it starts from the sampled version of the Wiener-Hopf equation

$$r_{xy}(k) = \sum_{m=0}^{M} h(n) r_{xx}(k-n); \quad k = 0, 1, \ldots, N, \tag{4.5.2}$$

where M is the regularization parameter, and (2) it projects both the output data $\underline{r}_{xy}$ and the unknown system function $\underline{h}$ onto the subspace of first N Walsh functions, thus obtaining the following system of equations

$$A_{ij}\, \hat{h}(i) = \hat{r}_{xy}(j); \quad A_{ij} = \sum_{p=0}^{N} \sum_{s=0}^{N} r_{xx}(p-s)\, \mathrm{wal}(i,s)\, \mathrm{wal}(j,p). \tag{4.5.3}$$

It is worth mentioning that experimental results [216] proved to be better for Walsh representations than other (Fourier, Chebyshev, Legendre and Laguerre) representations. As for a general discussion of linear time invariant system identification using random data and the Walsh transform, it is postponed until after the discussion of random data spectral analysis and basis restricted transformations.

Turning our attention from system identification to system control, an elegant and useful usage of Walsh analysis in a practical application has been made by Chen and Hsiao [217,218]. They are not so much interested in questions of abstract optimality but in problems of workable implementations of control theory. Given the often impracticability of theoretical solutions, they consider the problem of determining sub-optimal feedback laws for linear systems with quadratic performance criteria, and propose the implementation of (often time-varying) feedback gains by means of piece-wise constant gains obtained through Walsh representations. First, a Kronecker product formulation of the solution to the linear dynamic equations

$$\dot{\underline{x}} = A\,\underline{x} + B\,\underline{u} \tag{4.5.4}$$

is obtained in terms of Walsh functions. Such an approach is based on the fact that the integral of the $(2^n=)$ N-dimensional vector $\underline{w}$ of Walsh functions can be expressed approximately as a multiplication with a sparse matrix P_N,

$$\int \underline{w}\, dt = P_N \cdot \underline{w} \tag{4.5.5}$$

which makes it more advantageous to consider the Walsh representation of $\dot{\underline{x}}$ rather than $\underline{x}$. Then, given the Walsh representations of the m-dimensional vectors $\dot{\underline{x}}$ and $A\underline{x}_0$ and the k-dimensional vector $\underline{u}$ via the $m \times N$ matrices C and E and $k \times N$ matrix D respectively,

$$\dot{\underline{x}} = C \cdot \underline{w} \qquad A \cdot \underline{x}_0 = E \cdot \underline{w} \qquad \underline{u} = D \cdot \underline{w} \tag{4.5.6}$$

the solution is simply

$$\underline{x} = C\, P\, \underline{w} + \underline{x}_0, \tag{4.5.7}$$

where C is obtained from the matrix equation $C = ACP + E + BD$. Due to the sparsity of P and E, Chen and Hsiao were able to provide an elegant and efficient Fortran subroutine for solving this matrix equation. Alternatively, block pulse functions [88] could be used in place of Walsh functions [219,220].

Chen and Hsiao adopt a similar approach to the optimal control problem with quadratic performance index

$$= \frac{1}{2} \int_0^{t_f} (\underline{x}^t Q \underline{x} + \underline{u}^t R \underline{u}) dt. \tag{4.5.8}$$

Their solution is notationally too complicated to warrant reproduction; fact is that the method which they propose proves to be much simpler in analysis and easier in implementation than previous ones.

The same problem has been studied independently by Burkhardt [221,222]. He considered a quadratic time-variant cost functional and used Walsh functions as basis in a Ritz-method solution in order to

obtain a suboptimal extremal problem rather than be faced with a non-linear matrix Ricatti equation. He provides a solution similar to that of Chen and Hsiao, but his is a more complete study because he is also concerned with the convergence of the solution (it does converge), and he extends the results both to tracking problems (when the cost-functional weighs the deviation of the state vector from a prescribed trajectory) and to problems with final state constraints.

Before proceeding, it is necessary to review the theory of spectral representations of random processes, because it is essential for the forthcoming discussion concerning spectral estimation, system identification and filtering. Partly, this subject has been already discussed in Chapter III, where it was seen that the Walsh transform is a Karhunen-Loeve expansion for q-adic stationary random processes because it diagonalizes the autocovariance of the process so that the spectral coefficients are uncorrelated, and each of them can be estimated or processed independent of the others. In practical applications, it is rather seldom that one finds q-adic stationarity; but, in general, the corresponding Karhunen-Loeve transformation may be prohibitively complicated to implement [11]. It is then customary to perform some other transformations which are preferred to be unitary because the energy is invariant under them. They do not diagonalize the autocovariance matrix, and hence, do not decorrelate the spectral coefficients. But, if the errors introduced by neglect of off-diaongal coefficients are small, then what has been lost in theoretical goodness may be compensated by the

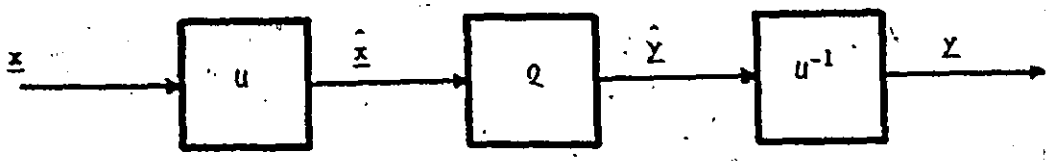


Figure 4.4. Illustration of the principle of intermediate domain processing.

practicability (efficiency) of implementation, with its gain in simplicity and speed of processing. This is the fundamental idea of basis restricted transformations (BRT) [134,178]. Figure 4.4 describes this process in which the random data vector is first transformed by the unitary transformation U, then a processing (not necessarily linear) is performed on the diagonal elements of the transform vector, followed by the inverse unitary transformation. In principle, one could use two separate unitary transformations [223], but it is more customary to use only one. The processing in the transform domain is described by a diagonal matrix operator

$$\underline{z} = U^{-1} \cdot Q \cdot U \cdot \underline{x}. \tag{4.5.9}$$

The restriction is twofold: (1) it imposes a suboptimal but computationally efficient transformation which is more or less independent of the signal, and (2) it presupposes a diagonal processing operator Q. If Q is linear, then one recaptures the concept of generalized Wiener filtering [137,180].

Following the same steps as in classical Wiener filtering theory, it is easy to show that for a linear diagonal matrix Q to provide the best estimate (in a mean square sense) of a vector $\underline{z}$ based on the observation of the vector $\underline{x}$ (both $\underline{z}$ and $\underline{x}$ are supposed to be of zero mean), it is necessary that its components should be

$$q_i = \frac{[U \cdot R_{xz} \cdot U^{-1}]_{ii}}{[U \cdot R_{xx} \cdot U^{-1}]_{ii}} \tag{4.5.10}$$

where R_{xz} and R_{xx} denote the cross- and auto-correlation matrices for $\underline{z}$ and $\underline{x}$. The residual error is found to be [134]

$$e^2 = trace\left(R_{zz} - \sum_i \frac{[U \cdot R_{xz} \cdot U^{-1}]_{ii}^2}{[U \cdot R_{xx} \cdot U^{-1}]_{ii}}\right). \qquad (4.5.11)$$

The error is, of course, a function of both U and the signals, and one can compare the theoretical efficiency of various unitary transforms (for estimating the same kind of signals !) by comparing the corresponding residual errors. As regards a comparison with the unrestricted optimal mean square error estimator, $H_{op} = R_{xz} \cdot R_{xx}^{-1}$, the suboptimality of the BRT scheme resides in the fact that it employes a Q designed as if U would diagonalize both R_{xz} and R_{xx}.

The vector formed of the diagonal components of $U R_{xx} U^{-1}$ is called the U-power spectrum of the random process $\underline{x}$. If the orthonormal functions defining the transformation U form a group under multiplication, then the U-spectrum admits a convolution interpretation; i.e. it may be considered to be the U-transform of a correlation function. This was the basic idea behind Gibbs' definition of "logical autocorrelation function" as a mean of local dyadic autocorrelation functions [55]; its Walsh transform being the Walsh-power spectrum P_w,

$$R_L(t) = \lim_{T\to\infty} \frac{1}{2T} \sum_{t=-T}^{T} 2^{-n/2} \sum_{\tau=0}^{2^n-1} x(t \oplus \tau) x(\tau) \overset{W}{\leftrightarrow} P_w(\lambda). \qquad (4.5.12)$$

With respect to the above definition, it is important to stress the fact that it refers not to dyadic stationary processes but to discrete real-

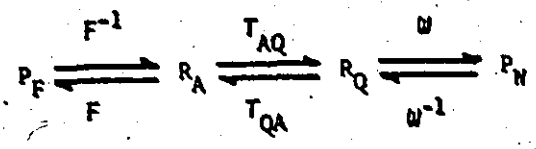


Figure 4.5. The relationship chain between Fourier and Walsh power spectra.

valued arithmetic stationary ones which are (usually) described by the arithmetic autocorrelation function $R_A(t)$ and its Fourier transform, the (Fourier) power spectrum P_F

$$R_A(t) = \lim_{T\to\infty} \frac{1}{T} \sum_{\tau=0}^{T-1} x(t+\tau)x(\tau) \overset{F}{\underset{F}{\rightleftharpoons}} P_F(\nu). \tag{4.5.13}$$

The relationship between the two autocorrelation functions has been studied by Gibbs and Pichler [146]; they showed that it can be expressed by a linear (non-singular) transformation T_{AL}. It can be generalized to q-adic analysis and the mean R_Q of local q-adic autocorrelation functions,

Th. 4.5a

$$R_Q(t) = q^{-n/2} \sum_{\tau=0}^{q^n-1} R_A[(\tau \overset{q}{-} t)-\tau]; \quad R_Q = T_{AQ} \cdot R_A \tag{4.5.14}$$

In fact, the theorem can be generalized even further to any unitary transformation and any random processes having enough symmetry to permit a reconstruction of the autocovariance matrix from its diagonal elements in a unitary representation.

Robinson [147] has studied the structure of the matrix T_{LA} (i.e. T_{AQ}^{-1} for the dyadic case) and found that it can be represented as a product of diagonal or at most block diagonal matrices

$$T_{LA} = \prod_{k=1}^{n} C(k) \cdot I_N \cdot E_N \tag{4.5.15}$$

where E_N is diagonal, I_N is the identity matrix of order N, and C(k) are block diagonal matrices with 2^{k-1} identical entries, each block being itself a sparse non-singular square matrix of order 2^{n-k+1}. Fast algorithms have been developed [142,143] for implementing the matrix multiplication by T_{LA} (or T_{AL}), so that it becomes computationally efficient to estimate the arithmetic autocorrelation of stationary data via an intermediate estimation of the Walsh power spectrum P_W followed by the required matrix conversion:

$$P_W(\lambda) = \frac{1}{M/N} \sum_{k=0}^{(M/N)-1} P_w^{(k)}(\lambda) = \frac{1}{M/N} \sum_{k=0}^{M/N-1} \left[\frac{1}{N} \sum_{i=0}^{N-1} x(i+kN)\,wal(\lambda,i)\right]^2 \tag{4.5.16a}$$

$$R_A = T_{LA} \cdot R_L = T_{LA} \cdot W^{-1} \cdot P_W. \tag{4.5.16b}$$

This procedure is, indeed, much faster than any other method; for example, it is shown in [142] that for $N = M = 2048$ the approximate numbers of multiplications/additions necessary with this method are $6 \cdot 10^3/6 \cdot 10^4$, compared with $2 \cdot 10^5/2 \cdot 10^5$ for the FFT method and $2 \cdot 10^6/2 \cdot 10^6$ for the direct lagged product method. An extra stage could be added to obtain an estimate of the Fourier power spectrum

$$P_F = FR_A = FT_{LA}W^{-1}P_W \tag{4.5.17}$$

which, for N sufficiently high, still proves to be more efficient than the FFT method, especially if M is greater than N.

$$T = \begin{bmatrix} h_T(0) & h_T(-1) & h_T(-2) & h_T(-3) & h_T(-4) & h_T(-5) & h_T(-6) & h_T(-7) & h_T(-8) \\ h_T(1) & (0) & (-1) & (-2) & (-3) & (-4) & (-5) & (-6) & (-7) \\ h_T(2) & (1) & (0) & (-1) & (-2) & (-3) & (-4) & (-5) & (-6) \\ h_T(3) & (2) & (1) & (0) & (-1) & (-2) & (-3) & (-4) & (-5) \\ h_T(4) & (3) & (2) & (1) & (0) & (-1) & (-2) & (-3) & (-4) \\ h_T(5) & (4) & (3) & (2) & (1) & (0) & (-1) & (-2) & (-3) \\ h_T(6) & (5) & (4) & (3) & (2) & (1) & (0) & (-1) & (-2) \\ h_T(7) & (6) & (5) & (4) & (3) & (2) & (1) & (0) & (-1) \\ h_T(8) & (7) & (6) & (5) & (4) & (3) & (2) & (1) & (0) \end{bmatrix}$$

$$Q = \begin{bmatrix} h_Q(0) & h_Q(2) & h_Q(1) & h_Q(6) & h_Q(8) & h_Q(7) & h_Q(3) & h_Q(5) & h_Q(4) \\ h_Q(1) & (0) & (2) & (7) & (6) & (8) & (4) & (3) & (5) \\ h_Q(2) & (1) & (0) & (8) & (7) & (6) & (5) & (4) & (3) \\ h_Q(3) & (5) & (4) & (0) & (2) & (1) & (6) & (8) & (7) \\ h_Q(4) & (3) & (5) & (1) & (0) & (2) & (7) & (6) & (8) \\ h_Q(5) & (4) & (3) & (2) & (1) & (0) & (8) & (7) & (6) \\ h_Q(6) & (8) & (7) & (3) & (5) & (4) & (0) & (2) & (1) \\ h_Q(7) & (6) & (8) & (4) & (3) & (5) & (1) & (0) & (2) \\ h_Q(8) & (7) & (6) & (5) & (4) & (3) & (2) & (1) & (0) \end{bmatrix}$$

Figure 4.6. Linear system matrices for a discrete domain of order 9.
T = time invariant system matrix (2N-1 distinct elements).
Q = 3-adic invariant system matrix (N distinct elements).

As presented above, the relationships $R_A \leftrightarrow R_L$ are exact and not approximate; it is true that P_N considers only the diagonal elements of the Walsh BRT transformation of the covariance matrix, but due to the symmetry (arithmetic stationarity) of the process and due to the abelian group structure of the transform, no information is lost such that the entire covariance matrix is recoverable (through the T_{LA} transform) from P_N. But there remains the problem of determining the errors involved in estimating P_N, and their propagation through the intermediate transforms.

Since it is possible, in principle, to estimate the secondary statistics of stationary data via an intermediate domain processing, it should also be possible to estimate TIL systems via estimates of intermediate models. The study of this problem was initiated by Pearl [141]. He concentrated his attention on a dyadic Walsh intermediate processing and a finite discrete case ($N = 2^n$ points), for which TIL and 2-TIL systems are describable by matrix equations

$$\underline{Y} = T \cdot \underline{X} \quad [T]_{ij} = h_T(i-j) \tag{4.5.18a}$$

$$\underline{Y} = D \cdot \underline{X} \quad [D]_{ij} = h_D(i \oplus j). \tag{4.5.18b}$$

In a first stage, the problem consists in searching the set $\{D\}$ of dyadic matrices for the element D_0 which best approximates the matrix T. The criterion of optimality is the minimization of the euclidean distance between the outputs of D and T over an ensemble of uniformly distributed input signals. The theory is similar to the one

concerning BRT estimations, and Pearl proves the diagonal matching theorem,

Th. 4.5b $D_0(T)$ is the dyadic matrix having the same diagonal elements as T under a Walsh similarity transformation

$$N^{-1}D_0(T)\cdot N = diag[N^{-1}\cdot T\cdot N] \tag{4.5.19}$$

the residual error being determined by the off-diagonal elements of $N^{-1}TN$ The theorem, which is in fact valid for any unitary transformation U, shows that a 2-TIL model may be obtained directly by measuring the response of the system to Walsh functions or dyadic stationary processes. But it must be mentioned that this projection from T onto D is not, in general, a one-to-one correspondence and, moreover, it is not an isomorphic mapping of TIL operator algebra into q-TIL operator algebra because, e.g., the cascade of two TIL systems does not imply a cascading of q-TIL models.

The second stage of the problem concerns the relationship between the "time" representation of the system and its intermediate domain model, and for the case of time-invariant and dyadic systems, Pearl has showed that

$$h_D(j) = \frac{1}{N}\sum_{m=0}^{N-1} h_T[(m\oplus j)-m]. \tag{4.5.20}$$

This idea can be generalized to any unitary transformation which has an

abelian group structure, i.e. the sets of rows or columns of the matrix U are closed under "element-wise" multiplication (to put it succintly - the unitary transformation has to be a generalized Fourier transform). For q-adic Walsh transforms (including the DFT as the 1-dimensional N-adic case), let us denote by Q the q-adic system matrix having $[Q]_{ij} = h_q(i \oplus_q j)$. (See Fig. 4.6.) Then,

Th. 4.5c the relationship between the impulse response of a LTI system and the impulse response of its optimal q-LTI model is

$$h_Q(j) = \frac{1}{N} \sum_{m=0}^{N-1} h_T[(m \oplus_q j) - m] = \frac{1}{N} \sum_{m=0}^{N-1} h_T[m - (m \oplus_q j)]. \qquad (4.5.20)$$

The third stage of the problem concerns the recovery of the original system from its intermediate model. As mentioned above, the projection (4.5.19) from {T} onto {Q} is not a one-to-one correspondence, such that it is not always possible to (uniquely) obtain T from Q. Pearl was able to show that recovery is possible for

a) TIL symmetric systems for which

$$h_T^s(t) = h_T^s(|t|); \quad t = -N+1, \ldots, -1, 0, 1, \ldots, N-1, \qquad (4.5.21a)$$

such that

$$h_D = H_{TD}\, \underline{h}_T^s = T_{AL}\, \underline{h}_T^s ; \qquad (4.5.21b)$$

and

b) TIL causal systems, for which

$$h_T^c = 0 \quad \text{for} \quad t \leq 0, \qquad (4.5.22a)$$

such that

$$\underline{h}_D = H_{TD}\,\underline{h}_T^C = \frac{1}{2}\,(T_{AL}+C)\underline{h}_T^C, \qquad (4.5.22b)$$

where T_{AL} is Robinson's conversion matrix, and C is a sparse matrix having 1's in the first column and 0's elsewhere.

Concerning this problem, it is possible to offer a more general statement:

Th. 4.5d the projection $Q_0(T) : \{T\} \rightarrow \{Q\}$ is a one-to-one correspondence for any TIL matrix T having at most N distinctive elements.

The proof is simple, making use of the fact that, by definition, T has at most $2N-1$ distinct elements (because $T_{ij} = h_T(i-j)$), whereas Th. 4.5c provides only N equations (because there are only N distinct elements in Q). Some extra $N-1$ h_T-structure equations (like the symmetry or the causality conditions) are necessary to permit the complete determination of T from Q. As presented here, this theorem invalidates Pearl's assertion that one cannot recover a TIL system from its circular (DFT) projection; his statement should have been qualified by specifying its validity only in the absence of the additional $N-1$ structure conditions for T. There remains, of course, the open problem of determining the goodness (bias and variance error) of estimates of dyadic (or q-adic) models (and hence the goodness of TIL system estimates) from finite and noisy measurements.

Intermediate Walsh domain processing of signals is another field of applications for Walsh analysis. The underlying theory has been enunciated in the context of basis restricted transforms [134,178] and of generalized Wiener filtering [137], but there were numerous applications preceding the formulation of the theory [224]. As shown before, the processing is definitely suboptimal unless the data has q-adic symmetry, its only redeeming feature being the lower cost and the higher speed of implementation. This is why such Walsh domain processings were applied almost exclusively to those problems where the factors of cost and speed are of paramount importance; the most noteworth examples being speech and image processing. Apart from considerations of efficient implementation, the use of Walsh transforms has proven to be rather disappointing. This can be explained first and foremost by the lack of "affinity" between the signals studied and the multi-dimensional structure of the Walsh transform. Secondly, the mean square error criterion with respect to which it has been customary to optimize processors using intermediate unitary transforms is not a good fidelity criterion from the point of view of human perception of either speech or image information content. It is expected that the introduction of adaptive control of processors will improve their capability. Some contributions in this direction have been already announced [224,225]. Also, attention has been brought recently to the importance of phase in image processing [226].

There are three clearly defined classes of objectives for intermediate domain processing.

(1) Data compression. This was discussed theoretically by Pearl [134] and Ahmed and Rao [184] amongst others. It encompasses two problems - one seeking an increased rate of data transmission/reception (see the example in [203]), the other seeking bandwidth compression. This is especially important for TV signals where real-time constraints are so stringent as to render unusable transforms which, although more suitable theoretically, are too slow to implement [227]. An interesting application of Walsh transform to adaptive bandwidth compression of video signals has been recently reported in [228]. Results on speech data compression were reviewed in [229]. The subject of signal multiplexing [230,231,232] should be also part of this class of applications.

(2) Signal recovery. This can be subdivided into two distinctive subclasses: one concerning noise reduction (the "convolution" problem), the other concerning distortion reduction and signal restoration (the "deconvolution - Fredholm integral equation" problem). Both aspects are important for image processing, and the use of Walsh transforms for achieving these objectives has been discussed by numerous researchers, among whom Andrews has the greatest contribution [233].

(3) Signal analysis. This refers to pattern recognition and feature extraction; also to the already discussed subject of spectral analysis. The use of Walsh transforms in such problems has been discussed in [234]. Unless the signal has affinity with q-adic symmetry, the prospects for Walsh analysis are not very encouraging. The Walsh transform has been used with some success for the study of gastro-intestinal signals

[235]; and recently, a report has been made [236] on the use of Walsh (and other) transforms for textual measurements in connection with the analysis of satellite images of the earth.

There are many other engineering problems for the solution of which it is possible to use Walsh analysis. Most interesting among them are those concerning the analysis of non-linear systems [237,238]. The efforts of Harmuth and his collaborators for the use of Walsh analysis in radar (concerning not the processing of the received signal but the emission and reception of Walsh waves) should also be mentioned [239,240, 241].

CHAPTER V

ON SIGNAL PROCESSING USING FAST CONVOLUTIONS IN FINITE FIELDS

5.1 CONVOLUTIONS AND FOURIER TRANSFORMS IN FINITE FIELDS

This chapter presents an investigation into some aspects concerning both the theory of convolutions in finite rings (or fields) and number theoretic transforms, as well as their applications in signal processing. As everywhere else throughout this dissertation, attention is concentrated exclusively towards OFF-line, transform-based methods of signal processing, with only scant remarks concerning sequential, ON-line implementations of signal processors.

The use of number theoretic transforms (as Fourier transforms in finite rings are sometimes called in engineering literature) for signal processing represents, in a sense, a third stage in the attempt to simulate linear processors via fast and error-free numerical computations of convolutions of functions in $\mathbb{C}[R]$. The first stage of this attempt consisted of the use of Fast Fourier transform algorithms to implement discrete finite Fourier transforms of complex-valued functions [24]. This implies the use of a finite group G as domain, or equivalently, the imposition of periodicity to all functions considered; whence the name of period (cyclical) convolutions [243]. But speed was sometimes insufficient and the computational errors were prohibitive for certain applications. To increase the first and to decrease the latters, there were attempts to replace the DFT with the Walsh or other, faster, transforms.

In the specific case of Walsh transforms, this implied the consideration of q-adic translation invariant systems, or, equivalently, the existence of a finite abelian group domain having the characteristics of a multi-dimensional space. As shown in the preceding chapter, the WT is definitely faster and less error-prone than the DFT, but its range of applications (and the effectiveness of its advantages) is attenuated due to its lack of affinity with real-world processing needs.

While the first stage of the effort to perform fast convolutions implied the use of a finite domain function space, the second stage implies the use of a finite co-domain function space. Restricted to an infinite domain, this case presents little interest from the point of view of fast numerical transform methods and has been little, if at all, discussed from the perspective of signal processing in the sense of filtering. The topic has been amply discussed in the context of coding and finite state machines (see the review article by Jury and Tsypkin [244], the book chapters of Gill in [117] and Kalman in [121], as well as the recent article of Sain [112]).

The third stage, the one to be discussed presently, implies the use of a function space having both a finite domain and a finite co-domain. The domain is an abelian group G, so that one can associate translation invariance with the function space. Also, the domain is finite, so that convolutions can be performed, in certain cases, via fast transforms. The co-domain is given the structure of a finite field K or, more generally, of a finite ring L, so that numeric computations are

performed exactly, with no truncation or round-off errors. But, of course, there are errors associated with simulating a $\mathbb{C}[R]$-convolution by one in a finite signal space $K[G]$. For one thing, the domain is finite and discrete, such that the convolutions are periodic. In addition, if one deals with sampled versions of analog signals, it is necessary to take into account the phenomenon of frequency aliasing [17,245]. For another thing, numeric computations cannot be performed with real numbers but only with rational ones. Also, the finiteness of the co-domain imposes an "amplitude aliasing" in the sense that the correspondence between the finite ring L and the infinite ring of rationals or integers (the only numbers to admit finite representations) is not bijective, each element of L representing an entire infinite (residue-) class of rationals or integers. (The "natural" injective mapping of the elements of a finite set A into the set of integers $\mathbb{Z}$ will be denoted by $i_A(\cdot)$.) Nevertheless, if the signals are bounded, it is possible to choose L such that the phenomenon of amplitude aliasing does not affect the end result of the convolution.

The mathematical theory related to convolutions and Fourier transforms in finite rings or in Galois fields [246] has been developed since long in the context of algebra and number theory [247,248], but its engineering significance has become apparent only within the recent development of numeric signal processing. The idea of performing signal processing in finite fields has been suggested by Knuth [249] and Good [250] amongst others. Pollard [251] and especially Nicholson [66] have formu-

lated the self-contained algebraic theory of Fourier-transforms in finite rings or fields and discussed the possibility of implementing them via fast algorithms.

Nicholson considered the function space $J[G]$, where J is an integral domain with identity and G is an abelian group of order N and exponent m. This function space is an algebra if addition is defined pointwise and multiplication is defined as a convolution

$$(f * g)(t) = \sum_{\tau \in G} f(\tau)\, T_{\circ}^{t} g(\tau) = \sum_{\tau \in G} f(\tau) g(t \circ \tau) \qquad (5.1.1)$$

$$f, g \in J[G]$$

The existence of an isomorphism F (called Fourier transform) of this algebra into a pointwise multiplication algebra is conditioned (necessarily and sufficiently) by:

(1) the existence of an inverse of $i_J^{-1}(N)$ in J, which is trivial if J is a finite field K containing $i_K^{-1}(N)$; and

(2) the existence of a primitive m-th root of 1 in $J^{\times}$, the multiplicative monoid of J.

The characters of G over $J^{\times}$ are the homomorphisms $\chi_\nu : G \to J^{\times}$. They form the dual group Γ which is isomorphic to G. They can be viewed as forming a complete set of "orthogonal" vectors, the Fourier transform being just a representation with respect to the characters as basis. The symmetries of χ_ν and the properties of the Fourier transform

they define are formally the same as those expressed in Chapter I for the function space $\mathcal{C}[R]$. If J is a Galois field K isomorphic to Z_k (k has to be a prime), the existence of F is conditioned by the requirement that m be a divisor of (k-1). If G is a finite cyclic group isomorphic to the additive group Z_N, where N divides k-1, the characters can be expressed as

$$\chi_\nu(t) = \zeta^{\nu t} \tag{5.1.2}$$

where ζ is an arbitrary, but fixed, primitive N-th root of $1 \in K$, and "νt" is understood as $i_\Gamma(\nu) i_G(t)$. The indexing of the characters by $\nu \in Z_N$ will be discussed in Section 6.4. If G is a q-adic group, then one deals with a "finite field Walsh-Fourier transform". If the co-domain is a ring isomorphic to Z_M, where $M = \prod_{i=1}^{m} p_i^{r_i}$ and the p_i's are primes, then the condition $\zeta^N = 1$ requires N to divide $O(M)$ = the greatest common divisor of $\{p_1-1, \ldots, p_m-1\}$. A finite ring Fourier transform can be implemented by a fast algorithm, of which the precise nature depends on whether or not N is highly composite.

The first engineering paper concerning the use of such transforms for signal processing purposes is due to Rader [159]. He showed that finite field Fourier transforms are engineeringly attractive if (1) N is highly composite to permit fast algorithm implementations, (2) multiplication by ζ^k is simple, and (3) the co-domain is chosen to facilitate binary implementations of "arithmetic" operations. Given these considerations, Rader proposed the use of fields modulo a Mersenne

prime (which are of the form 2^p-1 and are closely related to Euclid's perfect numbers [248]) or modulo a Fermat number. He was also concerned with the word-length constraints necessary to avoid amplitude aliasing. His work has been continued by Agarwal and Burrus [252], who detailed the use of "Fermat number transforms", and implemented them both in software and hardware. Fermat numbers are of the form $2^{2^t}+1$, and they are primes for $t = 1, 2, 3$, and 4. Computations modulo Fermat primes lend themselves to simple implementations via binary logic, but there remain the problems of (1) amplitude aliasing if only moderate length registers are used, (2) representation of (-1) by 2^t-long registers, and (3) the restriction to relatively short convolutions (N small). The latter difficulty may be alleviated by implementing a one-dimensional convolution via multi-dimensional Fourier transforms [253].

In a later review paper [35], Agarwal and Burrus broadened the discussion to rings other than those modulo a Fermat number in an attempt to obtain more convenient implementations and to avoid amplitude aliasing problems. Concerning these latter problems, Brule [254] has proposed an interesting scheme in which amplitude aliasing is avoided not by performing a difficult-to-implement convolution and Fourier transform in a very large ring, but by performing many easy-to-implement transforms in small fields or rings, followed by use of the Chinese remainder theorem to reconstruct the result in the large co-domain. In essence, the Chinese remainder theorem [248] states that any positive integer can be determined modulo a number $M = \prod_{i=1}^{m} p_i$ if the p_i's are distinct

primes and the residues of the unknown integer with respect to each of them are known. Reed and Truong [255], who showed how to perform convolutions between sequences of complex numbers via fast Fourier transforms in Galois fields $GF(k^2)$, have also proposed a similar scheme [256] based on the Chinese remainder theorem. Recently, Justesson proposed the use of number theoretic transforms for decoding Reed-Solomon codes [217].

The analysis to be presented in the following sections of the chapter is not concerned with problems pertaining to simple hardware or software implementations of number theoretic transforms, but with conceptual problems pertaining to the significance of filtering and the theory of linear processors in finite rings or fields. Section 5.2 discusses signal processing in finite fields, while Section 5.3 introduces harmonic differential calculus in finite fields.

5.2 SIGNAL PROCESSING IN FINITE FIELDS

The problems of signal processing in the space of functions into a Galois field resemble, up to a certain point, those of coding theory or the theory of finite state automata. This is so because their underlying mathematical discipline is the same: the study of transformations on a signal space $K[S]$ whose elements are functions from a countable set S into a finite field K (or into a finite ring in a more general treatment). In practice, S is either the set $\mathbb{Z}$ of integers (with ordinary addition, isomorphic to the infinite cyclic group) or a subset

Z_N (with addition modulo N,[2], isomorphic to a finite cyclic group). The feature distinguishing between the two theories mentioned above is conceptual, and hence major: it consists in the different nature (origin and quality) of the information content of the signal to be processed. In coding theory [258], the signals are (at least initially) "deterministic", being created as elements of a fixed alphabet (subset of the signal space), according to certain stringent procedures. It is only afterwards, in the course of being propagated along the channel, that these signals are corrupted by noise; so the task of information extraction is to identify each received signal with one of the elements of the prescribed alphabet. The situation is totally different in the context of filtering theory, where the information carried by the signal does not pertain to its original creation but to the "distortion" (shaping) introduced by the phenomenon or system under investigation; so the task of processing the signals consists in separating (for analysis) the "deterministic distortion" introduced by that particular system from some other, unwanted, distortions (noise).

A pre-requisite to any meaningful discussion of filtering in an optimal sense requires the formulation of an error criterion. Unfortunately, the error criterion based on the Hamming distance [188], which is the cornerstone of coding theory is of no use in filtering theory, where there is no finite prescribed alphabet. Attempts at defining a mean square error criterion have been, so far, unsuccessful, because it has not been possible to define a functional $\rho : K[S] \times K[S] \to K$

fulfilling the requirements for an inner product, so that it could be used to obtain a norm, and hence a metric, on the signal space. The functional defined by

$$\langle f , g \rangle = \sum_{t \in S} f(t)\, g(t) \tag{5.2.1}$$

has some of the properties of an inner product, but it does not satisfy the postulate that $\langle x , x \rangle = 0$ should imply $x = 0$. The root of the problem resides in the fact that it is not possible to define a concept of distance within a finite field unless the injective mapping $i_K : K \to \mathbb{Z}$ is employed to transcend the problem to the realm of "non-finite" numbers. In that case, a potential candidate is the Lee metric [259], defined by

$$d(f_1,f_2) = \sum_{t \in} abs(i_K(f_1(t)) - i_K(f_2(t))), \tag{5.2.2}$$

where $abs(\cdot)$ denotes the absolute value function. Similarly, one can define a mean square distance. Another possible avenue appears if K is a Galois field $GF(p^n)$, in which case, the trace $tr_{K/F} : K \to F \triangleq GF(p)$,

$$tr_{K/F}(\cdot) = (\cdot) + (\cdot)^p + \ldots + (\cdot)^{p^{n-1}}, \tag{5.2.3}$$

has all the properties of a norm, with the triangular inequality replaced by equality [248]. This represents a generalization of the dyadic norm [107] introduced by Gibbs for $p = 2$. There remains much work to be done concerning the use of such error criteria for designing optimal processors in finite fields. No mention has been made in the literature with regard to this problem of optimal processing (with the obvious exception of [34]).

the general attitude being that of using finite field transforms only for simulating convolutions between real- or complex-valued functions. Thus, attention has been accorded not to the problem of finding the optimal finite field processor, but to that of finding the finite field or ring which permits a convenient (implementation-wise) simulation of real number convolutions; whence all the efforts to avoid amplitude aliasing.

As for the analysis of finite systems [117,121], a state space description is very powerful. Massey and Sain [260] were among the first to attempt finding explicit interconnections between the theories of error-correcting codes, finite state machines and continuous systems. But a state space description implies causality, which presupposes an ordered domain; this is not the case for processors which have been specifically designed to take advantage of fast Galois field Fourier transform algorithms. Alternatively, one can use a transform approach with the help of the finite discrete Laplace transform [261,262,263]. The relationship between such a transform and a Fourier transform in a finite field has been analyzed by Cohn-Sfetcu and Gibbs [34].

Systems in a finite signal space $K[G]$ which are characterized by linear operators commuting with each translation operator T are said to be linear translation invariant systems; and the relationship between input u and output y can be described by a convolution product of the form (5.1.1),

$$y(t) = (h * u)(t) \triangleq \langle h , T_t^t u \rangle, \qquad (5.2.4)$$

where h is called the impulse response of the system. As is to be expected, the input-output analysis of Galois-field linear translation invariant systems appears to be closely analogous to the classical analysis of linear "time invariant" systems describing relationships between complex-valued signals:

(i) the system response to an impulse δ, defined by $\delta(1) = 1$, $\delta(t) = 0 \quad (t \neq 0)$, is h;

(ii) the Fourier transform maps the convolution product $y = h * u$ into the pointwise product $\hat{y} = \hat{h} \cdot \hat{u}$, where $\hat{h}$ is called the transfer function of the system;

(iii) the characters are the eigenfunctions of the system, in the sense that each is undistorted by it apart from a constant factor independent of t: if $u = \chi_\nu$ then $y = \hat{h}(\nu) \cdot u$.

Random signals can be easily incorporated into this transform approach, especially weakly (N-adic) stationary signals, that is, signals characterized by a constant expected value and by covariance functions that depend only on the difference modulo N of the arguments; for example,

$$r_{vv}(t_1,t_2) \triangleq E[v(t_1)\, v(t_2)] = r_{vv}(t_2 \ominus t_1). \tag{5.2.5}$$

Power- and cross-spectral density functions are defined as the Galois-field Fourier transforms of the auto- and cross-covariance functions, respectively; and input-output relationships for linear (N-adic) shift variant systems take classical forms, for example,

$$r_{uy} = h * r_{uu} \underset{F^{-1}}{\overset{F}{\rightleftharpoons}} \hat{r}_{uy} = \hat{h} \cdot \hat{r}_{uu}. \qquad (5.2.6)$$

At first sight, these N-adic stationary processes may seem nothing more than mathematical curiosities; this is not so; such a process may be viewed as the output of a linear (N-adic) translation-invariant system excited by a white noise process (whose covariance function is non-zero for equality of arguments: $r_{ww}(t_1,t_2) = P_0\ \delta(t_2 \dot{-} t_1)$).

The problem is that all these relationships in the "transfer domain", although very simple from a mathematical point of view, have not acquired any direct engineering significance [35]. This explains why, unlike complex-field filters, which are usually designed in the frequency domain, finite field filters are designed in the time domain, by specifying the impulse response h (obtained from complex-field model considerations in the frequency domain); and measures (painful as regards complication and cost of hardware and software) have to be taken to avoid amplitude aliasing. It was this search for an engineering significance of transform domain operations which prompted the development of harmonic differential calculus in Galois fields which is to be presented in the next section. It also prompted the exploration of the concept of generalized frequency as a measure of the speed of variation of a function (Chapter VI). Since each character is an eigenfunction of a linear translation invariant system, it is natural (and important) to look for the linear operators whose eigenfunctions are the characters.

5.3 HARMONIC DIFFERENTIAL CALCULUS IN FINITE FIELDS

There have been numerous attempts at defining differentiation in finite fields or, more generally, to define it for any spaces of discrete-valued functions. For instance, Nathanson [264,265], who has been preoccupied mostly with the study of periodicity in discrete sequences, has defined some difference and integration operators for binary sequences. More interesting and of a much wider range and scope has been the research carried by Thayse, Deschamps and their collaborators. They made contributions not only in developing Akers' original idea [266] of Boolean difference [267.108] (see also the independent work of Reed [268] and Rudeanu [269]), but also in developing a unitary theory of discrete functions and difference operators [270] given either a lattice [195] or a ring or field structure [196] for the functions' co-domain. Thayse et al consider the functions' domain to be a (finite) direct sum of finite sets S_i; and they view a difference operator as a measure of the function's variation due to a certain shift of the variable. "Variation" should be understood in the sense of the structure of the co-domain: conjunction or disjunction if the co-domain is a lattice, difference modulo m if the co-domain is the ring $\mathbb{Z}_m$, etc.... In general, Thayse and his collaborators are mostly interested in generalizing aspects of Boolean calculus to the theory of multiple-valued switching circuits [197,198]. For instance, they provide counterparts of the concepts of prime implicants and prime implicates, and discuss their extraction by a technique very similar to that of Lechner, but not in the transform domain. It is interesting to note that, although he was aware of the relationship between Boolean

differential calculus and harmonic analysis [108,194], Thayse, who developed a theory of operators on functions from $GF(p^n)$ into $GF(p)$ which have many properties similar to those of the classical differentiator [271], did not pursue the matter of the relationship differentiation - harmonic analysis beyond the case $GF(2^n)$. This overlook may be explained by the perspective adopted - that of switching circuits, where there is more interest in Taylor-McLaurin type expansions of functions rather than in convolutions or their Fourier representation. The same is true regarding the work of Benjauthrit and Reed [27], who, independently of Thayse et al, have extended the concept of Boolean difference to spaces of functions from $(GF(p^n))^m$ into $GF(p^n)$. The outlook adopted herein is different because, as mentioned before, we are specifically interested in convolutions, Fourier transforms and operations in the transform domain, looking as we are, for those operators whose relationship with the characters is similar to the one between the Newton-Leibniz differentiator and the exponential functions.

The harmonic differential calculus in finite fields has been developed following the same principles as those adopted by Gibbs in his definition of harmonic differential calculus in the complex-number field. In fact, given the existence of a similar formalism describing Fourier transforms in either the complex-number field or in a finite field, it is to be expected that a formal similarity exists between the harmonic differential calculus in the complex-number field and that in a finite field. This is indeed the case, as it has been proven by Cohn-Sfetcu and

Gibbs [34]. The presentation to follow is restricted to only enunciating the basic theorems, their proofs, which are simple enough, being, at most, only sketched.

Def. 5.3a A harmonic differential operator $\mathcal{D}$ in $K[G]$ is any linear operator in $K[G]$ whose set of eigenfunctions is the set of characters of G over K. There are many such operators in each such space. Uniqueness is ensured by requiring that the eigenvalue corresponding to each character χ_ν is uniquely related to the index ν of the character,

$$\mathcal{D}\chi_\nu = \beta(\nu)\chi_\nu. \tag{5.3.1}$$

The map $\beta : G \to K$ is analogous to the contraction mapping introduced by Gibbs and Ireland in defining harmonic differentiation for complex-valued functions on a finite abelian group [73]. It is customary to define β via the natural injections i_K and i_G

$$\beta(\cdot) = i_K^{-1}(i_G(\cdot)); \tag{5.3.2}$$

this lends "credibility" to the usual notation

$$\mathcal{D}\chi_\nu = \nu\cdot\chi_\nu. \tag{5.3.3}$$

Th. 5.3a The harmonic derivative $\mathcal{D}f$ of a function $f \in K[G]$ whose Galois-field Fourier transform is $\hat{f}$ can be computed by

$$(\mathcal{D}f)(t) = \mathcal{D}(N^{-1}\sum_{\nu\in G}\hat{f}(\nu)\,\chi_\nu(t)) = N^{-1}\sum_{\nu\in G}\nu\cdot\hat{f}(\nu)\cdot\chi_\nu(t). \tag{5.3.4}$$

Th. 5.3b The harmonic derivative of a function $f \in K[G]$ is $0 \in K[G]$ if and only if f is a constant function.

Th. 5.3c The harmonic differentiator commutes with each translation operator:

for each $\tau \in G$,

$$DT^{\tau} = T^{\tau}D \tag{5.3.5}$$

Th. 5.3d As an immediate corollary of Th. 5.3a, if the Fourier transform of f is $\hat{f}$, then that of Df is the function ϕ defined by

$$\phi(\nu) = \nu \cdot \hat{f}(\nu). \tag{5.3.6}$$

Th. 5.3e The harmonic derivative of the convolution product of two functions is equal to the convolution product of either of the functions and the derivative of the other (this follows easily from Theorem 5.3d):

$$D(f * g) = (Df) * g = f * (Dg). \tag{5.3.7}$$

The harmonic derivative can be extended to arbitrary orders $s \in \mathbb{Z}$ via the definition of the harmonic delta function δ,

$$\delta(t) = N^{-1} \sum_{\nu \in G} \chi_{\nu}(t) = \begin{cases} 1 & t = 0 \\ 0 & t \neq 0 \end{cases}, \tag{5.3.8}$$

and its derivatives of order s

$$D^{s}\delta = N^{-1} \sum_{\nu \in G-\{0\}} \nu^{s}\chi_{\nu} \tag{5.3.9}$$

thus,

$$\mathcal{D}^s f = (\mathcal{D}^s \delta) * f = N^{-1} \sum_{\nu \in G-\{0\}} \nu^s f(\nu) \chi_\nu . \qquad (5.3.10)$$

As can be seen, harmonic differentiation has many of the features of Newton-Leibniz differentiation. But, in general, harmonic differentiation on $K[G]$ does not obey the product rule

$$(fg)' = fg' + f'g. \qquad (5.3.11)$$

Gibbs [107] has proved that harmonic differentiation in the dyadic field obeys a generalized form of the product rule. It is conjectured here that a generalized product rule holds whenever $G = (GF(p^n))^a$ and $K = (GF(p^n))^b$, where n, a and b are integers and p is a prime number.

As for the relationship between the harmonic differentiator in a finite field and the increment operator $\Delta^\tau : G \times K[G] \to K[G]$,

$$\Delta^\tau f(t) = T_+^\tau f(t) - f(t), \qquad (5.3.12)$$

it can be shown that,

Th. 5.3f the harmonic derivative on $K[Z_N]$ can be expressed as

$$\mathcal{D}f = \sum_{\tau \in Z_N'} \frac{\chi(M,\tau)}{1-\chi(1,\tau)} (T_+^\tau f - f) = \sum_{\tau \in Z_N'} a_N(\tau) \Delta^\tau f, \qquad (5.3.13)$$

where $Z_N' = Z_N - \{0\}$, and M denotes the residue class containing the greatest integer less than or equal to $\frac{1}{2}(N+1)$. To prove Theorem 5.3f, one has to show that expression 5.3.13 satisfies Def. 5.3a. The case $\nu = 0$ is trivial because $\chi(0,t) = 1$, and the derivative of a constant

function is 0. In case that $\nu \neq 0$, then

$$\sum_{\tau \in Z_N'} \frac{\chi(M,\tau)}{1-\chi(1,\tau)} [\chi(\nu,t\$\tau)-\chi(\nu,t)] = -\chi(\nu,t) \sum_{\tau \in Z_N'} \frac{\chi(M,\tau)}{1-\chi(1,\tau)} [1-\chi(\nu,\tau)]$$

$$= -\chi(\nu,t) \sum_{\tau \in Z_N'} \chi(M,\tau) \sum_{\mu=(-1+\delta gn(n))/2}^{n-(1+\delta gn(n))/2} \chi(\mu,\tau)^{\delta gn(n)}$$

$$= -\chi(\nu,t) \sum_{\mu=\ldots} \sum_{\tau \in Z_N'} \chi(\mu\$M,\tau)^{\delta gn(n)}$$

$$= -\chi(\nu,t) \sum_{\mu=(-1+\delta gn(n))/2}^{n-(1+\delta gn(n))/2} i_K^{-1}(-1)^{\delta gn(n)}$$

$$= \beta(\nu)\, \chi(\nu,t)$$

where $n \triangleq i_G(\nu)$ and $y \triangleq i_G(\eta)$. In this calculation, we have used the facts

(i) that

$$\chi(\nu,t) = \chi(\nu,1)^t = \chi(1,t)^\nu; \tag{5.3.14}$$

(ii) that, for α in any field, and for $n \in \mathbb{Z}-\{0\}$,

$$1-\alpha^n = (1-\alpha) \sum_{y=(-1+\delta gn(n))/2}^{n-(1+\delta gn(n))/2} \alpha^{y\delta gn(n)}; \tag{5.3.15}$$

and (iii) that

$$\sum_{\tau \in Z_N'} \chi(\nu,\tau) = i_K^{-1}(N\delta(\nu)-1). \tag{5.3.16}$$

For G, an n-dimensional q-adic group, an expression similar to that in [72] can be proven likewise.

As can be seen from Th. 5.3f, the harmonic differentiator is a "close cousin" of the differential operators defined by Thayse et al or by Benjauthrit and Reed for $f : GF(p^n) \to GF(p)$. In a sense, it is both a particularization and a generalization of them: a particularization because it is a weighted average of the function's "variation" precisely along the unit directions in the multi-dimensional domain $\mathbb{F}_p^n$; and a generalization because it is valid for function spaces other than those in which the domain and co-domain are so directly related as in the case of multiple-valued logic differential calculus.

With regard to harmonic differential equations, we have the easy

Th. 5.3g the homogeneous linear first-order harmonic differential equation (initial value problem)

$$Dx = a \cdot x \qquad x(t_o) = x_o \neq 0 \in K \tag{5.3.17}$$

is soluble if and only if there is in G an element equal to $i_G^{-1}(i_K(a))$; if a solution exists, it is unique and is given by

$$x(t) = x_o \chi(a, t \ominus t_o). \tag{5.3.18}$$

Th. 5.3h The non-homogeneous harmonic differential equation

$$Dx = a \cdot x + u \tag{5.3.19}$$

is soluble if and only if either of the following conditions holds:

(i) there is no element of G such that it is equal to $i_G^{-1}(i_K(a))$;
(ii) there is such an element η, but $\hat{u}(\eta) = 0$; if a solution exists, it is given by

$$x = \sum_{\nu \in G} (\nu \circ \eta)^{-1} \hat{u}(\nu) x_\nu . \tag{5.3.20}$$

In the first case, the solution is unique while in the second one the indeterminate coefficient of x_ν is fixed by an appropriate initial condition. The same approach can be applied to equations of higher order.

It is also possible to develop a harmonic state-space analysis of linear processors in $K[G]$. This has been done by Cohn-Sfetcu and Gibbs [34], who solved the multiple input-multiple output system equations

$$\begin{aligned} D\underline{x} &= A\underline{x} + B\underline{u} \\ y &= C\underline{x} + E\underline{u}, \end{aligned} \tag{5.3.21}$$

The results were formally the same as those in Section 3.4, and will not be repeated here. There remains the problem of interpreting these results, because, as in the case of processors in $\mathcal{C}[G_q]$, the finite order of the functions' domain makes difficult the introduction of the concept of causality and, hence, of the concept of state as it is generally understood in systems science.

5.4 CONCLUSIONS

The investigations presented in this chapter should be viewed only as a probing into the problems of signal processing in finite fields; as an initial step on the way towards the long-range objective

of freeing optimal digital signal processing from the constraint of referring it to some ideal optimal analogue model. The present objective was to clarify the concepts of filtering signals which are elements of function spaces in which both the domain and the co-domain are finite discrete sets, albeit having the structure of an abelian group and a field respectively. It was in the scope of providing some engineering significance for operations performed in the transform domain related to certain Fourier transforms that the harmonic differential calculus has been developed. Much work remains to be done. On one hand, there should be an exploration for those real life phenomena which may be best modelled as systems operating in finite fields. An example is the work of Remler [273] who studied pattern recognition by convolution techniques in finite fields as a model for sensory perception of human bodies. On the other hand, both the relationship between harmonic differential calculus and the differential calculus of Thayse, as well as their utility in connection with the analysis and synthesis of multiple-valued logic circuits or other similar circuits operating with discrete functions remain to be explored.

CHAPTER VI

ON A CONCEPT OF GENERALIZED FREQUENCY

6.1 INTRODUCTION

The development of digital signal processing techniques and the introduction of "new" discrete finite transforms have been associated by an effort at providing meaning to the spectral coefficients obtained through such transform representations. As usual, this meaning has been sought in terms of familiar concepts, emphasis being placed on the concept of frequency. This is only natural, given the resemblance between many of these new transforms and the classical Fourier representation. Of these transforms, the most discussed in the above-mentioned sense is the Walsh transform, in connection with which Harmuth [37,88] has defined the concept of sequency, i.e. the ordering of the dyadic Walsh functions according to their number of sign-changes. The claim that sequency is a generalization of the concept of frequency has played an important role in the development and proliferation of applications of the Walsh transform. The problem of meaning and ordering of spectral coefficients has been also raised in connection with number theoretic transforms [35].

The concept of frequency is seldom the concern of mathematicians, but it is a fundamental tool for applied scientists and engineers. Unfortunately, the word "frequency" is one of the most taken-for-granted words

in engineering terminology; assumed to be known from elementary trigonometry, it is almost never defined or explained in a general mathematical context. Applied scientists consider frequency as a measure of the speed of variation of a function; but this comes about only indirectly; through the fact that a rapidly varying signal has a Fourier transform representation with large coefficients for the complex exponentials of large index.

This chapter analyzes the concept of frequency as it is currently understood, and provides a consistent and unique criterion for generalizing this concept to function spaces other than the customary space of complex-valued functions on the real line. The crucial starting point of the analysis consists in accepting the applied scientists' consideration of frequency as a parameter which, by ordering the set of Euler functions, $\{\exp(j2\pi\nu t);\ \nu,t \in R\}$, serves to measure the speed of variation of a signal.

The theses expounded in this chapter are:

1) that the concept of frequency cannot be defined in terms of global characteristics of <u>any</u> function, but only in terms of ordering a set of functions forming the basis of a transform representation;

2) that the concept of generalized frequency should be defined only with respect to the group characters forming the orthogonal set used for the appropriate type of Fourier transform, so as to provide a unique and consistent ordering of these functions;

3) that the generalization should be made by extrapolation of the most essential features of the concept of frequency as it is already

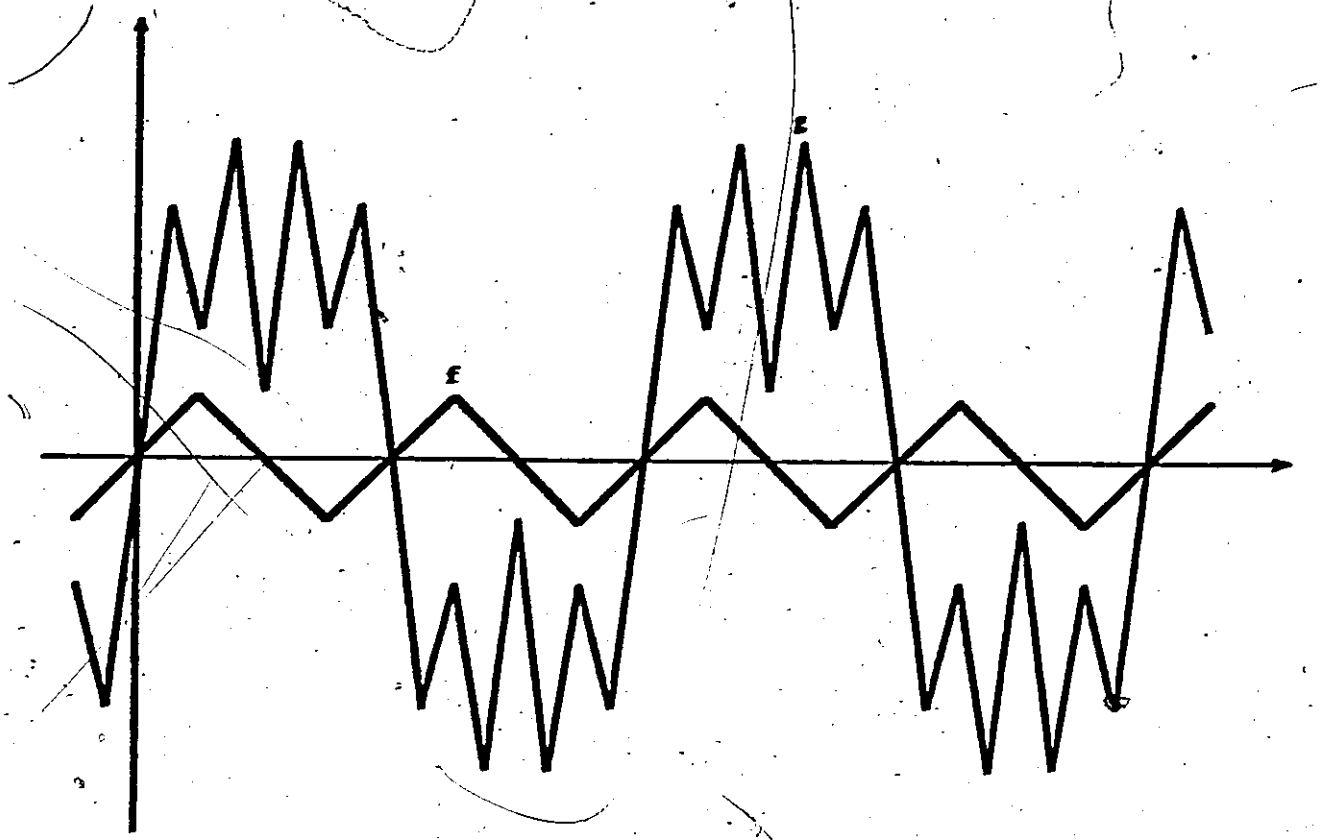


Figure 6.1. Illustration of the fact that the number of zero-crossings do not reflect the speed of variation of a function (even if periodic and of zero mean).

understood and used in connection with the set of complex exponentials and classical Fourier analysis.

The infinite variety of functions precludes from the outset any possibility of generalizing the concept of frequency on the basis of some basic global characteristic of any function. All attempts at achieving this objective while preserving the understanding of frequency as a measure of the speed of variation have met with certain failure. Such an attempt has been made recently by Ahmed et al [63], who proposed to define "generalized frequency as one-half the average number of zero crossings per unit time". First of all, a function may vary wildly and still have no zero crossing. Hence, a first restriction must be imposed on the class of functions which may be ordered by such a criterion. Let this restriction be that of zero-mean. But even then, the definition is fallacious. To see this, it is sufficient to look at Figure 6.1 and notice that although the function f has a higher number of zero crossings, it has a much lower speed of variation than the function g.

It is more appropriate to define a concept of generalized frequency only with respect to the ordering of a (restricted) set of functions which may be comparable from the point of view of speed of variation. In their turn, these functions should form a basis for the representation and characterization of a sufficiently large class of signals.

Such an attempt at defining generalized frequency has been made by Naylor [123] in connection with the analysis of time-varying, discrete-

time linear systems. He used two orthogonal transforms to obtain a diagonal "generalized frequency transfer function matrix". The key to his idea of generalized frequency is the "association of a frequency increment Δf" with each of the functions (column vectors) forming the basis of the "direct transformation". But no meaning is sought or given to such an arbitrary ordering of the basis functions (it is arbitrary because the orthogonality of the transform is an invariant under any permutation of the basis functions).

If meaning is to be given to an ordering of the basis functions, then this meaning should be sought in connection with the principle of creation of the respective set of basis functions. This reasoning, together with the relationships between groups, linear displacements, characters and Fourier-type function representations lie at the foundation of the second thesis enunciated above.

In addition, and subsequent to Harmuth's efforts, there have been numerous other attempts at studying the concept of sequency and defining a concept of generalized frequency [94,75,274,275,276,86,87,107,79]. At best, all of these efforts have met with only partial success because either they were not preceded by a proper and/or complete analysis of the concept of frequency as it is presently understood, or they were restricted to transform representations which simply do not admit a frequency-type interpretation for the ordering of its basis functions. Accordingly, it is first necessary to investigate the set of complex exponentials and the concept of frequency to reveal what it really means, i.e. to determine

which of its properties are essential enough to act as invariants under generalization.

Such a study, based on the previous partial analyses of [79] and [276], has been recently reported by Gibbs and Cohn-Sfetcu [17], and forms the subject of Section 6.2. Section 6.3 is concerned with a definition of generalized frequency. Section 6.4 studies the case of one-dimensional discrete Fourier transforms in complex number and finite number fields, while Section 6.5 is preoccupied with the relationship existing between sequency and the concept of generalized frequency defined in Section 6.2.

6.2 CLASSICAL FOURIER ANALYSIS, COMPLEX EXPONENTIALS AND FREQUENCY

In an elementary sense, the word "frequency" has the meaning of the inverse of the period of any periodic function, whereas in applied sciences, this word is more intimately connected with the trigonometric or complex exponential functions.

Applied scientists usually describe natural phenomena via real or complex-valued functions of a continuous variable (called, henceforth, with no loss of generality, time). This variable spans the real line, which has the structure of an additive group, the operation of addition representing the physical concept of displacement in a homogeneous (one-dimensional) space. It has been shown in the previous section that the concept of frequency can not be defined as a global parameter of any type of function. Such a concept can be defined only relative to a meaningful ordering of a set of functions forming the basis for a function represent-

ation. The primacy of the set of complex exponentials among all the sets of orthogonal functions resides in the fact that, in a first stage [158], the relationships between natural phenomena were (and are) modelled by linear time-invariant systems in the function space $\mathcal{C}[R]$. These systems are described by convolution equations, and their eigenfunctions are the complex exponentials. A representation in terms of complex exponentials is very convenient because, as eigenfunctions of linear time-invariant systems, they are not modified by such systems apart from an "amplification" of their modulus and a shift of their phase, both independent of the time.

From a mathematical viewpoint, these properties of the complex exponentials appear as a consequence of their being the characters of the additive real group. As such, they form the basis for a Fourier representation which performs an isomorphic mapping between commutative function algebras with convolution and pointwise multiplication as the respective products.

Linear time-invariant systems are also described by linear differential equations with constant coefficients. The eigenfunctions of the Newton-Leibnitz differential operator are the exponential functions:

$$Df(t) = \frac{d}{dt} f(t) = \lambda f(t) \quad \& \quad Df(0) = 1 \Longrightarrow f(t) = e^{\lambda t}. \qquad (6.2.1)$$

The additional restriction that the eigenfunction be bounded compels λ to be imaginary and $f(t)$ to be a Euler function with an absolute value identically equal to 1:

$$|f(t)| < M \Longrightarrow f(t) = e^{\lambda t} = e^{j2\pi\nu t} \triangleq \chi_\nu(t) \qquad |f(t)| = 1. \qquad (6.2.2)$$

These complex exponential functions are ordered according to the real number ν, called their frequency. It is customary to call $\omega \triangleq 2\pi\nu$ the angular frequency and $\phi \triangleq \omega t = 2\pi\nu t$ the phase of the complex exponential expression $\exp(j2\pi\nu t)$. More generally, any complex-valued function can be represented in polar form as

$$f(t) = \Lambda_f(t)e^{j\phi_f(t)} \qquad (6.2.3)$$

where the modulus Λ_f is a non-negative real-valued function and the phase ϕ_f is a function whose range is either the interval $(-\pi,\pi]$ or $[0,2\pi)$. The number e^j is viewed as the 2π-th order root of unity, since $(e^j)^{2\pi} = 1$.

NOTE: For reasons of mathematical elegance and consistency with Walsh-Fourier representations as well as with Fourier representations of functions in finite fields, it is advisable to consider the polar representation in terms of an imaginary phase function $\psi_f \triangleq j\phi_f$ and the corresponding imaginary angular frequency $\lambda \triangleq j\omega \triangleq j2\pi\nu$, and to regard the number e as the $j2\pi$-th root of 1. Thus, $f = \Lambda_f e^{\psi_f}$. This contention is supported by the fact that in all relations expressing FT properties the factors j and ν (or ω) are indissolubly connected and always appear together. This fact has been recognized in Laplace transform theory, where a Fourier transform is viewed as the restriction of the Laplace transform to the imaginary axis.

From the point of view of polar-form expressions of complex-valued functions, the Euler functions enjoy the distinction of being the only functions of unit modulus having a constant local frequency. The local angular frequency is defined as the slope (infinitesimal relative variation) of the phase function with respect to time [11]. There have been many discussions regarding the concept of instantaneous frequency [277,278], but the controversy does not apply in this case because the analysis is restricted to functions which are group characters and have a non-zero constant ($=1$) modulus. This linear variation of the phase is one of the remarkable features of the complex exponential, and can be used for defining "generalized characters" [79]; as well as for defining generalized frequency as a scalar proportional to the slope of the phase.

$$\nu = \frac{1}{2\pi j}\left(\frac{d}{dt}\,\Psi_\chi\right)(t) \triangleq \frac{1}{2\pi j}\lim_{\tau\to 0}\frac{\Psi_\chi(t+\tau)-\Psi_\chi(t)}{\tau} \triangleq \frac{1}{2\pi j}\,(\Delta\Psi_\chi)(t) \qquad (6.2.4)$$

(Use is made of the two symbols $\mathcal{D}$ and Δ to distinguish between the harmonic differential operator and the slope operator, which in general do not coincide.) Defined likewise, frequency can be viewed as a measure of the speed of variation of the characters, and can be used for ordering them accordingly.

Alternatively, since all the complex exponentials have identical, unit modulus, their speed of variation can be described by the number of rotations around the origin which their polar vectors make in a unit of time. This concept of frequency as the number of rotations of the polar vector in the complex plane is strongly connected with the "winding number"

concept. This concept is used in complex analysis [279] to indicate the number of times a closed (piecewise differentiable) curve γ in the complex plane z winds around a fixed point w not on the curve;

$$n(\gamma,w) = \frac{1}{2\pi j}\int_{\gamma}\frac{dz}{z-w} = -n(-\gamma,w). \tag{6.2.5}$$

Sometimes called "the index of the point w with respect to the curve γ", this concept is needed in the enunciation of Cauchy's residue theorem. As for the frequency of a complex exponential, it is its winding number with respect to the origin of a segment of its graph described in a unit of time.

$$\nu(\chi_{\nu}) = n(\chi,0) = \frac{1}{2\pi j}\int_{\chi}\frac{dz}{z} \tag{6.2.6}$$

But, since $dz/z = d(\ln z)$ and $\ln\chi = \Psi_{\chi}$,

$$\nu(\chi_{\nu}) = \frac{1}{2\pi j}\int_{\chi} d(\Psi_{\chi}), \tag{6.2.7}$$

meaning that the frequency of a complex exponential is the number of times the phase function spans its range in a unit of time.

To conclude, the analysis carried out above has shown that the preponderence of the set of complex exponentials among all other complete orthonormal sets of functions used for representing complex-valued functions stems from the fact that it is the set of characters of the group of real numbers with arithmetic addition. Consequently, the complex exponentials are also the (bounded) eigenfunctions of linear time-invariant systems described by either convolution or linear differential equations. It is

very important to notice that, as characters, the complex exponentials have identical (unit) modulus, such that their speed of variation may be uniquely determined in terms of variation of their phase.

As for the concept of frequency, it exhibits the following principal features:

a) the imaginary angular frequency $\lambda = j\omega = j2\pi\nu$ is the eigenvalue corresponding to the complex exponential as eigenfunction of the classical differentiator (harmonic);

b) the imaginary angular frequency is equal to the local frequency (defined as the slope of the phase) which is a characteristic constant (independent of time) of each complex exponential;

c) the frequency ν is the winding number with respect to the origin of the complex plane of the segment of graph which the complex exponential describes in unit time; i.e. it is equal to the number of times the phase function covers its range in unit time. Since the modulus of a character is always 1, this property appears as a generalization of the concept of frequency as the inverse of the period, to be used, if possible, even in case that the characters are not periodic; e.g. in the case of Walsh functions.

6.3 A CONCEPT OF GENERALIZED FREQUENCY

Following the analysis of the classical concept of frequency, it is now possible to extrapolate its essential features and define them as the features characterizing the concept of generalized frequency. Such a

concept is general only in the sense that it applies to all (and only to) the spaces of functions having a LCA group as domain and a field as co-domain. For each such function space, this concept is instrumental in ordering, in a sense of increased speed of variation, the set of characters spanning the space. The ordering of the character set is possible only subsequent to the definition of an order on the domain of the function set.

Then, the generalized frequency of a character can be defined:

a) formally, as the eigenvalue corresponding to the character as eigenfunction of the harmonic differential operator;

b) locally, as the everywhere equal slope of the character's phase function; and

c) globally, as the "unit-time" winding number of the character's graph with respect to the origin of the co-domain.

The next two sections deal with the application of this concept to function spaces other than the space of complex-valued functions on the real line. A systematic approach consisting of the following three stages is adopted.

1) Given a space of functions defined on a LCA group G with values in a field K, determine the set of characters forming the basis for a Fourier transform. The discussion will be mainly concerned with functions on finite, discrete domains [66,251] because this is the case implementable through numeric analysis, and because its concepts are

easier to grasp. The definition and analysis can be fairly easily extended to infinite, continuous domains.

2) Define a harmonic differential operator such that the characters are its eigenfunctions. This has been done by Gibbs and others [73, 80,39], their work being discussed in Sections 2.5 and 5.3. Also, define the corresponding slope operator.

3) Define the generalized frequency according to the three criteria enunciated above, and show that they all lead to the same ordering of the characters.

6.4 GENERALIZED FREQUENCY AND ONE-DIMENSIONAL FINITE DFT's

The problem of defining the frequency of discrete complex exponentials has been discussed many times before. Two opposite approaches can be distinguished. One accepts the notion of "digital" frequency without any question; this is the approach generally adopted by applied scientists who regard discrete signals as samples of continuous ones, such that the concept of digital frequency is a natural extension of that in classical Fourier analysis [280,21], for example. Hence, the numerous instances of discussing (and proposing algorithms for) the reordering of DFT spectral coefficients according to increasing frequency, without giving much thought to the meaning of frequency for signals which are defined on a finite abelian group. In contrast, the other approach considers this problem from an abstract, mathematical point of view [282]. Its conclusion is that, since there is no unique ordering of finite abelian

groups, there is no point in even mentioning the word frequency. The properties of the transform are invariant to changes in the order of the domain (i.e. permutations of rows or columns in the transform matrix) such that, whereas it is normal to talk about a transform domain, it is not acceptable to give it a frequency interpretation.

These two divergent approaches can be reconciled once a fixed "time-order" and a principal N-th root of unity, ξ_N, are accepted. (The latter is necessary to define a unit of variation - the variation from the value 1 to the value ξ_N.) Then, the concept of generalized frequency defined in the preceding section is applicable and brings a meaningful order to the transform domain. The fixed time-order is normally supplied by the natural injection mapping $i_G : G \to Z_N$ (or P_N) discussed in Chapter V.

For the space of complex-valued functions on a finite abelian group, the complex exponentials of Section 6.2 are replaced by the characters (called sometimes discrete complex-exponentials)

$$\chi_\nu : Z_N \to K \qquad \chi_\nu = \xi^{\nu t}, \tag{6.4.1}$$

where $\xi_N = e^{j2\pi/N}$ is the principal N-th root of unity. The phase function $\Psi_{\chi_\nu}(t)$ is the exponent modulo N of ξ_N; i.e., symbolically,

$$\Psi_{\chi_\nu}(t) = \nu t = \log_{\xi_N}(\chi_\nu(t)), \tag{6.4.2}$$

In general, any function $f : Z_N \to K$ can be expressed in polar form as

$$f(t) = A_f(t)\xi_N^{\Psi_f(t)}$$

$$A_f : Z_N \to R_+ \triangleq [0,\infty) \quad \text{and} \quad \Psi_f : Z_N \to (-N/2, N/2] \tag{6.4.3}$$

As for the concept of frequency, it is straightforward to show that the parameter ν characterizing the discrete complex exponentials meets criteria (a), (b) and (c) of Section 6.3 for defining generalized frequency. Indeed,

$$\mathcal{D}\, \xi_N^{\nu t} = \nu\, \xi_N^{\nu t} \quad \text{(by definition of } \mathcal{D}); \tag{6.4.4a}$$

$$\Delta_N \Psi_{\chi_\nu}(t) = \Delta_N(\nu t) = ((t \dot{+} 1)\nu - t\nu)_{\bmod N} = \nu \quad \text{(independent of } t); \tag{6.4.4b}$$

$$\frac{1}{N} \sum_{t \in Z_N} \Delta_N(\log_{\xi_N} \chi_\nu(1)) = \frac{1}{N} \sum_{t \in Z_N} \Delta_N \Psi_{\chi_\nu}(t) = \nu. \tag{6.4.4c}$$

Δ_N is the relative variation (or slope) "differential" operator - it serves to measure the variation of the function for the smallest linear displacement of its argument. For a finite (discrete) abelian group domain, the smallest displacement is a unit translation: $t \to t \dot{+} 1$.

A similar analysis can be carried out for the space of functions defined on Z_N and with values in a field isomorphic to Z_F, with addition and multiplication modulo F (since there cannot be any confusion, these operations will be denoted by the same symbols as those used for real numbers). The characters are now

$$\chi_\nu : Z_N \to Z_F; \quad \chi_\nu(t) = \zeta^{\nu t} = \zeta^{\Psi(t)}; \quad \zeta^N = 1 (\bmod F) \quad \nu \in Z_N \tag{6.4.5}$$

and the phase function $\Psi(t)$ is also to be taken modulo N. As shown in Chapter V, the expressions for the harmonic differential and the slope operators are formally the same as for a space of complex-valued functions on Z_N, such that, as above, it may be concluded that the parameter ν belonging to the dual group, which multiplies the variable t to produce the character's phase function (defined as a certain logarithm of the character) fulfills all the requirements for being called the generalized frequency of that character.

$$\mathcal{D}\xi^{\nu t} = \nu\xi^{\nu t} \quad \text{(by definition of } \mathcal{D}\text{)} \tag{6.4.6a}$$

$$\Delta_N{}^{\Psi}\chi_\nu(t) = \Delta_N(\nu t) = ((t\dot{+}1)\nu - t\nu)_{\text{mod } N} = \nu \tag{6.4.6b}$$

$$\frac{1}{N}\sum_{Z_N}\Delta_N(\log_\xi \chi_\nu(t)) = \frac{1}{N}\sum_{Z_N}\Delta_N{}^{\Psi}\chi_\nu(t) = \nu \tag{6.4.6c}$$

In fact, for one dimensional Fourier transforms, the characters are periodic functions such that they could have been ordered according to the inverse of their period, with no need for a conceptual discussion and definition of frequency. It has been introduced here first, to show that this inverse of the period of a character has all the essential features of the concept of frequency, and, secondly, to pave the way for the more difficult discussion of Walsh functions and their ordering.

6.5 SEQUENCY AND GENERALIZED FREQUENCY

The ordering of Walsh functions (especially the existence of many possible orderings and their inter-relationships) has been the subject of

many investigations ever since their introduction by Walsh [49,285,75,52,51, 284, etc.]. It is one of the main points of this dissertation to stress the fact that, because of their "multi-dimensionality", there can not be any unique ordering of Walsh functions. Their domain is a LCA group whose elements (even in the case of non-finite groups) are all of (same) finite order q, such that the domain can be viewed as a q-dimensional space, the points of which can not be uniquely ordered. The fact that this domain can be mapped on a subset of the real line has no bearing upon the essential fact that the set of characters of this LCA group admits many equally significant order criteria, none of which fulfills all the requirements for being a generalized frequency in the sense of Section 6.3.

Of the various Walsh functions orderings proposed and discussed, the one most related to the concept of frequency is that of ordering the cal and sal subsets of Walsh functions according to their number of sign-changes (zero-crossings) per period or unit "time interval". This criterion of ordering the Walsh functions, together with its resemblance with the ordering of cosine and sine functions according to their number of zero-crossings per unit interval, lie at the foundation of the concept of sequency. This concept has played a catalytic role in the development of Walsh analysis and the proliferation of its numerical applications ever since its definition by Harmuth in 1968. It was also the subject of controversy, most of it related to claims that this concept is a generalization of the concept of frequency.

Most of the objections raised were concerned not with its "evident" relationship with the concept of frequency, but with the fact that the concept of sequency, as defined by Harmuth, was not general enough to be applicable even to q-adic Walsh functions. Hence, the numerous attempts at generalizing the concept of sequency itself [94,276,275,274]. The main point raised was that the relationship sequency - frequency should not be based on an analogy cal/sal - cosine/sine functions in the space of real-valued functions, but on an analogy Walsh function - complex exponentials in the space of complex-valued functions, so that the criterion of ordering should not be the number of sign-changes but the variation of phase. This is, of course, correct, but the formulae provided for the computation of generalized sequency are of little use, either because they are incomplete [275,276] (in the sense of considering the phase variation along an unclosed segment of the character's graph; a result of disregarding the phase jump from the point N-1 to the point N), or because the formulae provided had been forcefully contrived to permit the recovery of the character's index as its generalized sequency (see [274], in which the choice of basis vectors and certain multiplicative factors is rather arbitrary).

A different approach to the generalization of sequency has been adopted by Gibbs [79], who, instead of considering the total variation of the phase of a Walsh function, considered its local variation, defined as its harmonic derivative. Correctly, it should have been the "slope" derivative; the fact that no major difficulty arose for dyadic Walsh

functions can be explained by the proportionality which exists between the two kinds of differential operators in (but only in) dyadic analysis. Apart from this aspect, the generalization had not been conclusive because it led to (at least) two equally valid orderings in the sense of increased local sequency.

Neither of these attempts at providing generalizations of the concepts of sequency and frequency have been successful because, as mentioned before,

i) they were not based on a systematic study of the concept of frequency, and

ii) they disregarded the multi-dimensional character of the Walsh transform.

The fact that sequency is not a generalized frequency can be proved by showing that although it fulfills the requirements (a) and (b) of Section 6.3 (as any other ordering of Walsh functions can be made to do!), it does not fulfill the last requirement.

Indeed, in q-adic Walsh analysis, any complex-valued function on $\mathbb{Z}_q^n$ can be represented by a polar expression

$$f(t) = A_f(t)\, \xi_q^{\Psi_f(t)} \tag{6.5.1}$$

where ξ_q is the principal q-th root of 1, and the phase function Ψ_f takes values in the interval $[0,q)$. The Walsh functions have unit amplitude, their phase being expressed in terms of the product between the modules, $\underline{\nu} = (\nu_{n-1}, \ldots, \nu_0)$ and $\underline{t} = (t_{n-1}, \ldots, t_0)$:

$$\Psi_{wal_{\underline{\nu}}}(t) = (\underline{\nu}\ ,\ \underline{t}) = \sum_{i=0}^{n-1} \nu_i t_i \in \mathbb{P}_q \qquad (6.5.2)$$

As in the case of classical vector analysis, the multi-dimensionality of $\mathbb{P}_q^n$ permits the introduction of a module differential operator $\underline{\mathcal{D}} = (\mathcal{D}_0, \ldots, \mathcal{D}_{n-1})$ to describe the "behaviour" of a function along each of the basis directions. The characters are the eigenfunctions of $\mathcal{D}$ with module eigenvalues $\underline{\nu}$:

$$\underline{\mathcal{D}}\chi_{\underline{\nu}} = \underline{\nu}\chi_{\underline{\nu}} = (\nu_{n-1}, \ldots, \nu_0)\chi_{\underline{\nu}} \qquad (6.5.3)$$

The eigenvalue $\underline{\nu}$ is also the "module-slope" (gradient) of the phase of the character

$$\underline{\nu} = \underline{\Delta}^{\Psi}\chi_{\underline{\nu}} \qquad (6.5.4)$$

The existence of non-trivially distinct bases of $\mathbb{P}_q^n$ implies the existence of more than one kind of module harmonic differentiator and slope operator, each kind having a different (according to ordering) species of Walsh functions as eigenfunctions.

The two equations above may be put into a scalar (one-dimensional) form by using the invertible norm which effects the bijective mapping between $\mathbb{P}_q^n$ and $\mathbb{P}_{q^n}$. For instance, denoting

$$(\mathcal{D}f)(t) \triangleq \sum_{i=0}^{n-1} q^{n-i-1}(\mathcal{D}_i f)(t) \qquad (6.5.5a)$$

$$(\Delta f)(t) \triangleq \sum_{i=0}^{n-1} q^{n-i-1} (\Delta_i f)(t), \qquad (6.5.5b)$$

it is possible to conclude that the index $\nu = \sum_{i=0}^{n-1} q^{n-i-1} \nu_i$ fulfills the requirements (a) and (b) of Section 6.3 for the Walsh-Kaczmarz species of Walsh functions, and could be called the generalized frequency of the corresponding character.

But this would be a rash conclusion, because the sequency does not satisfy the third requirement, the one concerning the total variation of the character's phase along its closed graph. If it were true that sequency is a generalized frequency, then one should be able to prove that

$$E_1 \triangleq \frac{1}{q} \sum_{t \in P_N} \Delta \Psi_{wal_\nu}(t) = \frac{1}{q} \sum_{t=0}^{n-1} [\Psi_{wal_\nu}(t+1) \ominus \Psi_{wal_\nu}(t)] = \nu \qquad (6.5.6)$$

where, as usual, $N = q^n$, $\ominus$ represents addition modulo q, and t and ν are non-negative integers less than N being equal to the norms of the respective n-tuples $\underline{t}$ and $\underline{\nu}$, $(\underline{t}) = \sum_{i=0}^{n-1} q^i (\underline{t})_i$ If the point N is identified with the origin 0 as it is customary for finite one-dimensional domains (they are finite abelian groups), this conjecture turns out to be false. What one can prove, given the Walsh-Kaczmarz phase function

$$\Psi_{wal_K}(t) = \sum_{i=0}^{n-1} (\nu_{n-i-1} \ominus \nu_{n-i}) t_i, \qquad (6.5.7)$$

is that [33]

$$E_2 \triangleq \frac{1}{q-1} \sum_{t=0}^{N-2} [{}_{\Psi}wal_{\nu}(t+1) \overset{\circ}{\Psi} {}_{\Psi}wal_{\nu}(t)] = \nu \qquad (6.5.8)$$

The proof makes use of the fact that as t spans the set $\{0, 1, \ldots, N-2\}$, the n-tuple $\underline{y} = \underline{(t+1)} \overset{\circ}{-} \underline{t} = (\ldots, (t+1)_i \overset{\circ}{-} t_i, \ldots)$ takes the form $(0, \ldots, 0, 1, 1, \ldots, 1)$ with i 1's at the right in $(q-1)q^{n-i}$ occasions. Then,

$$E_2 \triangleq \frac{1}{q-1} \sum_{t=0}^{N-2} [{}_{\Psi}wal_{K}(t+1) \overset{\circ}{\Psi} {}_{\Psi}wal_{K}(t)]$$

$$= \frac{1}{q-1} \sum_{t=0}^{N-2} [\sum_{i=0}^{n-1} (\nu_{n-i-1} \overset{\circ}{\nu} \nu_{n-i})(t+1)_i \overset{\circ}{\sum} \sum_{i=0}^{n-1} (\nu_{n-i-1} \overset{\circ}{\nu} \nu_{n-i}) t_i]$$

$$= \frac{1}{q-1} \sum_{t=0}^{N-2} \sum_{i=0}^{n-1} [\nu_{n-i-1} \overset{\circ}{\nu} \nu_{n-i}][(t+1)_i \overset{\circ}{-} t_i]$$

$$= \frac{1}{q-1} \sum_{t=0}^{N-2} \sum_{i=0}^{n-1} (\nu_{n-i-1} \overset{\circ}{\nu} \nu_{n-i}) y_i \qquad (6.5.9)$$

$$= \frac{1}{q-1} [\ldots + (\nu_{n-1} \overset{\circ}{\nu} \nu_n) \overset{\circ}{+} \ldots \overset{\circ}{+} (\nu_{n-i} \overset{\circ}{\nu} \nu_{n-i+1})(q-1)q^{n-i} + \ldots]$$

$$= \frac{1}{q-1} \sum_{i=1}^{n} \nu_{n-i}(q-1)q^{n-i} = \sum_{i=0}^{n-1} \nu_i q^i = \nu.$$

Then, instead of having $E_1 = \nu$, as it should be if ν were a generalized frequency, one has that

$$E_1 \triangleq \frac{1}{q} [(q-1)E_2 + {}_{\Psi}wal_{\nu}(N) \overset{\circ}{\Psi} {}_{\Psi}wal_{\nu}(N-1)] = \frac{1}{q} [(q-1)E_2 + \nu_0]. \qquad (6.5.10)$$

This means that $E_1 = \nu$ only provided that one accepts the convention that the phase jump from the point $N-1$ to the point N is not ν_0 but $\nu = \sum_{i=0}^{n-1} \nu_i q^i$. For one-dimensional domains ν_0 is ν, and hence, $E_1 = \nu$ as it should be. For multi-dimensional domains and non-trivial Walsh functions, taking the final phase jump as being equal to ν would be an unjustified artifice because it leads to the admission of phase jumps greater than q.

The above considerations can be easily extended to Walsh-type functions with values in finite fields. At the same time, an extension to the "inhomogeneous" space $\mathbb{F}_{q_o}^{n_o} \oplus \mathbb{F}_{q_1}^{n_i} \oplus \ldots \oplus \mathbb{F}_{q_{ij}}^{n_i}$, for which a harmonic differentiator may be defined along the principles exposed in [72], involves some merely notational complications.

This, then, explains the limitations of the concept of sequency and the difficulties encountered in its generalization to q-adic Walsh functions. Of course, one can use the ad-hoc formula "generalized sequency $\nu = E_2$", but its significance is minor. The definition and acceptance of the concept of sequency as a generalized frequency was facilitated in dyadic analysis by the felicitous particularity of the case $q = 2$, when the (real-valued) characters exhibit that alternately even and odd symmetry, and the role played by the factor $(q-1)$ in Eqn. 6.5.8 is obscured by the fact that it is equal to 1. It is only surprising that nobody has questioned the use of the formula E_2 instead of E_1; i.e. the use of the unclosed graph of the character (E_2 ignores the

last phase jump to the point N) for measuring the total phase variation. Again, it should be stressed that the concept of sequency is a valid and a proven-to-be-useful tool in the context of dyadic Walsh analysis. The objections raised in this section were not directed against it per se, but against the claims that it represents a generalization of the concept of frequency.

After this analysis of the concepts of sequency and frequency, the investigation will be directed towards the study of spectral analysis techniques in function spaces no longer having as domain an additive group but a multiplicative one.

CHAPTER VII

MULTIPLICATIVE CONVOLUTION: FAST NUMERIC IMPLEMENTATION AND SPECTRAL ANALYSIS OF MULTI-COMPONENT SIGNALS

7.1 INTRODUCTION

This chapter investigates the use of digital signal processing techniques, specifically the discrete Fourier transform and digital filters, for performing fast convolutions and deconvolutions in the space of complex-valued functions having a multiplicative abelian group as domain. As mentioned in the introductory chapter, this investigation was prompted by the need to analyze a certain class of multi-component signals appearing in magnetic resonance experiments. Accordingly, the presentation will be less concerned with theoretical details, the accent being put on the practical solution of the problem at hand.

This problem - the analysis of multi-component signals - is not an easy one; in fact, it is one of the most time consuming, critical and frustrating stages in the collection and interpretation of experimental data. Multi-component signals carry information encoded in the nature (shape) of their components, their amplitudes, widths, locations and their number. Assuming linearity, the process of extracting this information is mathematically equivalent to that of representing the signal $x(t)$ in a functional space spanned by a set of basis functions $\{k(t,s)\}$.

$$x(t) = \int_S u(s)k(t,s)ds; \quad t \in T, \tag{7.1.1}$$

where T and S are, usually, intervals on the real line. Of course, in practice, this often takes the discrete form

$$x(n) = \sum_{m=1}^{M} u(m)k(m,n). \tag{7.1.2}$$

In principle, it is always possible to use an iterative curve fitting procedure, but such an approach is painstakingly labourious, and requires good a priori numerical information in order to guarantee convergence in an acceptably finite time [285,286]. In contrast, transform techniques, if they exist, do not require any a priori numerical information, are independent of the number of components in the signal, and can be made amenable to automation.

In many cases, the problem is simplified by assuming that the set of kernels $\{k(t,s)\}$ is generated by some operation performed on an unique basic waveform $k(t)$. The classic example is the case of $k(t)$-generation by continuous (additive) translation on the real line, when Eqn. 7.1.1 becomes a convolution representation:

$$x(t) = \int_S u(s)k(t-s)ds = u(t) * k(t); \quad t \in T. \tag{7.1.3}$$

Such a representation has many interpretations in practice. The filtering problem requires the computation of the integral (7.1.3) in order to find the output $x(t)$ of the TIL system with impulse response $k(t)$ to an excitation $u(t)$. The spectral analysis problem, on the other hand, requires

to solve (7.1.3) as an integral equation in order to compute the density function $u(t)$ describing the distribution (in "time") of the kernels $\{k(t)\}$ composing the given signal $x(t)$. Alternatively, the system identification problem requires the identification of the system function k from input (u) - output (x) data. For $T=S=R=(-\infty,\infty)$, the solution to these problems can be readily computed by applying Fourier transform techniques,

$$x = F^{-1}\{F\{u\}\cdot F\{k\}\} \qquad u = F^{-1}\{F\{x\}/F\{k\}\}, \tag{7.1.4}$$

which have become "numerically" attractive since the introduction of FFT algorithms.

This chapter is concerned with problems admitting a representation of the form (7.1.1) in which the basis functions are also generated from a unique kernel, but by a multiplicative translation process:

$$x(t) = \int_S u(s)k(ts)ds; \qquad t \in T. \tag{7.1.5}$$

Equation 7.1.5 can be viewed as the representation of the signal $x(t)$ by a linear superposition of components having the same nature and location but different widths and amplitudes. Such signals appear frequently in applied sciences, where basic components may take such varied forms as sinusoids, exponentials, gaussians, lorentzians, sin(t)/t - type curves, etc... Characteristic is the fact that the kernel k depends on the product of the arguments. The most obvious and frequently used representation of this kind is the Fourier transform, for which $k(t,s) = \exp(j2\pi ts)$.

A numeric Fourier transform has ceased to be a problem since the advent of the FFT [15,18]. However, when k is not a complex exponential or a simusoid, the numerical computation of the transform and, especially, the numerical inversion of (7.1.5) is a difficult problem. For example, considerable work has been done on the numerical inversion of the Laplace transform on the real line, for which the basic waveform is an exponential decay [287,288]. Other well known transforms of the form (7.1.5) are the Hankel transforms [289,290].

$$x(t) = \int_0^\infty u(s)(st)^{1/2} J_a(st)ds; \quad t \in (0,\infty) \tag{7.1.6}$$

where J_a is the a-th order Bessel function.

The representation (7.1.5) is sometimes called a Mellin convolution transform because, in certain conditions, it admits inversion by a transform of the same type.

$$u(s) = \int_T x(t)h(ts)dt; \quad s \in S; \quad T, S = (0,\infty) = R_+, \tag{7.1.7}$$

having a kernel $h(s,t)$ related to the kernel $k(t,s)$ via their Mellin transforms [289,290].

$$M\{k;v\}M\{h;1-v\} = 1; \tag{7.1.8}$$

where the Mellin transform is defined as

$$M\{x;v\} = \int_{R_+} x(t)t^{v-1}dt. \tag{7.1.9}$$

Widder [291] presents an elegant solution to the Mellin convolution integral equation in terms of an operator E which is related to the bilateral Laplace transform of the kernel k. Besides its theoretical interest, Widder's approach leads directly to the inversion of (7.1.5) via a power series representation of x(t). Finite Mellin convolution transforms (S , T are finite) as well as bilateral Mellin convolution transforms (S, T cover the entire real line) have been studied by Perry [292, 293], while the generalized transform has been studied by Zemanian [10].

All of these studies are purely analytic, neither of them being concerned with the problem of obtaining efficient algorithms for performing either the convolution of (7.1.5) or its deconvolution with functions described by sets of data. From a numeric point of view, a more appealing approach would be to relate the solution of (7.1.5) to the Fourier transform and its fast and relatively accurate numerical implementation. This is feasible by mapping the multiplicative convolution representation (7.1:5) into the "normal" convolution equation (7.1.3). This is the principle of the spectral analysis technique to be discussed later in the chapter. The idea of mapping the Mellin convolution into a normal one can be traced back to the work of Titchmarsh [30]. Gardner et al [294] were the first to attempt its use, in a laboratory situation, for the analysis of multi-component exponential decays, but they met with limited success due to the difficulty of performing a numerical Fourier transform at that time (1956-1959). Their method has been revived by Schlessinger [295] and, independently, by Smith and Cohn-Sfetcu [296] who made use of

the FFT algorithm. The latter have also showed the necessity of coupling the transformation with an adequate low-pass filtering to decrease the effect of computational noise and of high frequency noise usually present in experimental data. Cohn-Sfetcu et al [297,31] have shown that the same technique can be generalized and used in connection with basis functions other than exponential decays. Following their work, Siegman [39] has devised a fast algorithm for the computation of Hankel transforms, which are particularly useful in optical beam propagation problems. A similar technique for performing discrete Mellin transforms has been used independently in an attempt to recover images affected by coma-type aberrations [298].

Section 7.2 presents the theory on which the fast numeric transform and spectral analysis algorithms are based, while Section 7.3 is concerned with problems affecting an efficient implementation of these algorithms. A practical example is detailed in Section 74. Conclusions and ideas for possible improvements and research into this topic are discussed in Section 7.5.

7.2 FOURIER TRANSFORMS ON MULTIPLICATIVE ABELIAN GROUPS: AN ALGORITHM FOR SOLVING MELLIN CONVOLUTIONS

The basic problem for which this investigation has been started is that of devising a numerical algorithm for computing the spectrum $u(s)$ describing a signal $x(t)$ as a superposition of components having the same shape and location but different amplitudes and widths

$$x(t) = \int_{R_+} u(s)k(ts)ds; \quad t \in R_+ \tag{7.2.1}$$

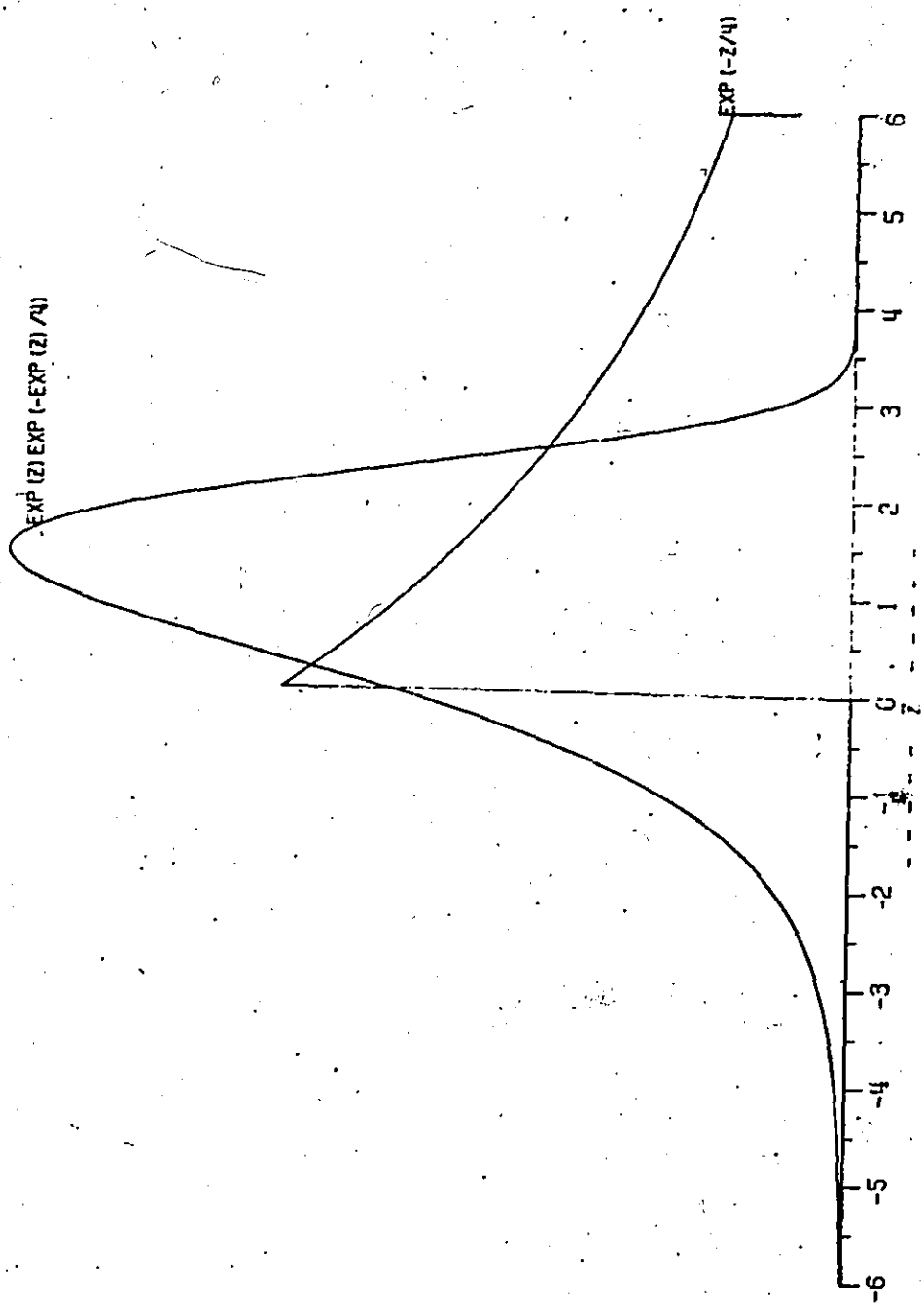


FIG. 7.1 THE EXPONENTIAL CHANGE OF VARIABLES

where R_+ is the positive real line. The solution is sought in terms of mapping this multiplicative convolution into a normal one.

A comparison of the arguments of the two kernels in (7.2.1) and (7.1.3) reveals that this can be achieved via a multiplication of both sides of (7.2.1) by t and a subsequent transformation of variables. The required transformation is the isomorphic mapping of the multiplicative group of positive real numbers $\dot{R}_+$ to the additive group of real numbers R, which is performed by the logarithmic function:

$$\alpha = \log_r t \qquad \beta = -\log_r s. \tag{7.2.2}$$

Then, (7.2.1) transforms into

$$\tilde{x}(\alpha) = \tilde{u}(\alpha) * \tilde{k}(\alpha) = \int_R \tilde{u}(\beta)\tilde{k}(\alpha-\beta)d\beta; \quad \alpha \in R, \tag{7.2.3}$$

where

$$\tilde{x}(\alpha) = r^{\alpha}x(r^{\alpha}); \quad \tilde{u}(\alpha) = u(r^{-\alpha}); \quad \tilde{k}(\alpha) = \ell n(r)r^{\alpha}k(r^{\alpha}). \tag{7.2.4}$$

Therefore, given the signal x and the basic waveform k, the spectrum u of the representation (7.2.1) can be obtained in the form

$$u(r^{-\alpha}) = \tilde{u}(\alpha) = F^{-1}\{F\{\tilde{x}/F\tilde{k}\}\}; \tag{7.2.5}$$

the functional relationship between $\tilde{u}(\alpha)$ and α being equivalent to that between $u(s)/s$ and $\log_r(1/s)$. Figure 7.1 indicates the effect which the transformation of variables has upon an exponential function, and Figure 7.2 illustrates a 6-component Gaussian signal and its spectrum obtained via this technique. There is normally no disadvantage in using

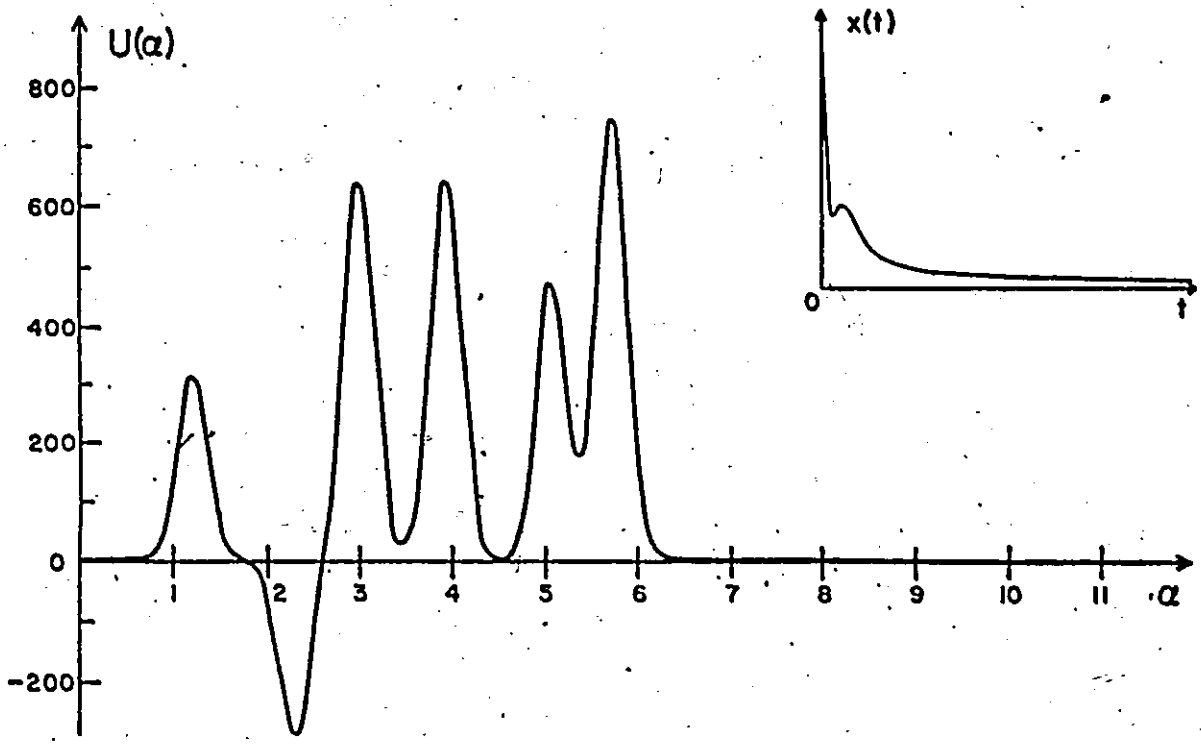


Figure 7.2. The spectrum $U(\alpha)$ of a superposition of six Gaussian functions of the form $A_i \exp(-t^2/T_i^2)$,

$$x(t) = 100\exp(-(t/3.3)^2) - 30\exp(-(t/10)^2) + 32\exp(-(t/20)^2) + 13\exp(-(t/50)^2) + 3\exp(-(t/160)^2) + 2.5\exp(-(t/300)^2).$$

$n = e$, $n = 1$, $N = 256$, $A_0 = 12.8$, triple time aliasing from $(-3A_0, 3A_0]$ into $(-A_0, A_0]$, $\nu_g = 35$. The finite widths of the peaks reflect the effect of the low-pass filtering. The amplitude and width of the i-th Gaussian component are given by $A_i = U(\alpha_i)/\exp(\alpha_i)$ and $T_i = \exp(\alpha_i)$ respectively, where α_i is the position of the i-th spectral peak. The errors in determining the amplitudes and widths were less than ±1.2% and ±0.8% respectively.

natural logarithms (r=e) and all the examples in this chapter have been computed with this choice of r.

A condition for the existence of the solution (7.2.5) is that the functions tx(t) and tk(t) belong to $L(0,\infty)$; i.e., they are integrable. If this is not the case, it is still possible to use this method provided that $t^{1/n}x(t)$ and $t^{1/n}k(t)$, where n>1, belong to $L(0,\infty)$. In such cases, the change of variables should be

$$\alpha = \log_r t \qquad \beta = -n \log_r s. \tag{7.2.6}$$

The relationship between $\tilde{u}(\alpha)$ and α is now equivalent to that between $u(s)/s^{1/n}$ and $\log_r(1/s)$. Figure 7.3 presents the spectrum obtained by analyzing a 5-component sin(t)/t signal using such a change of variables (n=2). Also, it is obvious that this approach to solving (7.2.1) can be generalized to kernels of the form

$$k(t,s) = k(t^m s^n); \qquad n, m \in R_+, \tag{7.2.7}$$

with the corresponding change of variables

$$\alpha = m \log_r(t); \qquad \beta = -n \log_r(s).$$

If there is interest not in the inversion of (7.2.1) but in obtaining a fast multiplicative convolution transform, then the solution is immediately obtained in the form

$$x(r^\alpha) = r^{-\alpha} F^{-1}\{F\{\tilde{u}\}\cdot F\{\tilde{k}\}\}. \tag{7.2.8}$$

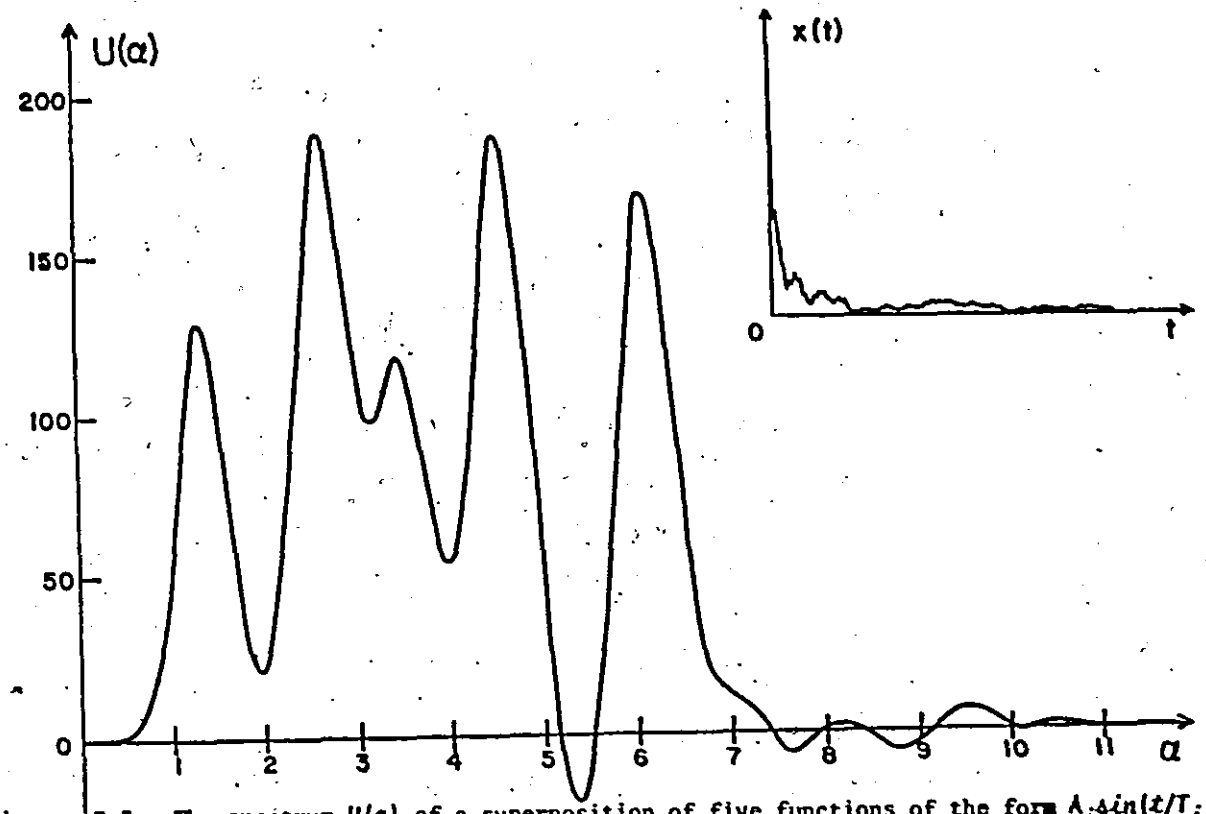


Figure 7.3. The spectrum $U(\alpha)$ of a superposition of five functions of the form $A_i \sin(t/T_i)/(t/T_i)$,

$$x(t) = 65\,\frac{\sin(t/4)}{t/4} + 45\,\frac{\sin(t/15.4)}{t/15.4} + 19\,\frac{\sin(t/35.7)}{t/35.7} + 17\,\frac{\sin(t/100)}{t/100} + 7\,\frac{\sin(t/500)}{t/500}\,.$$

$r = e$, $n = 2$, $N = 256$, $A_0 = 12.8$, double aliasing from $(-2A_0, 2A_0]$ into $(-A_0, A_0]$, $v_g = 20$. The finite widths of the peaks reflect the effect of the low-pass filtering. The amplitude and width of the i-th $\sin(x)/x$ component are given by $A_i = U(a_i)/\exp(a_i/2)$ and $T_i = \exp(a_i)$ respectively, where a_i is the position of the i-th spectral peak. The errors in determining the amplitudes and widths were less than ±3.5% and ±2% respectively.

As mentioned earlier, the computational advantages of this method lie in the possibility of implementing either (7.2.5) or (7.2.8) by using the DFT and its associated FFT algorithm. Before discussing the problems involved with a practical implementation of these solutions, it may be interesting to sketch a different and more general approach to the solutions (7.2.5) and (7.2.8). This can be done via the general theory of Fourier transforms on groups. The benefit is doublefold. First, it embeds this chapter's investigation into the avowed subject of the dissertation: topics on the theory and applications of generalized convolution transforms in digital signal processing and system theory. Secondly, and most important, it permits a generalization to domains other than the positive real line, a fact which is essential for developing a consistent theory of finite and discrete Mellin convolution transforms.

The function space of interest is the space of complex-valued functions on a multiplicative LCA group $\dot{G}$. More generally, the functions may take values in any field K. The domain of greatest interest is the multiplicative group of positive real numbers $\dot{R}_+$. (The dot above "underlines" the multiplicative and not additive nature of the group operation.) Right and left multiplicative translation operators are defined as particular cases of general, abstract translation operators on LCA groups (see Eqn. 1.1.1).

$$\dot{T}_{\rightarrow}^{\tau}x(t) \triangleq x(t/\tau); \quad \dot{T}_{\leftarrow}^{\tau}x(t) \triangleq x(\tau/t);$$
$$\tau \in \dot{G}; \quad \forall\, x \in K[\dot{G}]. \tag{7.2.9}$$

A multiplicative convolution can be viewed as a generalization of the process of taking weighted linear combinations of a function's multiplicative translates. Formally, a multiplicative convolution (denoted m-convolution) is the mapping $\dot{*} : K[\dot{G}] \times K[\dot{G}] \to K[\dot{G}]$ defined by

$$(f \dot{*} g)(\tau) \triangleq \int_{\dot{G}} \dot{T}_{+}^{\tau} f(t)\, g(t)\, \dot{dt}. \tag{7.2.10}$$

In contrast with Eqn. 1.1.2, Eqn. 7.2.10 uses the symbol $\dot{dt}$. This is done to stress the difference between the measure on the multiplicative group $\dot{R}_{+}$ and the measure on the additive group R. $\dot{dt}$ refers to the Haar measure on $\dot{G}$, and for the particular case of $\dot{R}_{+}$, the relationship between $\dot{dt}$ and dt on R is $\dot{dt} = dt/t$, such that $\dot{d}(t_1 t_2) = \dot{d}(t_1) + \dot{d}(t_2)$ as required, and (7.2.10) takes the Lesbegue form expression

$$(f \dot{*} g)(\tau) \triangleq \int_{0}^{\infty} f\left(\frac{\tau}{t}\right) g(t)\, \frac{1}{t}\, dt. \tag{7.2.11}$$

As any convolution on a LCA group, the m-convolution is commutative, associative and distributive with respect to K-addition in L(G), which becomes a commutative Banach algebra if multiplication is defined by a m-convolution product.

This multiplicative convolution algebra can be mapped into a pointwise multiplication algebra via the corresponding multiplicative Fourier transform. By definition, the group's characters are the homomorphisms $\dot{\chi} : \dot{G} \to \mathbb{C}$ satisfying

$$|\dot{\chi}(\nu,t)| = 1: \quad \forall\, t \in \dot{G} \tag{7.2.12}$$

$$\dot{\chi}(\nu,t_1)\dot{\chi}(\nu,t_2) = \dot{\chi}(\nu,t_1 \cdot t_2); \quad \forall\, t_1,t_2 \in \dot{G}.$$

They form the dual group $\dot{\Gamma}$. For $\dot{G} = \dot{R}_+$, the dual group is also (isomorphic to) $\dot{R}_+$, and the role of characters may be played by the functions

$$\dot{\chi}(\nu,t) = \exp(j2\pi \ell n(t\nu)) = (\nu t)^{j2\pi}. \tag{7.2.13}$$

It is also evident that

$$\dot{\chi}(\nu_1 \cdot \nu_2,t) = \dot{\chi}(\nu_1,t) \cdot \dot{\chi}(\nu_2,t).$$

We define the multiplicative Fourier transform (m-FT) of any function $f \in L(\dot{G})$ as

$$\dot{F}\{f\} \triangleq \hat{\dot{f}}(\nu) = \int_{\dot{G}} f(t)\dot{\chi}(\nu,1/t)\dot{dt}, \tag{7.2.14}$$

with the corresponding inverse transform

$$\dot{F}^{-1}\{\hat{\dot{f}}\} \triangleq \int_{\dot{\Gamma}} \hat{\dot{f}}(\nu)\dot{\chi}(\nu,t)\dot{d\nu} = f(t). \tag{7.2.15}$$

The fact that the m-FT of the m-convolution is equal to the pointwise multiplication of the m-FT's of the respective functions can be viewed as a trivial particularization of theorem (1.1.10); or it can be proved following a similar proof-path as for the classical Fourier or Laplace transforms. Indeed, using the Lesbegue form for $\dot{G} = \dot{R}_+$, it is easy to see that, if

$$x(t) = u(t) \dot{*} k(t) = \int_{\dot{R}_+} u(s)k(t/s)(ds/s),$$

then

$$\dot{\hat{x}}(\nu) = \int_{\dot{R}_+} x(t)\dot{\chi}(\nu,\tfrac{1}{t})\,\frac{dt}{t} = \int_{\dot{R}_+} \dot{\chi}(\nu,\tfrac{1}{t})\,\frac{dt}{t} \int_{\dot{R}_+} u(s)k\left(\frac{t}{s}\right)\frac{ds}{s}$$

$$= \int_{\dot{R}_+} u(s)\dot{\chi}(\nu,\tfrac{1}{s})\,\frac{ds}{s} \int_{\dot{R}_+} k(\tfrac{t}{s})\dot{\chi}(\nu,\tfrac{s}{t})\,\frac{s}{t}\,d(\tfrac{t}{s})$$

$$= \dot{\hat{u}}(\nu)\cdot\dot{\hat{k}}(\nu). \qquad (7.2.16)$$

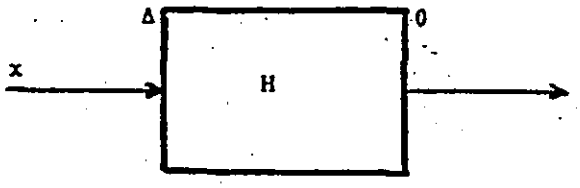
Such a presentation of the theory of multiplicative convolutions and Fourier transforms can be continued within the same pattern of particularizing to multiplicative LCA groups all the results of the general theory of Fourier transforms on abstract LCA groups, (so much about the power of going from the abstract and general to the concrete and particular case, in contrast with the purely analytic approach of proving separate theorems for all particular cases). We shall continue no longer although this approach may prove to be rich in new analytic results.

For the moment, our interest lies with the numerical solution of (7.2.1). It is evident that this equation can be rewritten as an m-convolution

$$x(t) = \int_{\dot{R}_+} u^{\dagger}(s)k^{\dagger}(\tfrac{t}{s})\,\frac{ds}{s}\,; \quad u^{\dagger}(s) = s\cdot u(s); \quad k^{\dagger}(\tfrac{t}{s}) = k(ts) \qquad (7.2.17)$$

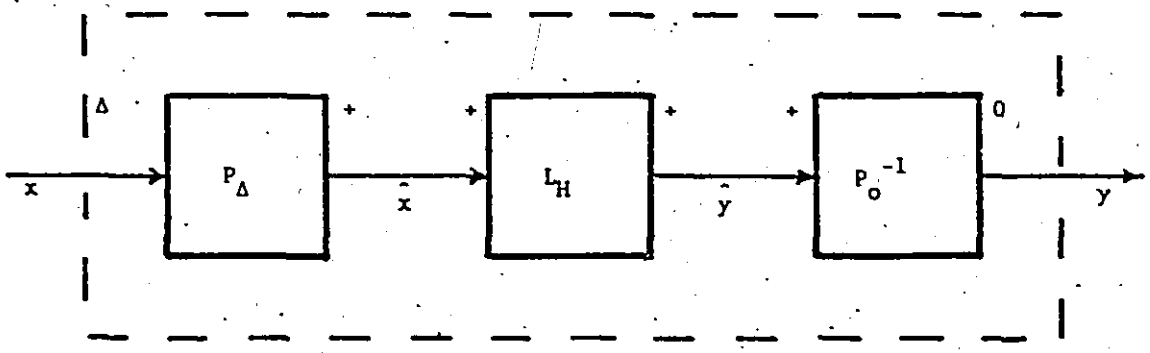
whose immediate solution can be written, following (7.2.16), as

$$u^{\dagger}(a) = \dot{F}^{-1}\{\dot{F}\{x\}/\dot{F}\{k^{\dagger}\}\}, \qquad (7.2.18)$$



$$H\{x_1 \Delta x_2\} = H\{x_1\} \; 0 \; H\{x_2\}$$

$$H\{c*x\} = c/H\{x\}$$



$$P_\Delta\{x_1 \Delta x_2\} = \hat{x}_1 + \hat{x}_2$$

$$P_\Delta\{c*x\} = c\hat{x}$$

$$P_o^{-1}\{\hat{y}_1 + \hat{y}_2\} = y_1 \; 0 \; y_2$$

$$P_o^{-1}\{c\hat{y}\} = c/y$$

Figure 7.4. Illustration of intermediate linear realizations of generalized superposition systems.

which is equivalent with (7.2.5). Similarly, it is possible to obtain the m-FT equivalent of (7.2.8).

The gain obtained by adopting this approach to the solution of (7.2.1) is, for the moment, mostly theoretical: the transformation involved in mapping (7.2.1) into (7.2.3) loses any traces of arbitrariness; and, more important, the theory covers more general spaces than $\mathcal{L}(\dot{R}_+)$, and is embedded in the context of Fourier transforms on groups.

As for the numerical implementation of the solution, although it is not difficult to provide a fast algorithm for the m-FT, it is more convenient to consider the solutions in the form (7.2.5) and (7.2.8), and thus make use of the vast knowledge available concerning the implementation of the Fourier transform via DFT and its fast algorithms.

But, before discussing in detail the numerical implementation, it is worth making another parenthesis and point to the relationship between the transform method described above and the generalized superposition theory of Oppenheim [299], which has found use in the non-linear filtering of multiplied and convolved signals [300]. Generalized superposition (homomorphic) systems (mappings) are described in Fig. 7.4 together with Oppenheim's idea of performing (or inverting) the mapping via an intermediate linear system (mapping). In order to obtain the linear image of the homomorphic system, he proposed the use of certain transformations P_Δ and P_o^{-1} affecting the co-domain of the working function space. The transform described here can be viewed in similar terms, with the difference

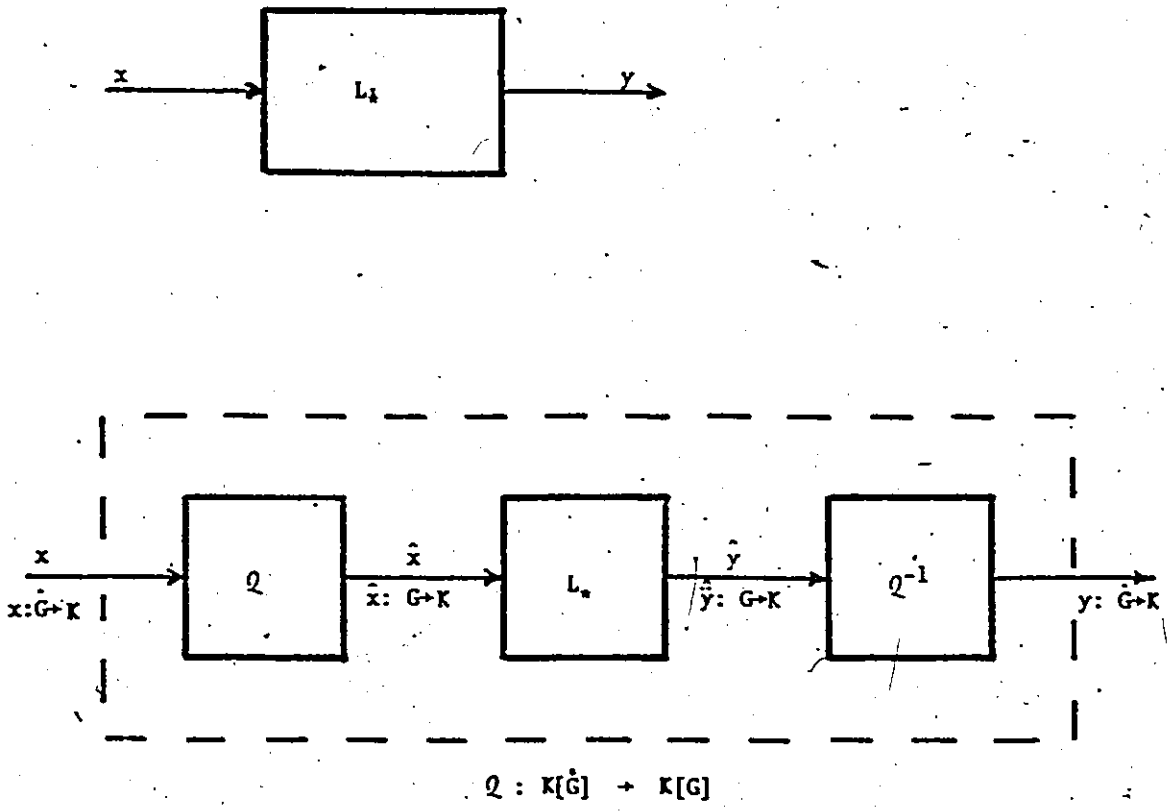


Figure 7.5. Illustration of a time-invariant realization of multiplicative-invariant linear systems.

that we are now concerned with the domain of the function space and not its co-domain. In contrast with the principle of generalized superposition we may talk now about a principle of generalized time-invariance (Figure 7.5 - remember q-TIL systems also), where $L_{\ddagger}$ signifies a convolution operator in a space of functions on an abstract LCA group $\{G,+\}$, whereas L_{+} signifies a convolution operator (TIL system) in the space of functions on an additive LCA group $\{G,+\}$, which usually, is the set of either additive reals or integer numbers.

7.3 NUMERICAL IMPLEMENTATION: GENERAL DISCUSSION

The numerical implementation of solution (7.2.5) involves, mainly, an exponential change of variables followed by the inversion of a (additive) convolution integral equation (i.e., a deconvolution) followed by another transform to obtain the final result u. The last transform is not absolutely necessary because the graph of $u(\alpha)$ versus α is equivalent to that of $u(s)/s$ versus $\ln(1/s)$.

The change of variables may not be necessary either, if data; i.e. the signal $x(t)$, is gathered at exponentially increasing intervals, and the kernel k is known analytically and is calculated by the computer. If, as is usual, the signal is sampled at equidistant intervals, then it is necessary to perform an interpolation which is by no means a simple procedure in this case. This is so because, as it will be seen, the deconvolution is extremely sensitive to the presence of noise (computational errors) in the data. The subject of correct interpolation is a problem in

itself, and will not be treated in depth here, although some remarks will be made in connection with the spectral analysis example to be presented in Section 7.4.

The problem of performing a numeric deconvolution or convolution is by far the one of greatest importance for a successful implementation of (7.2.5) and (7.2.8). For purely numeric data, it is possible to perform a deconvolution via a Monte Carlo approach [301] or by using de Marsilly's method as outlined in Chapter IV. If the analytic form of the kernel k is not known but it may be safely assumed that it is separated in frequency from the spectrum u (i.e. their Fourier transforms have non-intersecting supports), then a homomorphic deconvolution [300,19,22,302,303] may be attempted. If not, there remains to try a blind deconvolution [304]. But, usually, k is known analytically, and it makes much more sense to use straight DFT techniques in conjunction with the FFT algorithm.

The literature on the DFT and FFT is abundant [174,305,164] and the computational techniques have been thoroughly discussed and refined. Nevertheless, the implementation of a deconvolution process remains a delicate problem [29,306,307]. For low-frequency signals, as is usually the case, the deconvolution favours the high-frequency components of the data and, hence, of the noise accompanying the signal. What is worse, even if the original data are pure of any noise, the digital processing (including the FFT [308,19]) is bound to produce some noise due to inherent computational inaccuracies; and this noise is subsequently enhanced by the deconvolution process.

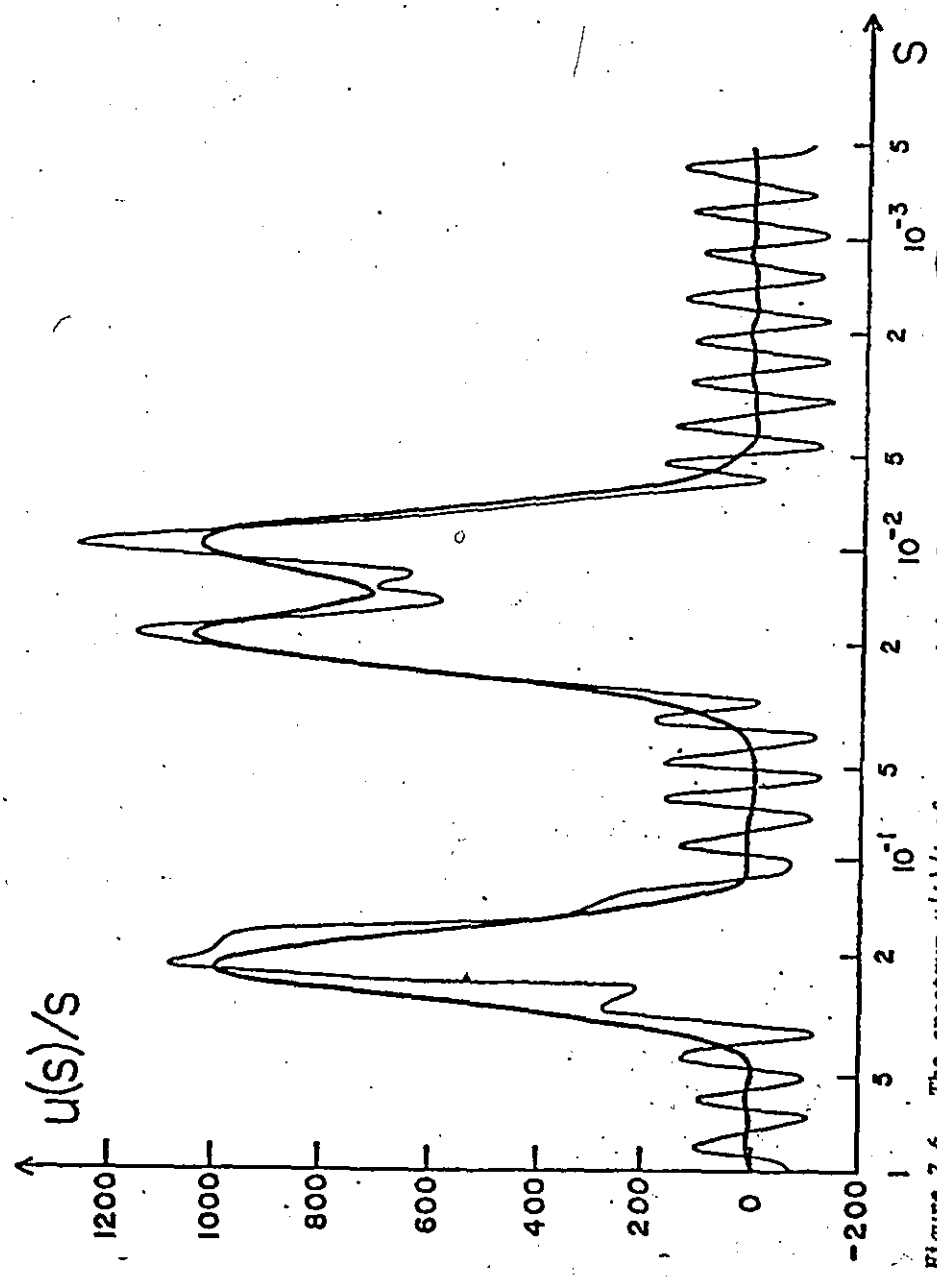


Figure 7.6. The spectrum $u(s)/s$ of a superposition of three Lorentzian functions, $x(t) = 20/(1+(t/5)^2) + 2/(1+(t/50)^2) + 1/(1+(t/100)^2)$,

a) Illustrating the effect of changing the Gaussian filter from $\nu_g = 24$ (heavy line) to $\nu_g = 28$ (light line). $\Delta t = 2$, $n = 1$, $N = 256$, $A_0 = 12.8$, and triple time aliasing.

To alleviate noise enhancement phenomena, the deconvolution must be coupled with a low-pass filtering performed in the final stages of the processing. Unfortunately, low-pass filtering has the effect of broadening the final spectrum; i.e., of diminishing the resolution. It is a fact of life that one always has to strike a compromise between noise purification and resolution enhancement. Computationally, this low-pass filtering poses no problems, since it can be implemented by a simple point-wise multiplication of $F\{x\}$ with the transfer function $\hat{h}(\nu)$ of the filter. The optimum filter to be used is dependent on the signal, the noise, and the criterion for optimality [25,26].

However, for signals composed of discrete peaks, good preliminary results may be normally obtained by employing a filter with a Gaussian transfer function

$$\hat{h}_g(\nu) = \exp(-\nu^2/\nu_g^2). \tag{7.3.1}$$

For separate pulses and white noise, this filter provides the best compromise between noise reduction and loss of resolution because it is characterized by the minimum duration-bandwidth product [11]. In practice, one would start with a hard filtering (ν_g small - low resolution but high S/N) and then go to lighter filterings (ν_g larger - high resolution but reduced S/N), selecting the compromise between resolution and S/N which best suits both his purpose and his cost function; i.e., the measure of the relative cost of not noticing small signal components versus cost of accepting false components. Figure 7.6a illustrates the effect of changing

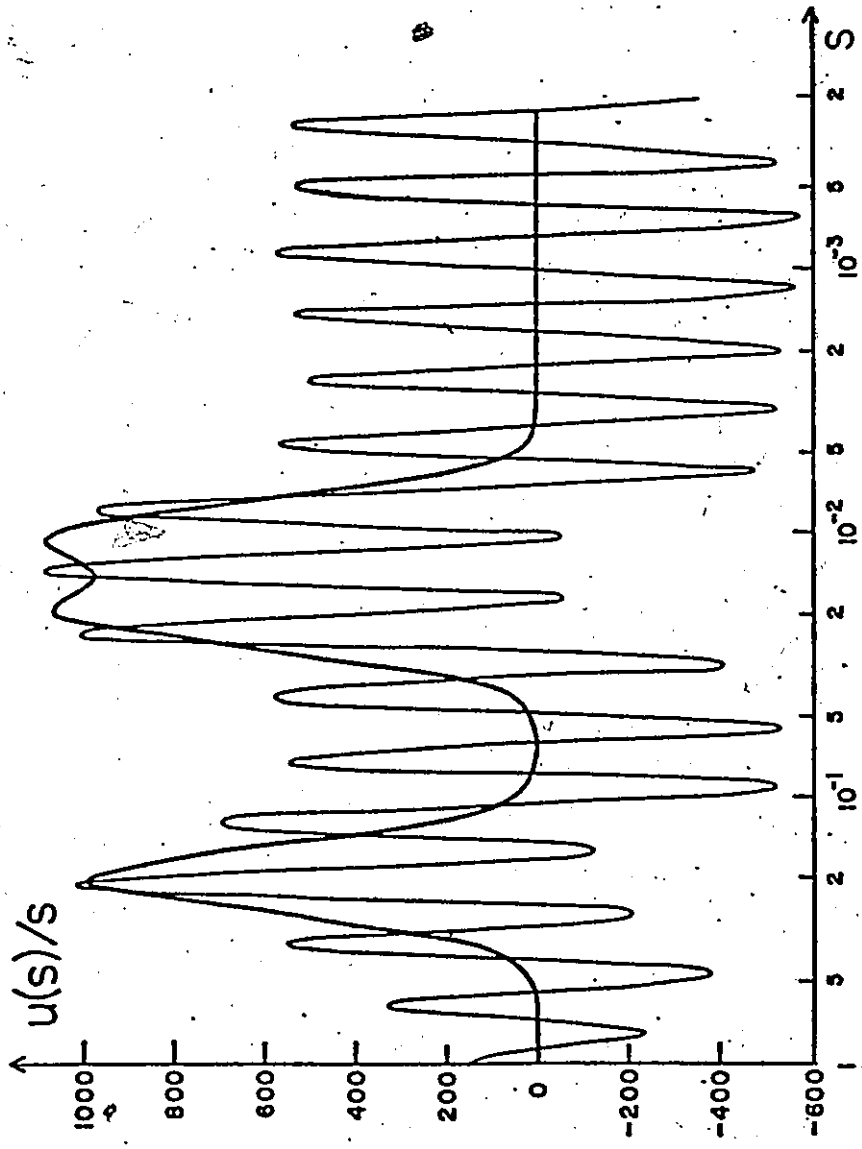


Figure 7.6. b) Illustrating the improvement in S/N due to triple time aliasing (heavy line) in comparison with no aliasing (light line). $\kappa = e$, $n_* = 1$, $N = 256$, $A_0 = 12.8$, $\nu_\theta = 20$.

the filter parameter ν_g when analyzing a 3-component Lorentzian signal.

The numerical implementation of (7.2.5) is also affected by data window effects and by the fact that the convolution associated with the DFT is cyclic rather than aperiodic, as required by (7.2.5). It is only natural to be so because instead of a FT on the entire real line we perform a DFT, which is a FT on a finite discrete group of order N. To avoid these problems, it is necessary to make sure that the functions $r^{\alpha}x(r^{\alpha})$ and $\ln(r)r^{\alpha}k(r^{\alpha})$ have, for all computational purposes, a finite support which is included in the α-interval $(-A_o,A_o]$ over which they (and the DFT) are computed. An increase of A_o beyond certain limits is detrimental since it leads to longer computation times and larger computational errors (N is proportional to A_o); instead, it is advisable to perform an aliasing (folding over) of the data inside the $(-A_o,A_o]$ interval with data values outside the interval [37]. As an illustration, Figure 7.6b presents the spectra of the 3-Lorentzian signal computed with and without such an aliasing.

As for the numerical implementation of (7.2.8), this is a much simpler problem since it involves a convolution rather than the inversion of one. Siegman [39] has discussed the problem of implementing (7.2.8) for the particular case of performing Hankel transforms:

$$g(\rho) = 2\pi \int_0^{\infty} r\, f(r) J_{\ell}(2\pi\rho r)\, dr \qquad (7.3.2a)$$

$$f(r) = 2\pi \int_0^\infty \rho g(\rho) J_\ell(2\pi\rho r)\, d\rho. \qquad (7.3.2b)$$

After the change of variables

$$r = r_o e^{a\alpha} \qquad \rho = \rho_o e^{a\beta}, \qquad (7.3.3)$$

where r_o, ρ_o, and a are scaling parameters; and a subsequent equidistant sampling in the α- and β-domains, Eqs. 7.3.2 transform into the discrete correlations

$$\tilde{g}_m = \sum_n \tilde{f}_n \tilde{J}_{m+n} \qquad \tilde{f}_n = \sum_m \tilde{g}_m \tilde{J}_{m+n} \qquad (7.3.4)$$

where

$$\begin{aligned} \tilde{g}_m &= \rho_m g(\rho_m); & \rho_m &= \rho_o e^{am} \\ \tilde{f}_n &= r_n f(r_n); & r_n &= r_o e^{an} \\ \tilde{J}_m &= 2\pi a\, \rho_o r_o e^{am} J_\ell(2\pi\rho_o r_o e^{am}). \end{aligned} \qquad (7.3.5)$$

Given the N-point vector $\tilde{f}$, its N-point Hankel transform $\tilde{g}$ can be obtained by performing only two DFT's of 2N points each, provided the 2N-point inverse DFT of $\tilde{J}$ is already stored in memory. It is necessary to perform 2N-point DFT's and not N-point DFT's in order to simulate linear convolutions by performing cyclic DFT convolutions. The sequence of operations is straightforward:

- padd the N-point $\tilde{f}$ vector with N extra zeroes, and perform its 2N-point direct DFT;

- multiply the result by the 2N-point inverse DFT of $\tilde{J}$;

- perform a 2N-point DFT; the first N points of the result represent the vector $\tilde{g}$.

The success of the operation is guaranteed by a proper scaling of the signal to be transformed; i.e., by ensuring that the sampling is sufficiently dense, and that it is done in a sufficiently large number of points (N) to ensure the negligibility of both $\tilde{f}$ and $\tilde{g}$ beyond that interval. Siegman provides a heuristic analysis for determining N and the scaling parameters r_o, ρ_o, and a in function of two sampling coefficients K_1 and K_2 (both to be chosen greater than 2) and a space-bandwidth product $r_{max}\rho_{max}$, where r_{max} and ρ_{max} are the values above which $\tilde{f}$ and respectively $\tilde{g}$ become negligible:

$$N = K_2 r_{max}\rho_{max} \ln(K_1 r_{max}\rho_{max}),$$

$$a = 1/(K_2 r_{max}\rho_{max}), \qquad (7.3.6)$$

$$r_o\rho_o = 1/(K_1 r_{max}\rho_{max}).$$

The experimental results which Siegman provides are very encouraging, and the speed and memory savings obtained through the use of the fast Hankel transform versus the use of a direct integration transform are impressive proofs of the value of the numeric solution of m-convolution transforms described in Section 7.2. The advantages and problems involved in using this technique for the decomposition of multi-component exponential signals are discussed at large in the next section.

7.4 ANALYSIS OF MULTI-COMPONENT EXPONENTIAL DECAYS

The impulse for studying means of solving the integral equation (7.1.5) has been the need for decomposing (in the sense of spectral analysis) multi-component signals which appear in pulsed NMR experiments (Figure 7.7). As with most other application oriented investigations, the problem was initially set exclusively in terms of exponential decay type signals; i.e., given the experimental signal $x(t)$, find the parameters M, A_1 ...., A_M and λ_1, ...,λ_M (which describe it in the following manner:

$$x(t) = \sum_{m=1}^{M} A_m \exp(-\lambda_m t). \tag{7.4.1}$$

It was only later that the full potential of the spectral analysis technique (henceforth called the exponential transform technique) has been fully recognized [297]. This section will present a succint review of some of the results obtained by M.R. Smith and the author [309], concerning the analysis of signals of the form (7.4.1). Although only exponential type signals are discussed, most of the problems encountered apply as well to other types.

The problem of analyzing multi-component exponential signals is one of long standing in applied sciences, particularly because the transfer function of a linear time invariant system may be determined by expanding its transient response in terms of exponential components. Analysis of a single exponential decay is fairly straightforward, but

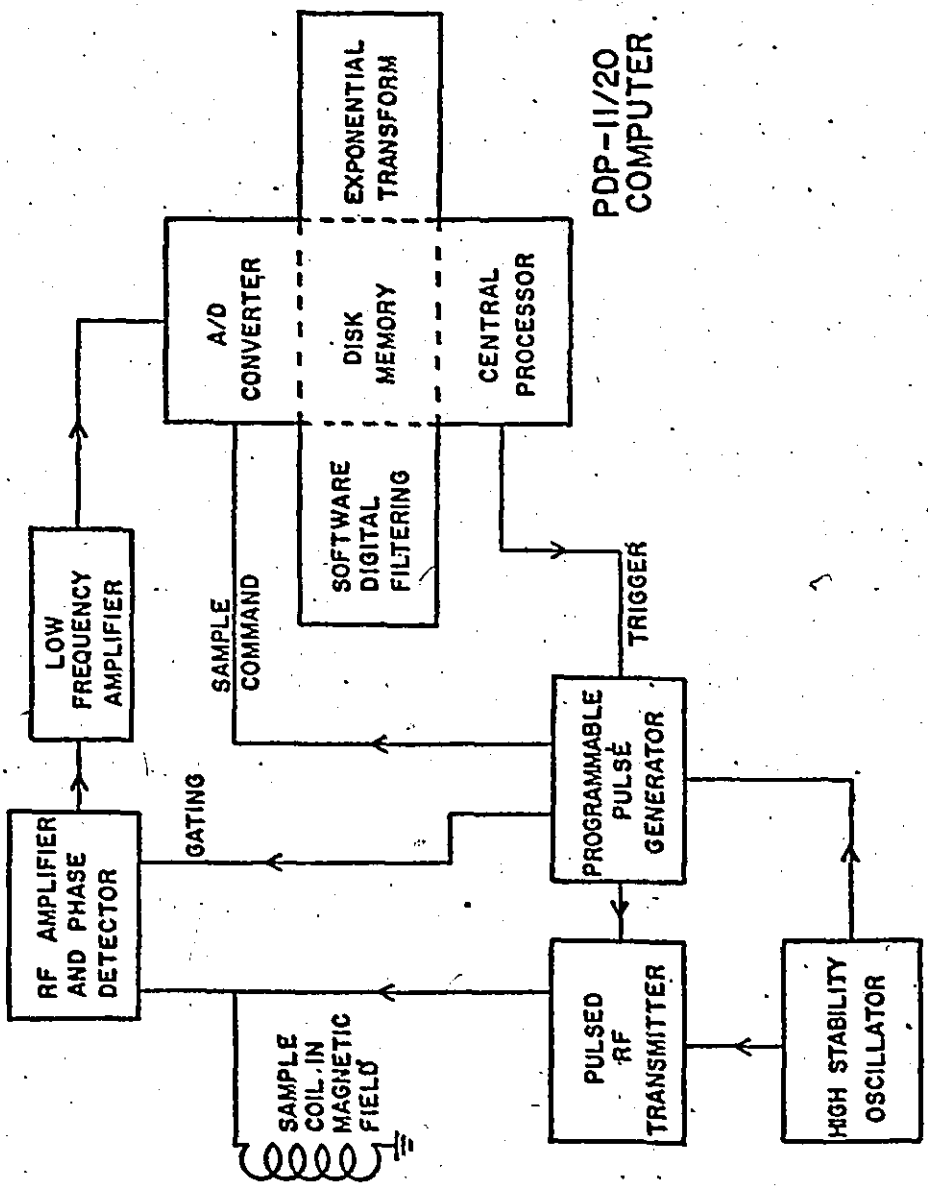


Figure 7.7. Block diagram of the pulsed Nuclear Magnetic Resonance spectrometer system.

that of multi-component decays is difficult, especially when the number of exponentials is unknown (or large), exponentials with similar decay constants are present, and when data is contaminated by noise and baseline offsets. Successive elimination of the slowest decay [310] is a simple technique which is satisfactory as long as there are few exponential components, and they are well resolved.

There are iterative techniques which synthesize multi-component exponentials and compare them to the original data until some goodness-of-fit criterion is met [31 ,31], but their applicability is limited by the finite computer time available, particularly since the presence of noise may influence their convergence. More sophisticated iterative techniques based on optimum least-squares approximations have been developed as generalizations of Prony's method (dated 1975!) which separates the search for the linear parameters A_M from that for the non-linear parameters λ_m. The main difficulty with this technique consists in the fact that it requires the solution to a non-linear eigenvalue problem. McDonough and Huggins [313] have proposed a hardware implmenetation of such a solution. Golub and Pereyra [285] and Osborne [286] present generalizations of Prony's technique in which the component signals have non-linear parameters, but not necessarily of the multiplicative type:

$$x(t) = \sum_{m=1}^{M} A_m k_m(t;b_1,b_2, \ldots, b_L) \qquad (7.4.2)$$

But, as with all other iterative techniques, their solutions are strongly dependent on a good initial guess of the unknown parameters, and they are

practically useless if M is unknown and-or relatively large. It is true, nevertheless, that, given a close initial guess of the parameters and a correct choice of M, they usually provide results which are more accurate than those obtained through other methods. (In addition, they are applicable even for signals which are composed of different types of kernels; but this is not of interest in the present case.)

It would be ideal if all these problems could be avoided by using a method which generates a spectrum (amplitudes A_m versus distribution of parameters λ_m) much in the same way the Fourier transform generates the spectrum of complex exponentials forming a signal. The advantages of such an approach are that:

- no a priori numeric information on the parameters is required;
- all the parameters are determined simultaneously (without any interference);
- broadened peaks would indicate the presence of (exponential) components with close (decay-) parameters without interfering with the analysis of other components;
- reliable analysis could be performed by personnel unskilled in curve-fitting (this is an important desideratum in the context of attempts at laboratory automation);
- even when not very accurate, the results may be used as a good initial set of parameters for a more refined iterative technique.

Three candidates for such exponential spectral analysis techniques have been considered in [309], where it was found that the only acceptable one is that described in the previous two sections. The transform based on the orthonormalization of a set of exponential functions [314,11] has been implemented, but the results proved to be discouraging. The distribution obtained provided a good fitting of the data, but the results were correct only if the non-linear parameters were integers. In any case, the method is not computationally efficient, and it requires to operate with extremely high precision arithmetic (the use of double precision wordlengths proved often not to be enough). The same negative conclusion has been drawn concerning the use of various algorithms for solving (7.4.1) through an inversion of a Laplace transform using only real-valued input data [288].

There remains the exponential transform spectral technique whose theory was described in Section 7.2, and whose practical implementation was discussed in Section 7.3. It is evident from the examples presented there that the technique is satisfactory when confronted with theoretically generated (simulated) data. But what about experimental data? The three major problems affecting the proper application of the technique refer to

- approximation of an aperiodic (linear) convolution by a cyclic one,
- the enhancement of noise present in the data, and
- the errors introduced by the interpolation and extrapolation procedures used to calculate the signal values at exponentially increasing intervals and over an extended range.

The first of these problems has been already discussed in a general context in Section 7.3. Concerning the positioning of the signal in the a-interval, it has been found experimentally that, for a single exponential analysis, data must be gathered over at least 22 time constants, and the peak of the transformed data must not be too close to either of the α-interval boundaries. The choice of sampling density (the word rate has a connotation of equidistant sampling) and total number of points N has to be made through a compromise; a large N is necessary for a good approximation of the FT by a DFT, but a large N also implies larger computational errors, particularly those associated with the interpolation and extrapolation necessary to calculate $r^{\alpha}x(r^{\alpha})$. Also, large N means longer computation times, N = 256 and $A_o = 12.8$ proved to be a satisfactory compromise.

Usually x(t) is gathered over a time interval of typically 6 to 10 times the longest time constant of the exponential present in the signal; such that an extrapolation is necessary to cover the entire 22 time constants interval mentioned above. It was found experimentally that a poor extrapolation is better than none. Unfortunately, such an extrapolation involves the portion of the data characterized by the poorest S/N, and it is rather critical.

Indeed, the major difficulty faced in the implementation of a digital spectral analysis is that of overcoming the noise problem. Real experimental signals are seldom free of any noise, and this noise may make

the final results meaningless; hence, the necessity of eliminating it one way or the other. By this, it is meant that the noise may be dealt with either at the source (by improvements of the experiment itself) or by an adequate processing of the final experimental data. Regarding pulse NMR and exponential decays, the latter can be obtained either in the form of signal averaging or in the form of digital filtering.

A particularity of exponential decay signals is that they decrease monotonically in time, and so does their S/N ratio. Smith and Cohn-Sfetcu [315] have shown that this particularity can be used to advantage if the signal is RC-filtered in reverse time direction. By reverse time filtering, it should be understood that, given a signal

$$x(t) = \sum_{m=1}^{M} A_m \exp(-\lambda_m t); \quad 0 < t \leq t_F = K\Delta t, \tag{7.4.3}$$

the input sequence to the digital filter is

$$z(k\Delta t) = x((K-k)\Delta t) = \sum_{m=1}^{M} A_m \exp(-\lambda_m (K-k)\Delta t). \tag{7.4.4}$$

A digital RC filter with a time constant $T_f \Delta t$ can be characterized by the difference equation

$$y(k\Delta t) = (1-c)z(k\Delta t) + cy((k-1)\Delta t), \tag{7.4.5}$$

where $c = \exp(-1/T_f \Delta t)$. The advantage of such a reverse time filtering over a normal forward filtering consists in the fact that, irrespective of the filter bandwidth, (1) it does not distort the exponential nature of

the decays, and (2) no starting transients are produced, because the signal increases smoothly from zero rather than starting abruptly. However, there is one distortion in the sense that the exponential decays composing the signal are unevenly enhanced, but in a known way: the longer their time constants, the more their relative amplitude is increased. Fortunately, this is to our advantage, because the long time constant components are the ones most affected by the noise. Consequently, pulsed NMR data with S/N better than 10^3 in the initial portion of the signal may be obtained either directly from the spectrometer and/or through an appropriate digital processing of the collected data.

It remains to discuss the interpolation of the experimental signal in order to estimate its values at exponentially increasing intervals. In principle, these values could be obtained without any interpolation via a direct exponential sampling. However, the pulse sequence applied during the NMR experiment and the use of digital filtering make more natural to sample the experimental signal at equidistant points. A simple linear interpolation between the two immediately adjacent points around $t_m = \exp(m\Delta\alpha)$ makes very inefficient use of the information contained by the sampled data since, for m large, it does not use all the points between t_{m-1} and t_{m+1}. Besides, the presence of noise drastically reduces the accuracy of such a simplistic evaluation, especially for m large - exactly where the S/N is already very poor.

In the end, an adaptive curve fitting procedure has been shown to yield satisfactory results [310]. This curve fitting procedure, besides providing the interpolation, introduces some smoothing of the data and increases the effective S/N. A fit in a semilogarithmic space - $\ln(x(t))$ vs t - has been adopted because the data appears quasi-linear there, thus permitting the use of a low-order polynomial fit without incurring excessive interpolation errors. A cubic fit interpolation is performed until there is an indication that only a single exponential component remains in the data [316], the remaining data being fitted to a single exponential curve [317]. The cubic fit interpolation parameters are monitored, and if it is concluded that, because of decreasing S/N, the fit is deviating from that expected for a multi-component exponential decay (e.g., the first derivative becomes positive), the curve fitting interval is increased and/or the cubic fit is replaced by a quadratic one. However, such an interpolation procedure places no constraints upon the continuity of adjacent cubic fits. This has the effect of introducing discontinuities into the interpolated function which are interpreted by the spectral analysis as fast exponential components and produce noise in the corresponding part of the final spectrum. This effect could be attenuated or avoided completely by employing a spline-fitting interpolation procedure which would ensure the continuity of the fits and their derivatives [318].

As mentioned earlier, more detailed discussions and results upon interpolation and spectral analyses of noisy data are presented in [309]. Simulated noisy data have been analyzed, with the conclusion that, in the

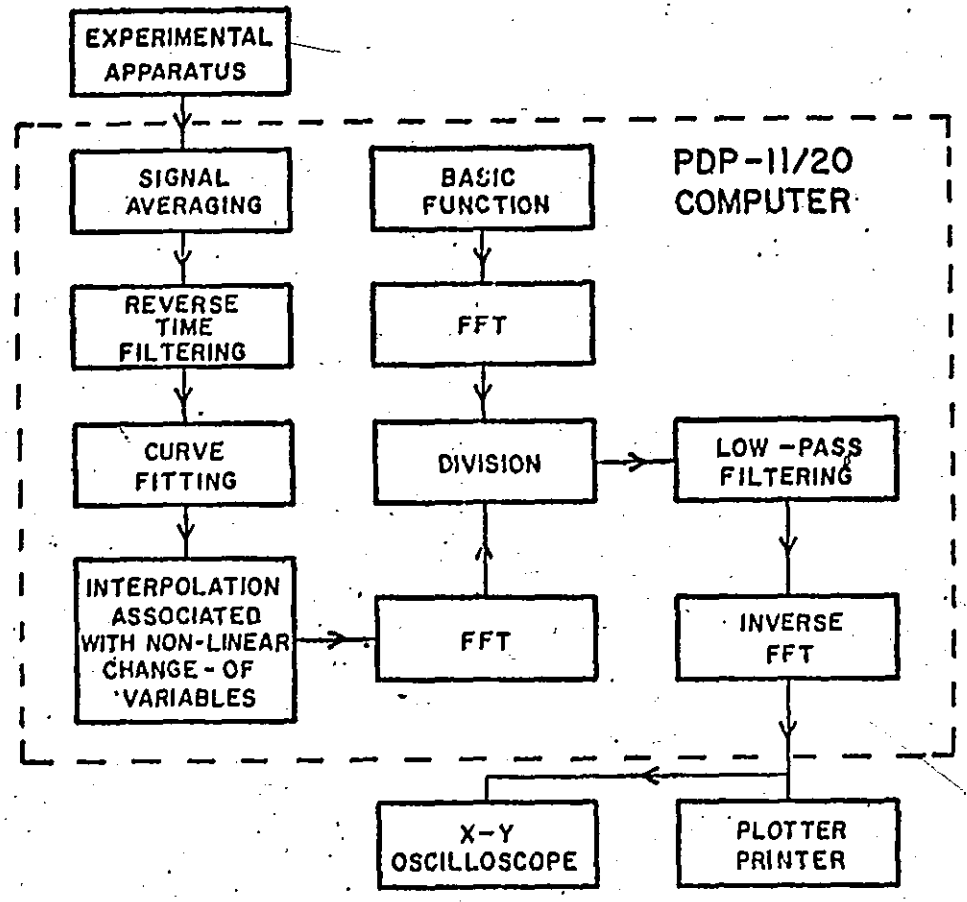


Figure 7.8. Flow chart of the signal processing performed upon the experimental data to produce an exponential transform spectrum.

absence of digital prefiltering, the S/N in the initial portion of the data has to be of the order of 10^3 to ensure acceptable end results. The influence of sampling jitter has also been investigated, and it was concluded that it is of little concern. Experimentally, it was found that a 10% rms error in time is equivalent to only 0.01% amplitude noise, since, for exponential decays, jitter produces amplitude-dependent noise, such that the effective S/N of the signal remains constant as time increases, whereas the S/N for normal amplitude noise decreases monotonically.

Figure 7.8 presents the signal processing flow chart which was used at the University of Calgary for analyzing pulsed NMR experimental data. The exponential transform spectrum of a set of NMR data of a breast biopsy sample tissue is shown in Figure 7.9. The data had been prefiltered by a digital low-pass filter in the manner described earlier. It can be clearly seen that there are four exponential components present, each corresponding to a different environment of the proton nuclei in the tissue sample. The effect of increasing the filter time constant can be also observed: there is an enhancement of the slower exponential decays, as it is evident from the increase in the relative amplitude of the left side spectral peaks. Better results are expected once the NMR instrumentation is improved; mainly, there is need for a detection system which is linear over a wider dynamic range, and which permits a better compensation (or, at least, determination) of baseline offsets.

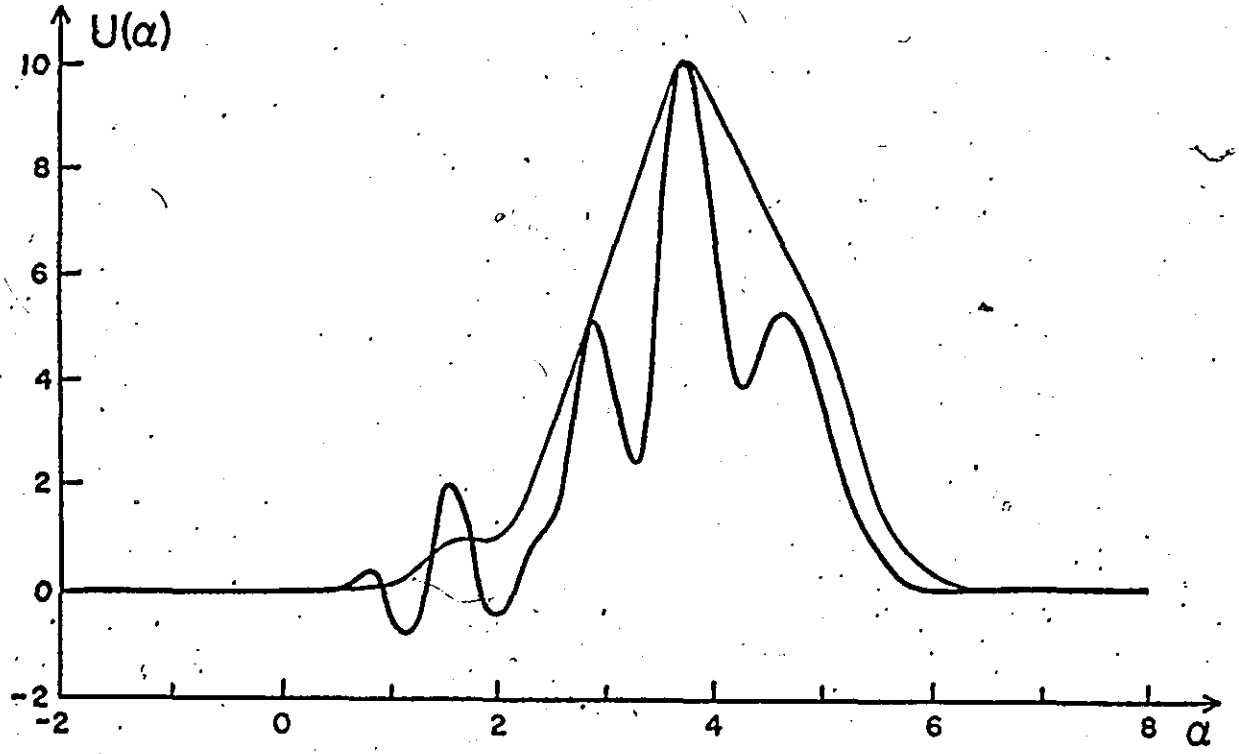


Figure 7.9. The exponential spectrum $U(\alpha)$ of experimental NMR data obtained from a breast biopsy sample. Four exponential components are indicated. Two Gaussian filters are employed: $\nu_g = 20$ (light line) and $\nu_g = 24$ (heavy line). $r = e$, $n = 1$, $N = 256$, $A_0 = 12.8$, no time aliasing.

a) Spectra of data after being reverse time filtered by a RC digital filter with a time constant of $10\Delta t$.

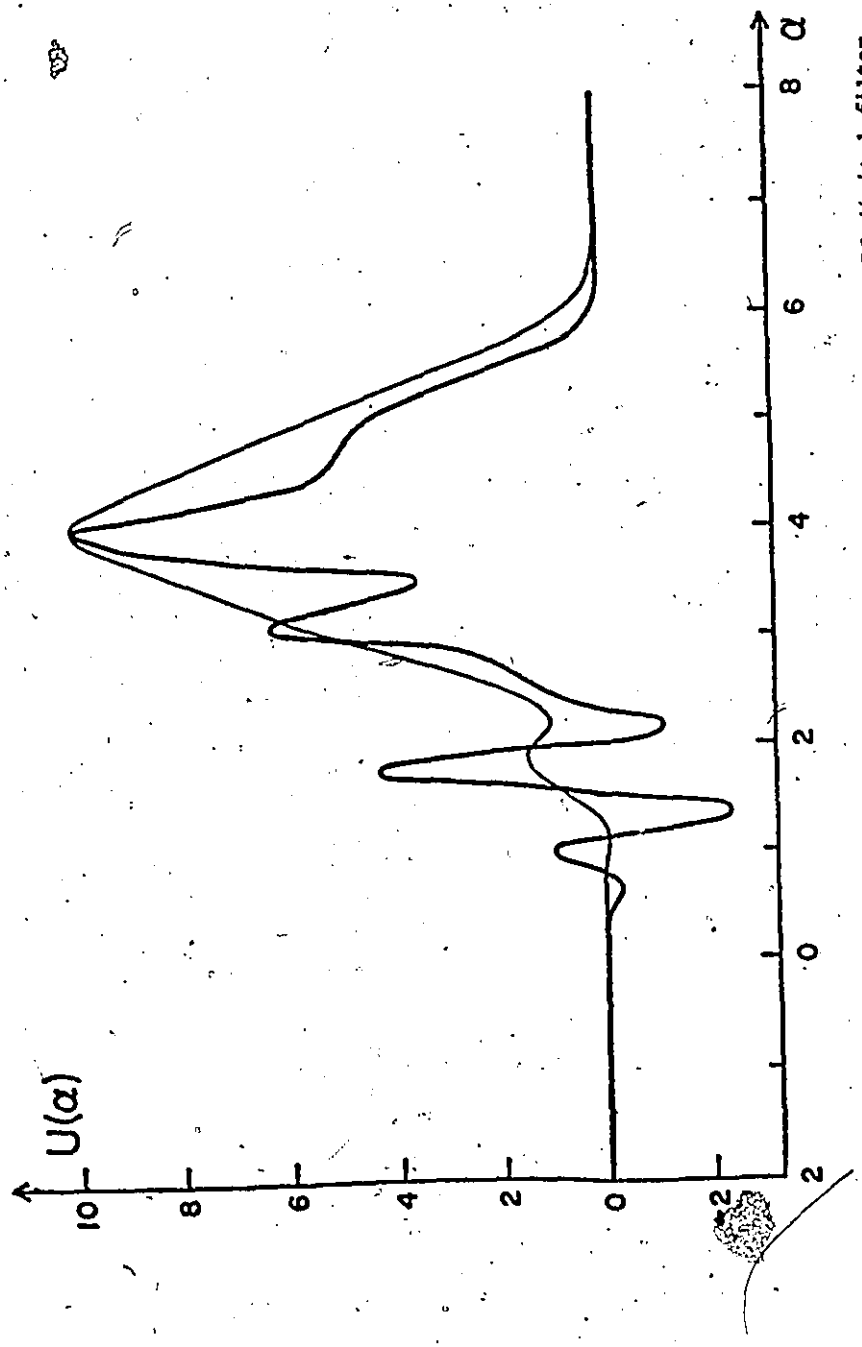


Figure 7.9. b) Spectra of data after being reverse time filtered by a RC digital filter with a time constant of $100\Delta t$.

A future practical implementation of the exponential transform technique is planned at McMaster University in connection with photon correlation and light beating spectroscopy [320,298], where data may be collected in the form of either a multi-component exponential decay or a multi-component Lorentzian signal.

7.5 CONCLUDING REMARKS

The investigation presented in this chapter has had the objective of providing an efficient spectral analysis technique for decomposing signals formed by a linear superposition of kernels (basic waveforms) of the same shape and location but various amplitudes and widths. In main, this objective has been attained.

First, the problem has been modelled as a multiplicative (Mellin) convolution integral equation. Solving such equations through a method amenable to an efficient numeric implementation forms the subject of Section 7.2, where a transformation of variables is used to map an m-convolution equation into a normal convolution one, solvable through Fourier transform techniques. The theory is presented both from an heuristic point of view and from the perspective of Fourier transforms on abstract abelian groups. As a byproduct, a new, faster and more accurate algorithm has been devised [39] for performing Hankel transforms.

Section 7.3 has been concerned with general problems affecting the numeric implementation of both the fast m-convolution and the m-deconvolution

(i.e., the multiplicative spectral analysis technique). The latter is a much more delicate operation because of its sensitivity to noise, due to the fact that the analysis implies a deconvolution process and a change of variables which enhance the noisiest part of the data. Also, it requires the collection or computation (interpolation) of the signal at increasingly large intervals. A correct interpolation is a rather difficult problem in itself, and this is the area where most of the improvements over present results could be achieved. In the end, the antagonism resolution versus S/N is the limiting factor in an experimental application of the technique described, a fact which is, nevertheless, characteristic to all spectral analysis techniques.

Compared to other procedures for analyzing multi-component signals, the spectral analysis technique described here has the advantages of being fast, relatively accurate and easy to implement, and, most important, of not requiring any a priori numeric information regarding the composition of the signal.

In addition, it is not limited to the analysis of discrete sums of components, since it is also capable of analyzing signals formed by a continuous distribution of such components. It was with this in mind that restraint was exerted from formulating the problem in terms of finding the amplitudes and widths of components in $x(t) = \sum_m A_m k(t/T_m)$; and the formulation in terms of the integral equation (7.2.1) was preferred instead.

The absence of a definite peak in a spectrum $u(s)$ may indicate either a continuous distribution of $k(t)$-type components, or the inadequacy of considering a $k(t)$-type basic waveform. The distinction between these two cases is rather fuzzy, but in real life applications, it can be made on account of the presence or not of a well-behaved spectrum u. This fact can be advantageously employed for determining the type of basis function when no theoretical model of the process under investigation is certain, and $k(t)$ is unknown. In such cases, the investigator should analyze the signal as a composition of various types of basis functions and decide on the most adequate one. For instance, suppose that there is uncertainty with regard to the correct $k(t)$, and the investigator applies both an exponential decay and a Gaussian spectral analysis to a signal which is actually composed of several Gaussian components. The exponential decay spectral analysis will yield a very noise-like continuous distribution of both positive and negative amplitude exponential decays, while the Gaussian transform spectrum will exhibit well-defined peaks. In such clear-cut cases, the investigator should, of course, choose the Gaussian model.

The application of this spectral analysis technique in connection with pulsed NMR experiments and exponential decays is discussed in Section 7.4. Pending dissemination, it is expected that this technique will find applications in other areas of scientific investigation (biology, chemistry, physics, radar, electrical engineering, optics, etc. ...) where multi-

component signals of the type analyzable by this technique occur. Also, the use of the exponential transformation for providing fast m-convolution transforms is expected to become advantageous beyond Siegman's application for fast Hankel transform algorithms.

CHAPTER VIII

SUMMARY AND SUGGESTIONS FOR FUTURE WORK

The investigation presented in this dissertation dealt, in broad lines, with engineering applications (signal processing and system theory) of convolution and Fourier transforms on locally compact abelian groups. Emphasis was placed on the study of such transforms on finite discrete groups because of the important role which they play in numeric applications. The purpose of this investigation was twofold, a fact which manifests itself in the structure of the thesis. On one hand, it contains a broad theoretical study of the significance of convolution and Fourier (-Walsh) transforms on q-adic groups, their character, and their applicability to a varied class of engineering problems. On the other, it describes an efficient practical solution to a well-defined signal processing problem, which also admits a modelling as a convolution transform on an abelian group, albeit a multiplicative one.

The wide range of possible applications of convolution and (Walsh-) Fourier transforms in signal processing and system identification and control leads to a conflict between the number of topics covered and the detail achieved in each topic. The compromise solution which has been attempted presents investigations of large areas of interest, but they are less rigorous than initially intended.

The dissertation begins with an introductory chapter which serves two functions. First and foremost, it lays the theoretical groundwork upon which the body of the thesis' investigation rests. Secondly, it acquaints the reader with the past research activity of the author [22], motivating thus the topics of research chosen, and explaining the line of thought adopted. The chapter starts from the two most basic assumptions made in constructing simple, mathematically manageable models of real-world phenomena and relationships: linearity and translation invariance. Linearity usually implies a function space with a field as co-domain and a restriction to linear operators, while translation invariance implies a function space with a domain which is an abelian group. Characteristically, the research reported in this dissertation is not directed to problems concerning the usual space of complex-valued functions of real variables, but to problems in other function spaces, specifically to complex-valued functions on either q-adic or multiplicative abelian groups and to finite-field-valued functions on finite and discrete abelian groups. From necessity (and for elegance), the mathematical background is presented in terms of functions on an abstract LCA group, with emphasis on the concepts of translation operators, convolutions, characters and Fourier transforms.

The first part of the thesis (Chapters I, II and III) deals exclusively with spaces of complex-valued functions on q-adic groups, and with Walsh functions and transforms. The research into this field was initiated in the hope of applying the dyadic Walsh transform for solving signal analysis and system identification problems related to magnetic

resonance spectroscopy. These problems admit elegant and efficient solutions via discrete Fourier transform techniques, and it was hoped that Walsh transform techniques could provide more (computationally) efficient solutions. This was not the case, and since similar results (in various other fields of research) were reported in the literature, it was decided to direct the investigation not towards a particular application of Walsh transforms, but towards obtaining a fundamental explanation of the experimental results already reported. The final objective of such an investigation is to delineate those engineering problems which are optimally suited to be solved using such techniques. The study is centered around three concepts (two mathematical and one engineering) which are fundamental to the theory and applications of Fourier transforms: harmonic differentiation, translation invariant linear operators, and generalized frequency.

Chapter II is dedicated to a presentation of Walsh functions and Walsh transforms. The presentation is rather lengthy, but, given the stated objective of the investigation and the wide range of guises under which Walsh functions appear in engineering applications, it was considered essential to introduce them from different points of view. Their classical, analytic definition is discussed first because this is how Walsh functions came to be devised, and how most of their properties were first proven. This is followed in Section 2.3 by a discussion of discrete finite Walsh functions (and Walsh transforms). After summarizing their usual definition through sampling of the "continuous" Walsh functions, attention is directed exclusively towards an algebraic definition of them as characters of a

finite q-adic group. This definition is much more elegant, and, in the end, much more advantageous because it leads to a natural embedding of Walsh analysis into the well-developed field of harmonic analysis on abelian groups. At the same time, this definition is more fundamental than others because it points directly to the characteristic structure of the Walsh transform and leads immediately to the explanation of experimental results concerning the engineering use of Walsh functions.

Following a group-theoretic definition, such an explanation is simple and straightforward: the Walsh functions are the characters of a function space having a multi-dimensional domain. The fact that they can be presented as functions on a one-dimensional domain is due to the existence of a bijective mapping between a q-adic group and a subset of the real line.

Section 2.5 is concerned with yet another definition of Walsh functions - as the eigenfunctions of certain linear operators which share some of the properties of classical, Newton-Leibnitz differential operators. Two theorems are proved after a brief presentation of past results concerning these operators called harmonic (Gibbs) differentiators. One theorem shows that a harmonic differentiator commutes with a translation operator, while the other one proves that this type of differential operators are not of the Ritt-Kolchin class because they do not obey the product rule [101].

Chapter III deals with the theory of q-adic translation invariant linear systems. As such, it represents a generalization of past work concerning dyadic linear systems [106]. The input - output description of q-TIL systems is analyzed in Section 3.2. This description concerns bounded linear transformations and Fourier transforms which map convolution function algebras into pointwise multiplication function algebras, and, hence, is formally identical with that of zero-state "time" invariant linear systems. This is a consequence of the fact that the I/O formalism is determined solely by the (locally compact) abelian group character of the domain, with no reference to the particular structure of the group involved.

The situation is different if one attempts to develop a state-space analysis of q-adic translation invariant systems. Formally, it is possible to construct a "harmonic state-space" theory of such systems, with the harmonic differentiator taking the place of the classical one. Section 3.4 does, indeed, present such an analysis, but its merit lies mainly not in developing this formalism, but in pointing out that such a formalism lacks conceptual significance. This is so because q-adic groups do not admit a group order, without which it is impossible to define a meaningful concept of state.

Apart from these theoretical considerations, Section 3.3 considers the problem of identifying (mostly man-made) dyadic systems. It is shown that, as expected, the I/O identification can be performed by algorithms

similar to those for identifying time invariant linear systems, only that Walsh functions take the place of complex exponentials, the Walsh transform replaces the classical Fourier transform, and stationary random processes are replaced by dyadic stationary ones. The identification of multiple input - multiple output q-adic systems is presented in Appendix A, together with a computer simulation example.

Following the theoretical developments of Chapters II and III, it was the role of Chapter IV to provide the answer to the question of optimal engineering usage of Walsh functions and the Walsh transform. After a brief introduction, the chapter continues with an analysis of the computational advantages of the Walsh transform. Two theorems are proved in the course of this analysis. One concerns the fact that of all the discrete Fourier transforms the dyadic and 4-adic Walsh transforms are the only ones with coefficients representable by rational numbers. The other theorem proves that no non-trivial z-transform of a sequence of length N can be evaluated by rational numbers of less than N digits [160].

Section 4.3 describes in depth the "multi-dimensional" character of the Walsh transform, and shows that this is an essential feature stemming from the q-adic group structure of the functions' domain. In its turn, this feature explains the unsuccessful attempts at employing the Walsh transform in place of the discrete Fourier transform for solving various signal processing problems [33]. Section 4.4 shows that, obviously, the problems which admit an optimal solution via Walsh analysis are exactly

those which can be modelled in spaces of functions on q-adic groups. A succinct review of such applications of Walsh analysis is presented, with emphasis on coding and on the classification and synthesis of switching functions.

Besides the above-mentioned problems to which Walsh functions may bring a theoretically optimal solution, there are many other engineering problems for which Walsh analysis may bring a computationally efficient solution. Section 4.5 reviews critically some applications of the Walsh transform in which its computational advantage offsets its "theoretical unsuitability". These problems can be classified into three major classes: signal representation and analysis, signal recovery, and system identification and control. Several examples of Walsh solutions to such problems are discussed, emphasis being placed on the concept of intermediate domain processing. An original contribution is made with respect to the possibility of using intermediate q-adic models for identifying translation invariant systems [320]. Also, it is shown that this is the area where much work remains to be done, especially concerning the goodness of such estimates.

Chapter V is concerned with signal processing and system theory on spaces of functions taking values in finite fields. The point of view adopted is original and different from that prevalent in other investigations into this topic. Interest is manifested not towards problems pertaining to simulations of complex-valued processors via convolutions in

finite fields, but towards an analysis of finite field systems per se. Specifically, a study is made of those operators whose eigenfunctions are the characters defining the Fourier transform on the respective function space. Harmonic differential operators in finite fields are defined, and some of their essential properties are studied [34]. Linear translation invariant systems in K[G] function spaces are analyzed from the point of view of both input-output and state-space description. The study presented in this chapter is only preliminary - an initial step towards an elaboration of a comprehensive theory of optimal signal processing in function spaces in which both the domain and the co-domain are finite and discrete sets.

The concept of frequency plays a fundamental role in applied sciences, and Chapter VI provides a consistent and unique criterion for generalizing it to function spaces other than the customary space of complex-valued functions on the real line. This generalization is made to concord with the applied scientists' view of frequency as a parameter which, by ordering the set of complex exponentials, serves to measure the speed of variation of a signal. The first step on the way to obtaining the general concept consists in an analysis of the features of the classical concept in order to extract those which are essential enough to be features of the general concept. This is done in Section 6.2. The concept of generalized frequency [17] is defined in Section 6.3. Section 6.4 applies this concept to one-dimensional discrete Fourier transforms, while Section 6.5 applies it to Walsh transforms. It is concluded [33] that, not with-

standing previous claims, sequency is not a generalized frequency according to the criteria established in this chapter. This should not come as a surprise, because the multi-dimensional structure of a q-adic group precludes any possibility of associating a concept of generalized frequency to a set of Walsh functions.

Chapter VII considers a specific signal processing problem: that of analyzing multi-component signals in which the components have the same location and shape, but different amplitude and "width" parameters. The analysis of such signals is essential for interpreting data gathered in a variety of experimental investigations. Methods for performing such analyses had already been suggested, but they are either computationally inefficient, or they require a priori numeric information concerning the composition of the signal. It is the purpose of this chapter to devise a technique which does not require any a priori information and which is, at the same time, computationally efficient [31]. The theory of the technique is presented in Section 7.2, where it is shown that the problem can be modelled as a convolution transform in a space of complex-valued functions on a multiplicative abelian group. Some elements of the theory of Fourier transforms on such groups are introduced.

A suitable transformation of variables is proposed to map a multiplicative deconvolution problem into a (normal) deconvolution problem in the $\mathcal{L}[R]$ function space so that use can be made of the considerable knowledge available with respect to the use of discrete Fourier transform

techniques for solving such problems [297]. Section 7.3 discusses the numeric implementation of the algorithm proposed, with emphasis on the compromise which has to be made between loss of resolution and decrease in signal to noise ratio.

Section 7.4 is dedicated to the particular case of multi-component exponential decays. It is shown that the technique proposed is suitable for analyzing experimental data (pulsed NMR) which is not excessively contaminated by noise [309]. The use of reverse-time digital RC filters [316] for enhancing the signal to noise ratio of such signals is also discussed. Attention is drawn to problems pertaining to the collection or calculation (through interpolation) of data at exponentially increasing intervals.

The research into the subject of multiplicative convolution transforms is by no means completed, and work remains to be done concerning both their theory as well as their efficient numeric implementation. The elaboration of a theory, and of fast algorithms for performing finite discrete multiplicative Fourier transforms may prove to be an interesting and fertile field of research. The success of the analysis technique presented in Chapter VII relies heavily on the success of performing numerical deconvolutions. This is still an open problem, and more efficient and less noise-sensitive deconvolution algorithms are actively investigated. Another delicate operation involved is the interpolation of experimental data, and more research is needed to obtain interpolation procedures which

lead to less distortions. This problem is compounded by the fact that the interpolation is dependent on the type of components in the signal, and no "universally" good interpolation procedure can be envisaged. Also, the transformation of noise under the required change of variables deserves a closer look, although this may prove to be a difficult problem to tackle due to the large number of variables involved.

The use of the analysis technique described in Chapter VII should be extended to experimental data formed of signal formed of components other than exponential decays, and obtained through experiments other than NMR. It has been already shown that this technique enables a fast and efficient computation of Hankel transforms. Other similar transforms should be also investigated.

Another idea which may prove fruitful and deserves a deeper investigation is that of generalized translation invariance, especially problems pertaining to the realization and/or estimation of generalized translation invariant systems via linear "time" invariant systems. Or, conversely - it has been seen how linear time invariant systems can be identified via estimates of q-TIL systems.

APPENDIX A

IDENTIFICATION OF MULTIPLE INPUT-MULTIPLE OUTPUT LINEAR DYADIC SYSTEMS

The problem of identifying multiple input-multiple output linear dyadic systems can be formulated identically as for the case of single input-single output linear dyadic systems with the only exception that now we are dealing with, for generality, a k-dimensional input signal and a m-dimensional output signal. Due to the linear restriction of the problem, it is evident that a k-input-m-output identification problem can be reduced to m k-input-1-output system identification problems.

We shall, then, present the algorithm for identifying the linear dyadic convolution system characterized by a k-dimensional (column) operator $H(t)$ such that the output $z(t)$ to a k-dimensional (column) random process $X(t)$ is the best approximate in the mean squared sense to the one-dimensional random process $Y(t)$. The proof follows the same logical steps as for the single input case and it will be omitted here due to lack of time and space. For concision, the algorithm is presented using matrix notation (a matrix is marked by a capital character). The interaction between the input signal and the model (system) is given by the following two equations:

$$z(t) = H^T(t) \,\theta\, X(t) = \sum_{i=1}^{k} [h_i(t) \,\theta\, x_i(t)] \qquad (A.1.a)$$

$$Z(z) = \hat{H}^T(\lambda) \cdot \hat{X}(\lambda) = \sum_{i=1}^{k} [\hat{h}_i(\lambda) \cdot \hat{x}_i(\lambda)] \qquad \text{(A.1.b)}$$

Let $R_{XX}(\tau)$ and $R_{Xy}(\tau)$ be the autocorrelation and the crosscorrelation matrices, respectively

$$R_{XX}(\lambda) = E[X(t) \cdot x^T(t \oplus t)] \qquad \text{(A.2.a)}$$

$$R_{Xy}(\lambda) = E[X(t) \cdot y^T(t \oplus \tau)] \qquad \text{(A.2.b)}$$

and also let $S_{XX}(\tau)$ and $S_{Xy}(\tau)$ be the input Walsh power density spectral matrix and the Walsh cross-spectral density matrix, respectively.

$$S_{XX}(\lambda) = WR_{XX}(\tau) = \begin{bmatrix} s_{x_1x_1}(\lambda) & \cdots & s_{x_1x_k}(\lambda) \\ \vdots & & \vdots \\ s_{x_kx_1}(\lambda) & & s_{x_kx_k}(\lambda) \end{bmatrix} \qquad \text{(A.3.a)}$$

$$S_{Xy}{}^T(\lambda) = WR_{Xy}{}^T = [s_{x_1y}(\lambda) \ \cdots \ s_{x_kx_k}(\lambda)]. \qquad \text{(A.3.b)}$$

Similarly, we can define an augmented spectral matrix $S_{yXX}(\lambda)$ as follows:

$$S_{yXX}(\lambda) = \begin{bmatrix} s_{yy}(\lambda) & s_{yx_1}(\lambda) & \cdots & s_{yx_k}(\lambda) \\ s_{x_1y}(\lambda) & s_{x_1x_1}(\lambda) & \cdots & s_{x_1x_1}(\lambda) \\ \vdots & \vdots & & \vdots \\ s_{x_ky}(\lambda) & s_{x_kx_1}(\lambda) & \cdots & s_{x_kx_k}(\lambda) \end{bmatrix} =$$

$$= \begin{bmatrix} S_{yy}(\lambda) & S_{Xy}(\lambda) \\ \hline S_{XX}(\lambda) & S_{XX}(\lambda) \end{bmatrix}. \qquad (A.4)$$

The optimum linear dyadic convolution system relating the output signal $y(t)$ to the input signal $X(t)$ is given by

$$\hat{H}(\lambda) = S_{Xy}(\lambda) \cdot S_{XX}^{-1}(\lambda). \qquad (A.5)$$

The goodness of the estimate is measured with the help of partial and multiple coherence functions. A partial coherence function represents the fractional portion of the mean square value at the output $y(t)$ which is contributed by the input $x_i(t)$ at the sequency λ when the presence of all the other inputs is taken fully into account.

$$\gamma_{x_i y \cdot X_-}^2(\lambda) = \frac{S_{x_i y \cdot X_-}(\lambda)}{S_{x_i x_i \cdot X_-}(\lambda) S_{yy \cdot X_-}(\lambda)} \qquad (A.6)$$

where $S_{x_i y \cdot X_-}$, $S_{x_i x_i \cdot X_-}$ and $S_{yy \cdot X_-}$ are the elements of the residual spectral matrix defined as

$$S_{x_i y \cdot X_-}(\lambda) = \begin{bmatrix} S_{yy \cdot X} & S_{y x_i \cdot X_-} \\ S_{x_i y X_-} & S_{x_i x_i \cdot X_-} \end{bmatrix} = A_i - B_i C_i B_i^T \qquad (A.7)$$

and where A_i, B_i, and C_i are matrices obtained through the following partitioning of the augmented spectral matrix

$$A_i(\lambda) = \begin{bmatrix} S_{yy} & S_{y x_i} \\ S_{x_i y} & S_{x_i x_i} \end{bmatrix} \qquad (A.8.a)$$

$$B_i(\lambda) = \begin{bmatrix} S_{yx_1} & \cdots & S_{yx_{i-1}} & S_{yx_{i+1}} & \cdots & S_{yx_k} \\ S_{x_ix_1} & \cdots & S_{x_ix_{i-1}} & S_{x_ix_{i+1}} & \cdots & S_{x_ix_k} \end{bmatrix} \qquad \text{(A.8.b)}$$

$$C_i(\lambda) = \begin{bmatrix} S_{x_1x_1} & \cdots & S_{x_1x_{i-1}} & S_{x_1x_{i-1}} & \cdots & S_{x_1x_k} \\ \vdots & & \vdots & \vdots & & \vdots \\ S_{x_{i+1}x_1} & \cdots & S_{x_{i-1}x_{i-1}} & S_{x_{i-1}x_{i+1}} & \cdots & S_{x_{i-1}x_k} \\ S_{x_{i+1}x_1} & \cdots & S_{x_{i+1}x_{i-1}} & S_{x_{i+1}x_{i+1}} & \cdots & S_{x_{i+1}x_k} \\ \vdots & & \vdots & \vdots & & \vdots \\ S_{x_kx_1} & \cdots & S_{x_kx_{i-1}} & S_{x_kx_{i+1}} & \cdots & S_{x_kx_k} \end{bmatrix} \qquad \text{(A.8.c)}$$

The multiple coherence function is an overall measure of the goodness of the model (system):

$$\gamma_{y\cdot X}^2(\lambda) = 1 - \frac{|S_{yXX}(\lambda)|}{S_{yy}(\lambda)|S_{XX}(\lambda)|} \qquad \text{(A.9)}$$

where $|S|$ denotes the determinant of the matrix S. It can be proved that the multiple coherence function represents the Walsh cross-spectral density matrix between $z(t)$, the model output, and $y(t)$, the actual output, i.e., it measures the fraction of power in the actual output accounted for by the model.

For exemplification purposes, we present here the identification in a noisy environment of a 2 input-1 output linear dyadic invariant system. As it is shown in Fig. A.1, the output $y(t)$ of the dyadic system, consisting of a sequency low-pass filter, $\hat{h}_1$ and a sequency bandpass filter $\hat{h}_2$, is observed in a noisy signal $x_n(t)$. This setup was simulated on a CDC 6400 computer via the equation

$$y(t) = [\hat{h}_1(\lambda) \cdot Wx_1(t) + \hat{h}_2(\lambda) \cdot Wx_2(t)] + x_n(t) \qquad (A.10)$$

where t and λ take values from 0 to 31. The transfer functions were generated and stored in the computer. The Walsh transform W has been computed by a Fast Walsh Transform algorithm [88]. For identifying this dyadic system, we have used as input signals, $x_1(t)$ and $x_2(t)$, two sequences of normally distributed random numbers which were generated by standard subroutines from the Fortran IV scientific library package. The noise sequence $x_n(t)$ was generated likewise, its rms amplitude being 1/100 that of either input signal. The sequency spectral density functions of the input and output signals were estimated by computing the average of 100 corresponding Walsh periodograms using Welch's direct method [321] with suitable variation for the Walsh transform. For instance,

$$s_{x_1x_2}(\lambda) = (1/100) \sum_{r=1}^{100} [Wx_{1,r}(t) Wx_{2,r}(t)] \qquad (A.11)$$

where $x_{1,r}(t)$ and $x_{2,r}(t)$ denote the kth x_1 and x_2 signals, respectively (each signal is a sequence of 32 numbers). This averaging represents an approximation of the expectation operator E. The identification procedure has been carried on according to Eqs. A.5, A.6 and A.9.

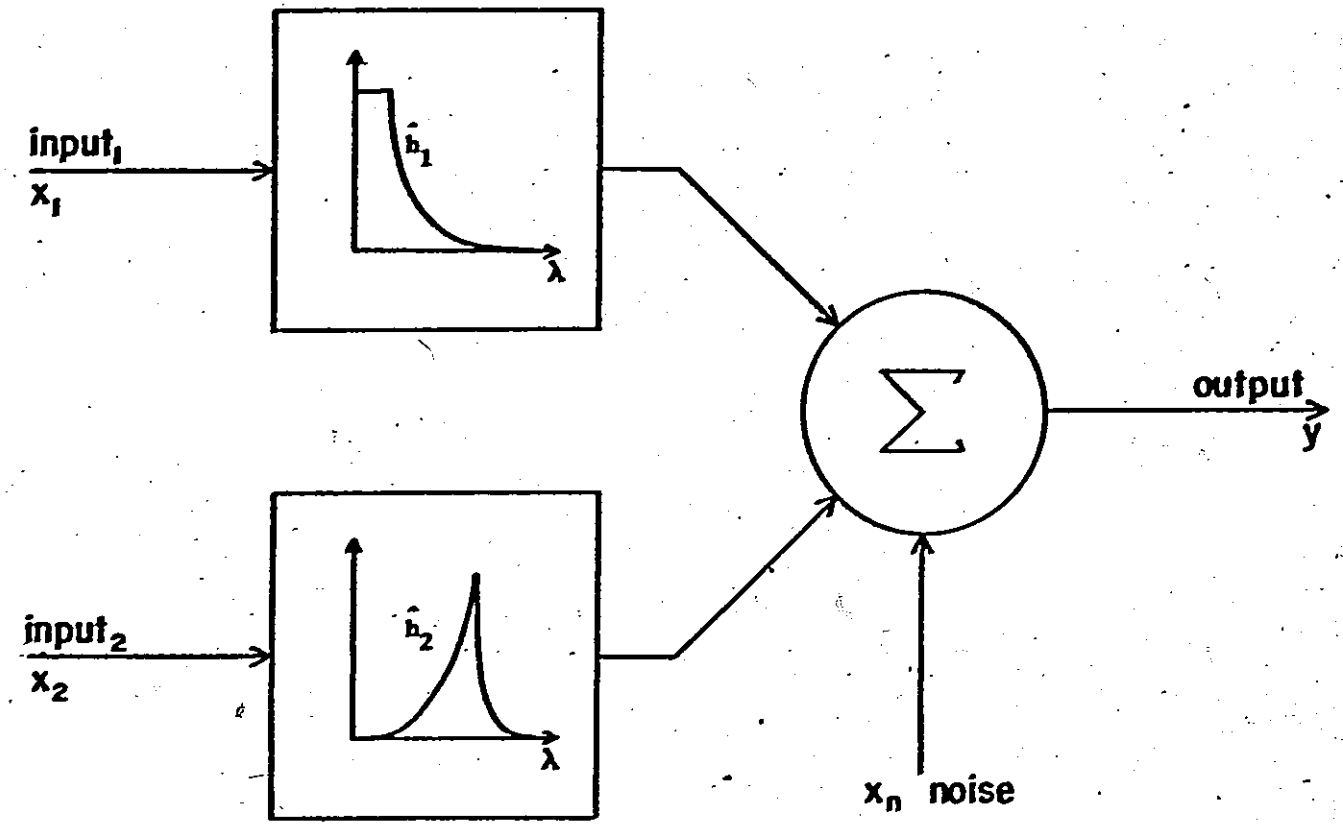


Figure A.1. Block diagram of simulated 2 input - 1 output linear dyadic invariant system operating in noisy environment.

The total computation time was 11.705 seconds (the computation of each set of Walsh periodograms requires the generation of the three random number sequences, 4 Walsh transforms, 8 point by point vector multiplications, and 9 vector additions). The model system resulting from this identification is presented in Fig. A.2(a); for comparison, the theoretical system is also represented by continuous lines. Fig. A.2(b) presents the coherence functions characterizing the identification and it shows that, as expected, the departure from unity of the coherence functions is a good local indication of the degree of faithfulness with which the model system estimates the real one.

A detailed analysis of the errors associated with the identification of dyadic convolution systems via Walsh spectral analysis exceeds the purpose set for this Appendix. It suffices to say that it follows closely the similar analysis for the identification of linear time-invariant systems. A general discussion of these errors can be found in [150] while a more detailed one is presented in [133, 322].

We would like to conclude by mentioning the fact that while many linear dyadically symmetric systems have been artifically created and then used for a variety of purposes (hence the need of developing adequate methods for their identification), the question regarding the existence of natural dyadic systems (or phenomena) still remains unconvincingly answered.

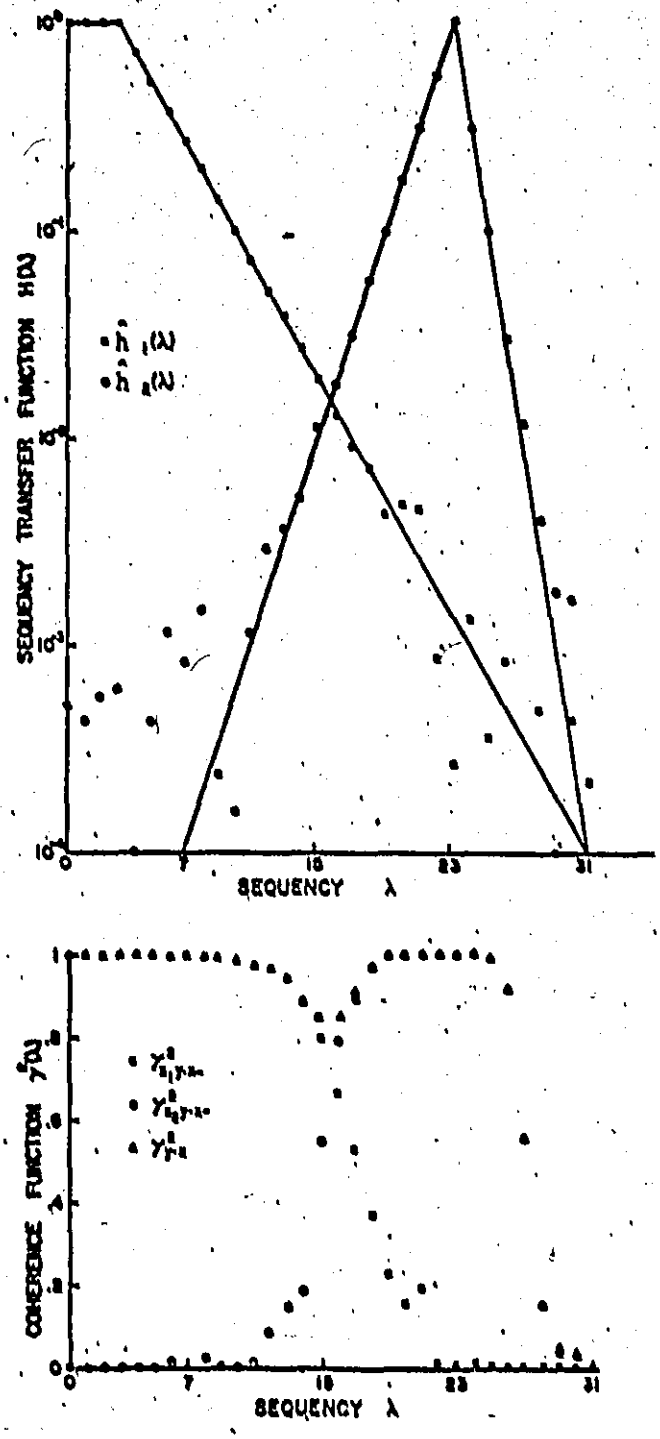


Figure A.2. (a) Sequency transfer functions of identified model system for situation depicted in Fig. A.1.
(b) Sequency coherence functions characterizing identification of system illustrated in Fig. A.1.

REFERENCES

1. J.G. Mikusinski, *The Calculus of Operators*. Oxford: Pergamon Press, 1959.

2. W. Rudin, *Fourier Analysis on Groups*. New York: Interscience, 1960.

3. G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*. New York: MacMillan, 1965.

4. R. Godement, *Algebra*. Paris: Hermann, 1968.

5. E.R. Lorch, *Spectral Theory*. New York: Oxford Univ. Press, 1962.

6. E. Hewitt, "The role of compactness in analysis", Amer. Math. Monthly, Vol. 67, pp. 499-516, 1960.

7. P.R. Halmos, *Measure Theory*. Princeton, N.J.: d.Van Nostrand, 1950.

8. W. Rudin, *Real and Complex Analysis*. New York: McGraw Hill, 1974.

9. H.J. Bremermann, *Distributions, Complex Variables, and Fourier Transforms*. Reading, Mass.: Addison-Wesley, 1965.

10. A.H. Zemanian, *Generalized Integral Transformations*. New York: Interscience, 1968.

11. L.E. Franks, *Signal Theory*. Englewood Cliffs, N.J.: Prentice Hall, 1969.

12. C.T. Chen, *Introduction to Linear System Theory*. New York: Holt, Rinehart and Winston, 1970.

13. L. Pontryagin, *Topological Groups*. Princeton, N.J.: University Press, 1946.

14. R.E. Edwards, *Fourier Series: A Modern Introduction*. Vol. 1, New York: Holt, Rinehart and Winston, 1967.

15. J.W. Cooley and J.W. Tukey, "An algorithm for the machine computation of complex Fourier series", Math. Comput., Vol. 19, pp. 297-301, 1965.

16. I.J. Good, "The interaction algorithm and practical Fourier analysis", J. Roy. Stat. Soc. (London), Vol. B-20, pp. 361, 1958.

17. J.W. Cooley, P.A.W. Lewis and P.D. Welch, "Application of the fast Fourier transform to computation of Fourier integrals, Fourier series, and convolution integrals", IEEE Trans. Audio and Electro Acoustics, Vol. AU-15, pp. 79-84, 1967.

18. F. Abramovici, "The accurate calculation of Fourier integrals by the fast Fourier transform technique", J. Comput. Phys., Vol. 11, pp. 28-37, 1973.

19. A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing*. Englewood Cliffs, N.J.: Prentice Hall, 1975.

20. L.R. Rabiner and B. Gold, *Theory and Applications of Digital Signal Processing*. Englewood Cliffs, N.J.: Prentice Hall, 1974.

21. D.G. Childers and A.E. Durling, *Digital Filtering and Signal Processing*. St. Paul, Minn.: West Publishing, 1975.

22. S. Cohn-Sfetcu, EPR *Spectrometer Sensitivity Studies: An Investigation of Digital Fourier Analysis in Optimizing the Performance of Various EPR Spectrometer Systems*. M.Sc. (Physics) Thesis, U. of Calgary, 1973.

23. H.A. Buckmaster and S. Cohn-Sfetcu, "Noise measurements in EPR sensitivity studies and design: Part I: Instrumentation, Part II: Applications", in *Magnetic Resonance and Related Phenomena*, Ed., V. Hovi, Amsterdam: North Holland, pp. 276-279 and 280-282, 1973.

24. S. Cohn-Sfetcu and H.A. Buckmaster, "A new method to measure the figure of merit of microwave detector diodes", IEEE Trans. Instrum. and Measurement, Vol. IM-23, pp. 102-103, 1974.

25. R.R. Ernst, *Sensitivity Enhancement in Magnetic Resonance*. *Advances in Magnetic Resonance*, Vol. 2, New York: Academic Press, 1966.

26. H.A. Buckmaster, S. Cohn-Sfetcu and D.J.I. Fry, "Optimum filtering for sensitivity enhancement in magnetic resonance", J. Mag. Res., Vol. 14, pp. 296-307, 1974.

27. M. Caprini, S. Cohn-Sfetcu and A. Manof, "Application of digital filtering in improving the resolution and the S/N of nuclear and magnetic resonance spectra", IEEE Trans. Audio and Electro Acoustics, Vol. AU-18, pp. 389-394, 1970.

28. S. Cohn-Sfetcu, M. Caprini and A. Manof, "Digital processing of experimental data using the discrete Fourier transform". *Proc. XVI Congress AMPERE*, pp. 822-831, Ed., I. Ursu, Amsterdam: North Holland, 1971.

29. S. Cohn-Sfetcu, M. Caprini and A. Manof, "Digital filter for resolution improvement", Nucl. Instrum. and Methods, Vol. 85, pp. 329-332, 1970.

30. E.C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*. Oxford: Clarendon Press, 1948.

31. S. Cohn-Sfetcu, M.R. Smith, S.T. Nichols and D.L. Henry, "A digital technique for analyzing a class of multi-component signals", Proc. IEEE, Vol. 63, pp. 1460-1467, 1975.

32. N.M. Blachman, "Sinusoids versus Walsh functions", Proc. IEEE, Vol. 62, pp. 346-354, 1974.

33. S. Cohn-Sfetcu, "The question of sequency: a spatial interpretation of Walsh analysis", Proc. IEEE 1976 Int. Symp. on Electromagnetic Compatibility, Washington, D.C., pp. 278-283, 1976.

34. C.M. Rader, "Discrete convolutions via Mersenne transforms", IEEE Trans. Computers, Vol. C-21, pp. 1269-1273, 1972.

35. R.C. Agarwal and C.S. Burrus, "Number theoretic transforms to implement fast digital convolutions", Proc. IEEE, Vol. 63, pp. 550-560, 1975.

36. S. Cohn-Sfetcu and J.E. Gibbs, "Harmonic differential calculus and filtering in Galois fields", Proc. 1976 IEEE Conf. on Acoustics, Speech and Signal Processing, Philadelphia, Pa., pp. 148-153, 1976.

37. H.F. Harmuth, "A generalized concept of frequency and some applications", IEEE Trans. Inf. Theory, Vol. IT-14, pp. 375-382, 1968.

38. J.E. Gibbs and S. Cohn-Sfetcu, "On the concept of generalized frequency", Proc. 1976 Int. Symp. Inf. Theory, Ronneby, Sweden, 1976.

39. A.E. Siegman, "A quasi-fast Hankel transform using the fast Fourier transform", to appear.

40. M. Kline, *Mathematical Thought from Ancient to Modern Times*. New York: Oxford University Press, 1972.

41. A. Haar, "Zur theorie der orthogonalen Functionen-Systeme", Matematische Annalen, Vol. 71, pp. 38-53, 1912.

42. P. Levy, "Sur une generalisation des fonctions orthogonales de M. Rademacher", Comment. Math. Helv., Vol. 16, pp. 144-152, 1944.

43. A.L. Brown and A. Page, *Elements of Functional Analysis*. London: Van Nostrand Reinhold, 1970.

44. L.J. Walsh, "A closed set of normal orthogonal functions", Amer. J. Math., Vol. 45, pp. 5-24, 1923.

45. P. Civin, "Multiplicative closure and the Walsh functions", Pacific J. Math., Vol. 2, pp. 291-295, 1952.

46. N.J. Fine, "On groups of orthonormal functions (I)", Pacific J. Math., Vol. 5, pp. 51-59, 1955.

47. S. von Kaczmarz, "Uber ein Orthogonal system", Comptes Rendus du Premier Congres des Mathematiciens des Pays Slaves, Warsaw, pp. 189-192, 1929. (See also S. Kaczmarz and H. Steinhaus, "Le system orthogonal de M. Rademacher", Studia Math., Vol. 2, pp. 231-247, 1930.)

48. R.E.A.C. Paley, "A remarkable series of orthogonal functions (I and II)", Proc. London Math. Soc., Vol. 34, pp. 241-264 and 265-279, 1932.

49. S. Toni, "Su un notevole sistema orthogonale di funzioni", Atti. Acad. Sci. Inst. Bologna, Cl. Sci. Fis., Ann. 246, Rend. XI, pp. 225-230, 1958. (In English translation by J.E. Gibbs et al, DES Rept. No. 12, National Physical Laboratory, England, 1972.)

50. R.P. Larsen and W.R. Madych, "Walsh like expansions and Hadamard matrices", IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-24, pp. 71-75, 1976.

51. E.F. Vandivere, "Symmetry analysis: a generalization of finite Walsh-Fourier analysis", IEEE Trans. Audio and Electro Acoustics, Vol. AU-20, pp. 225-226, 1972.

52. N.A. Alexandridis, "Relations among sequency, axis symmetry, and period of Walsh functions", IEEE Trans. Inf. Theory, Vol. IT-17, pp. 495-497, 1971.

53. N.J. Fine, "On the Walsh functions", Trans. Amer. Math. Soc., Vol. 65, pp. 372-414, 1949.

54. N.J. Fine, "The generalized Walsh functions", Trans. Amer. Math. Soc., Vol. 69, pp. 66-77, 1950.

55. J.E. Gibbs, "Walsh spectrometry, a form of spectral analysis well suited to binary digital computation", Report, National Physical Lab., Teddington, England, 1966.

56. P.L. Butzer and H.J. Wagner, "Approximation by Walsh polynomials and the concept of a derivative", Proc. 1972 Symp. Applic. Walsh Functions, Washington, D.C., pp. 27-29, 1972.

57. H. Butin, "A compact definition of Walsh functions", IEEE Trans. Comput., Vol. C-21, pp. 590-592, 1972.

58. T.M. Chien, "On representations of Walsh functions", IEEE Trans. on Electromagnetic Compat., Vol. EMC-17, pp. 177-180, 1975.

59. H. Rademacher, "Einige Satze von allgemeinen Orthogonal funktionen", Math. Annalen, Vol. 38, pp. 122-138, 1922.

60. C. Yuen, "Walsh functions and Gray code", Proc. 1971 Symp. Applic Walsh Functions, Washington, D.C., pp. 68-73, 1971.

61. H.E. Chrestenson, "A class of generalized Walsh functions", Pacific J. Math., Vol. 5, pp. 17-31, 1955.

62. N.J. Vilenkin, "On a class of complete orthogonal systems", in English Translation of R.P. Bras, Amer. Math. Soc. Translations, Vol. 28, pp. 1-35, 1963.

63. N. Ahmed, H.H. Schreiber and P.V. Lopresti, "On notation and definition of terms related to a class of complete orthogonal functions", IEEE Trans. Electromagnetic Compat., Vol. EMC-15, pp. 75-80, 1973.

64. B.L.N. Kennet, "A note on the finite Walsh transform", IEEE Trans. Inf. Theory, Vol. IT-16, pp. 489-491, 1970.

65. I. Kluvanec, "Sampling theorem in abstract harmonic analysis", Mat. Fyz. Cas. Slov. Acad., Vol. 15, pp. 43-48, 1965.

66. P.J. Nicholson, "Algebraic theory of finite Fourier transforms", J. Comput. and Syst. Science, Vol. 5, pp. 524-547, 1971.

67. G.F. Simmons, *An Introduction to Topology and Modern Analysis*. New York: McGraw Hill, 1963.

68. J.E. Gibbs, B. Ireland and N.J. Smith, "Generalizations of Bolzano's theorem and of Lagrange's mean-value theorem related to the discrete Walsh transform", Proc. 1973 Symp. Th. and Applic. Walsh and Other Non-sinus Functions, Hatfield, England, 1973.

69. J.E. Gibbs, "Sine waves and Walsh waves in physics", Proc. 1970 Symp. Applic. Walsh Functions, Washington, D.C., pp. 260-274, 1970.

70. J.E. Gibbs and B. Ireland, "Some generalizations of the logical derivative", DES Rept. No. 8, NPL, Teddington, England, 1971.

71. J.E. Gibbs, B. Ireland and D.R. Perkins, "A generalization of the dyadic expansion", DES Rept. No. 10, NPL, Teddington, England, 1971.

72. J.E. Gibbs, B. Ireland and J.E. Marshal, "A generalization of the Gibbs differentiator", Proc. 1973 Symp. Theory and Applic. of Walsh and Other Non-sinus Functions, Hatfield, England, 1973.

73. J.E. Gibbs and B. Ireland, "Walsh functions and differentiation", in *Applications of Walsh Functions and Sequency Theory*, Eds., H. Schreiber and G.F. Sandy, pp. 147-176, New York: IEEE Press, 1974.

74. F. Pichler, "Walsh functions, signals and systems", in *Applications of Walsh Functions and Sequency Theory*, Eds., H. Schreiber and G.F. Sandy, pp. 177-190, New York: IEEE Press, 1974.

75. C.K. Yuen, "Remarks on the ordering of Walsh functions", IEEE Trans. Computers, Vol. C-21, pp. 1452, 1972.

76. C.K. Yuen, "Function approximation by Walsh series", IEEE Trans. Computers, Vol. C-24, pp. 590-598, 1975.

77. S. Morgenthaler, "On Walsh-Fourier series", Trans. Amer. Math. Soc., Vol. 84, pp. 472-507, 1957.

78. G. Selfridge, "Generalized Walsh transform", Pacific J. Math., Vol. 5, pp. 451-480, 1955.

79. J.E. Gibbs, "Differentiation and frequency on the dyadic group", DES Rept. No. 8, NPL, Teddington, England, 1973.

80. P.L. Butzer and H.J. Wagner, "Walsh-Fourier series and the concept of a derivative", Applic. Analysis, Vol. 3, pp. 29-46, 1973.

81. N.J. Fine, "On groups of orthonormal Walsh functions (II)", Pacific J. Math., Vol. 5, pp. 61-65, 1955.

82. F. Pichler, *Das System der sal- und cal-Funktionen als erweiterung das systems der Walsh-funktionen und die Theorie der salund cal-Fourier Transformation.* Dissertation, Innsbruck, Univ., 1967.

83. H.C. Andrews and H.L. Caspari, "A generalized technique for spectral analysis", IEEE Trans. Comput., Vol. C-19, pp. 16-25, 1970.

84. H.C. Andrews and J. Kane, "Kronecker matrices, computer implementation and generalized spectra", J.A.C.M., Vol. 17, pp. 260-268, 1970.

85. N. Ahmed, K.R. Rao and R.B. Schultz, "A generalized discrete transform", Proc. IEEE, Vol. 59, pp. 1360-1362, 1971.

86. H.A. Helm, "Group codes and Walsh functions", Proc. 1971 Symp. Applic. Walsh Functions, Washington, D.C., pp. 78-83, 1971.

87. D.F. Elliot, "A class of generalized continuous orthogonal transforms", IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-22, pp. 245-254, 1974.

88. H.F. Harmuth, *Transmission of Information by Orthogonal Functions.* New York: Springer Verlag, 1972.

89. R.B. Crittenden, "On Hadamard matrices and complete orthonormal systems", Proc. 1973 Symp. Applic. Walsh Functions, Washington, D.C., pp. 16-18, 1973.

90. S.C. Kak, "On matrices with Walsh functions as the eigenvectors", Proc. 1972 Symp. Applic. Walsh Functions, Washington, D.C., pp. 27-29, 1972.

91. F.R. Pichler, "Walsh functions and linear system theory", Proc. 1970 Symp. Applic. Walsh Functions, Washington, D.C., pp. 175-182, 1970.

92. F.R. Pichler, "On state-space description of linear dyadic invariant systems", Proc. 1971 Symp. Applic. Walsh Functions, Washington, D.C., pp. 166-170, 1971.

93. J.E. Gibbs and J. Simpson, "Differentiation in finite abelian groups", DES Rept. No. 14, NPL, Teddington, England, 1974.

94. J.E. Gibbs, "Discrete complex Walsh functions", Proc. 1970 Symp. Applic. Walsh Functions, Washington, D.C., pp. 106-122, 1970.

95. T. Ito, "A note on a general expansion of functions of binary variables", Inf. and Control, Vol. 12, pp. 206-211, 1968.

96. T. Ito, "Application of Walsh functions to pattern recognition, and switching theory", Proc. 1970 Symp. Applic. Walsh Functions, Washington, D.C., pp. 128-137, 1970.

97. P.L. Butzer and H.J. Wagner, "A calculus for Walsh functions defined on R", Proc. 1973 Symp. Applic. Walsh Functions, Washington, D.C., pp. 75-81, 1973.

98. P.L. Butzer and H.J. Wagner, "On dyadic analysis based on the pointwise dyadic derivative", Analysis Mathematica, Vol. 1, pp. 171-196, 1975.

99. W. Rudin, *Functional Analysis*. New York: McGraw Hill, 1973.

100. I. Kaplansky, *An Introduction to Differential Algebra*. Paris: Hermann, 1957.

101. S. Cohn-Sfetcu, "On the relationship between Gibbs and Ritt-Kolchin differential operators", Proc. 1975 Symp. Theory and Applic. of Walsh and Other Non-sinus Functions, Hatfield, England, 1975.

102. J.F. Ritt, *Differential Algebra*. Amer. Math. Soc. Coll. Publ., Vol. 33, New York, 1950.

103. E.R. Kolchin, *Differential Algebra and Algebraic Groups*. New York: Academic Press, 1973.

104. J.E. Gibbs and M.J. Millard, "Some methods of solution of ordinary logical differential equations", DES Rept. No. 2, NPL, Teddington, England, 1969.

105. R.C. Hook, "On multi-variable discrete dyadic invariant systems", Proc. 1973 Symp. Theory and Applic. Walsh and Other Non-sinus Functions, Hatfield, England, 1973.

106. S. Cohn-Sfetcu, "On the theory of linear dyadic invariant systems", Proc. 1975 Symp. Theory and Applic. Walsh and Other Non-sinus Functions, Hatfield, England, 1975.

107. J.E. Gibbs, "The concept of frequency in dyadic analysis", Proc. 1975 Int. Symp. on Electromagnetic Compat., Wantreux, Switzerland, 1975.

108. A. Thayse and M. Davio, "Boolean differential calculus and its applications to switching theory", IEEE Trans. Comput., Vol. C-22, pp. 409-420, 1973.

109. J. Pearl, "Applications of Walsh transform to statistical analysis", IEEE Trans. System, Man and Cybernetics, Vol. SMC-1, pp. 111-119, 1971.

110. P.L. Butzer and H.J. Wagner, "A new calculus for Walsh functions with applications", Proc. 1973 Symp. Theory Applic. Walsh and Other Non-sinus. Functions, Hatfield, England, 1973.

111. C.K. Yuen, "Upper bounds on Walsh transforms", IEEE Trans. Comput., Vol. C-24, pp. 1273-1280, 1972.

112. M.K. Sain, "The growing algebraic presence in systems engineering: an introduction", Proc. IEEE, Vol. 64, pp. 96-111, 1976.

113. F.R. Pichler, J.E. Marshal and J.E. Gibbs, "A system theory approach to electrical measurements", Proc. 1973 Symp. Theory and Applic. of Walsh and Other Non-sinus. Functions, Hatfield, England, 1973.

114. J.H. Rosenbloom, "Physical interpretation of the dyadic group", Proc. 1971 Symp. Applic. Walsh Functions, Washington, D.C., pp. 158-165, 1971.

115. H.F. Harmuth, "Linear numbers versus dyadic groups as basis for models of time and space", unpublished report, 1973.

116. P.J. Orava, "Causality and state concepts in dynamical system theory", Int. J. Systems Sci., Vol. 4, pp. 679-691, 1974.

117. L.A. Zadeh and E. Polak, *System Theory*. New York: McGraw Hill, 1969.

118. L. Padulo and M.A. Arbib, *System Theory: A Unified State-Space Approach to Continuous and Discrete Systems*. Philadelphia, Pa.: W.B. Saunders, 1974.

119. P.L. Falb and M.I. Friedman, "A generalized transform theory for causal operators", SIAM J. Control, Vol. 7, pp. 452-471, 1969.

120. R.W. Brocket and A.S. Willsky, "Finite group homomorphic sequential systems", IEEE Trans. Automatic Control, Vol. AC-17, pp. 483-490, 1972.

121. R.E. Kalman, P.L. Falb and M.A. Arbib, *Topics in Mathematical System Theory*. New York: McGraw Hill, 1969.

122. F.R. Pichler, "Walsh functions and optimal linear systems", Proc. 1970 Symp. Applic. Walsh Functions, Washington, D.C., pp. 17-22, 1970.

123. A.W. Nailor, "Generalized frequency concepts of time-varying, discrete-time linear systems", IEEE Trans. Circ. Theory, Vol. CT-10, pp. 428-444, 1963.

124. K.B.K. La Barre, *A Transfer Technique for Linear Time-Varying Discrete Time Systems*. Dissertation, The University of Michigan, 1969.

125. C.K. Yuen, "On dyadic invariant systems", Proc. of IEEE, Vol. 1, pp. 166-169, 1973.

126. D.K. Cheng and J.J. Liu, "Walsh transform analysis of discrete linear systems", Proc. 1973 Symp. Applic. Walsh Functions, Washington, D.C., pp. 61-65, 1973.

127. D.K. Cheng and J.J. Liu, "Time domain analysis of dyadic invariant systems", Proc. IEEE, Vol. 62, pp. 1038-1040, 1974.

128. H. Gethoffer, "Convolution and deconvolution with Walsh functions", Proc. 1971 Symp. Theory and Applic. Walsh and Other Non-sinus Functions, Hatfield, England, 1971.

129. H. Gethoffer, "Algebraic theory of generalized convolution on cyclic groups", Proc. 1973 Symp. Theory and Applic. of Walsh and Other Non-sinus Functions, Hatfield, England, 1973.

130. C.T. LeDinh, P. Le and R. Goulet, "Sampling expansions in discrete and finite Walsh-Fourier analysis", Proc. 1972 Symp. Applic. Walsh Functions, Washington, D.C., pp. 265-271, 1972.

131. C.T. LeDinh, P. Le and R. Goulet, "An analysis of linear dyadic systems with MBL outputs", Proc. 1973 Symp. Applic. Walsh Functions, Washington, D.C., 1973.

132. P.A. Morettin, "Stochastic dyadic systems", Proc. 1973 Symp. Applic. Walsh Functions, Washington, D.C., pp. 290-293, 1973.

133. P.A. Morettin, "Estimation of the Walsh spectrum", IEEE Trans. Inf. Theory, Vol. IT-22, pp. 106-107, 1976.

134. J. Pearl, "Walsh processing of random signals", Proc. 1971 Symp. Applic. Walsh Functions, Washington, D.C., pp. 137-141, 1971.

135. S. Cohn-Sfetcu and S.T. Nichols, "On the identification of linear dyadic invariant systems", IEEE Trans. Electromagnetic Compat., Vol. EMC-17, pp. 111-117, 1975.

136. H.F. Harmuth, J. Kamol and S.S.R. Murthy, "Two dimensional spatial hardware filters for acoustic imaging", in *Applications of Walsh Functions and Sequency Theory*, Eds., H. Schreiber and G.F. Sandy, New York: IEEE Press, 1974.

137. R. Seviora, *Generalized Filtering*. Ph.D. Thesis, University of Toronto, Toronto, Ont., 1971.

138. W.A. Porter, *Modern Foundations of Systems Engineering*. New York: MacMillan, 1966.

139. L.A. Zadeh and C.A. Desoer, *Linear System Theory: The State-Space Approach*. New York: McGraw Hill, 1963.

140. A. Figa-Talamanca, "Translation invariant operators in L^p", Duke Math. J., Vol. 32, pp. 495-501, 1965.

141. J. Pearl, "Optimal dyadic models of time invariant systems", IEEE Trans. Comput., Vol. C-24, pp. 598-603, 1975.

142. P.V. Lopresti and H.L. Suri, "A fast algorithm for the estimation of autocorrelation functions", IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-22, pp. 449-453, 1974.

143. N. Ahmed and T. Natarajan, "On logical and arithmetic autocorrelation functions", IEEE Trans. Electromagnetic Compat., Vol. EMC-16, pp. 177-182, 1974.

144. F.E. Weiser, *Walsh Function Analysis of Instantaneous Non-Linear Stochastic Problems*. Dissertation, Polytechnic Inst. of Brooklyn, New York, 1964.

145. M. Maqusi, "On moments and Walsh characteristic functions", IEEE Trans. Communic., Vol. COMM-21, pp. 768-770, 1973.

146. J.E. Gibbs and F.R. Pichler, "Comments on transformations of Fourier power spectra into Walsh power spectra", Proc. 1971 Symp. Applic. Walsh Functions, Washington, D.C., pp. 51-54, 1971.

147. G.S. Robinson, "Logical convolution and discrete Walsh and Fourier power spectra", IEEE Trans. Audio and Electroacoust., Vol. AU-20, pp. 271-280, 1972.

148. B. Sawicki, "Estimation of power spectra of random processes by the use of Haar functions", Found. Contr. Eng., Vol. 1, pp. 47-58, 1975.

149. A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw Hill, 1965.

150. J.S. Bendat and A.G. Piersol, *Random Data Analysis and Measurement Procedures*. New York: Wiley Interscience, 1971.

151. K.I. Astrom and P. Eykhoff, "System identification - a survey", Automatica, Vol. 7, pp. 123-162, 1971.

152. A.P. Sage and J.L. Melsa, *System Identification*. New York: Academic Press, 1971.

153. December 1974 Issue of IEEE Trans. on Automatic Control.

154. R.M. de Santis, "Causality theory and system analysis", Proc. IEEE, Vol. 64, pp. 36-44, 1976.

155. G.W. Mackey, "The Laplace transform for locally compact abelian groups", Proc. Natl. Acad. Sci., Vol. 34, pp. 156-162, 1946.

156. G.W. Mackey, "Functions on locally compact groups", Bull. Amer. Math. Soc., Vol. 56, pp. 385-412, 1950.

157. R.E. Crochiere and A.V. Oppenheim, "Analysis of linear digital networks", Proc. IEEE, Vol. 63, pp. 581-595, 1975.

158. D. Slepian, "On bandwidth", Proc. IEEE, Vol. 64, pp. 292-300, 1976.

159. C.M. Rader, "A note on exact discrete Fourier transforms", IEEE Trans. Audio and Electroacoust., Vol. AU-21, pp. 558-559, 1973.

160. S. Cohn-Sfetcu, "On the exact evaluation of discrete Fourier transforms", IEEE Trans. Acoust., Speech and Signal Processing, Vol. ASSP-23, pp. 585-586, 1975.

161. B. Gold and C.M. Rader, *Digital Processing of Signals*. New York: McGraw Hill, 1969.

162. H.C. Andrews, "Degrees of freedom and computation requirements in matrix multiplications for Hadamard and other transforms", Proc. 1970 Symp. Applic. Walsh Functions, Washington, D.C., pp. 73-79, 1970.

163. L.R. Welch, "Computation of finite Fourier series", Space Programs Summary No. 37-37, Vol. IV, pp. 295-297 (Jet Propulsion Lab., Pasadena, Cal.), 1966.

164. H.D. Helms, J.F. Kaiser and L.R. Rabiner, Eds., *Literature on Digital Signal Processing; Author and Permuted Title Index*. New York: IEEE Press, 1970.

165. J.N. Bramhall, *An Annotated Bibliography on Walsh and Walsh Related Functions*. Tech. Memorandum TG 11 98B, John Hopkins Univ., Baltimore, Maryland, 1974.

166. R.R. Green, "A serial orthogonal decoder", Space Programs Summary, No. 37-37, Vol. IV, pp. 247-251 (Jet Propulsion Lab., Pasadena, Cal.), 1966.

167. B.F. Vandivere and R.L. Carrick, "Fast Walsh transform", Tech. Rept. TAM 212, Telcom Inc., MacLean, Va., 1967.

168. J.E. Welchel and D.F. Guinn, "The fast Fourier-Hadamard transform and its use in signal representation and classification", Tech. Rept. PRC 68-11, Melpar Inc., Falls Church, Va., 1968.

169. J.L. Shanks, "Computation of the fast Walsh-Fourier transform", IEEE Trans. Comput., Vol. C-18, pp. 457-459, 1969.

170. L.J. Ulman, "Computation of the Hadamard transform and the R transform in ordered form", IEEE Trans. Comput., Vol. C-19, pp. 359-360, 1970.

171. Y.Y. Shum and C.R. Elliott, "Computation of the fast Hadamard transform", Proc. 1972 Symp. Applic. Walsh Functions, Washington, D.C., pp. 177-180, 1972.

172. J.W. Manz, "A sequency ordered fast Walsh transform", IEEE Trans. Audio and Electroacoustics, Vol. AU-20, pp. 204-205, 1972.

173. H. Hubner, "Sequency ordered Walsh functions", Communic. of A.C.M., Vol. 13, pp. 511, 1970.

174. L.R. Rabiner and C.M. Rader, Eds., *Digital Signal Processing*. New York: McGraw Hill, 1972.

175. D.V. James, "Quantization errors in the fast Fourier transform", IEEE Trans. Acoust., Speech and Signal Processing, Vol. ASSP-23, pp. 277-283, 1975.

176. J. Pearl, "The Hilbert-Schmidt decorrelation measure for Fourier and Walsh transforms", Proc. 1972 Symp. Applic. Walsh Functions, Washington, D.C., pp. 293-297, 1972.

177. H.F. Harmuth, E. Briganti, J.E. Gibbs, H. Hubner, S.C. Kak, R. Lopez de Zavalia, F. Pichler, G.R. Redinbo, Y. Taki, and C.K. Yuen, "Comments on 'Sinusoids versus Walsh functions' ", IEEE Trans. Electromagnetic Compat., Vol. EMC-17, p. 194, 1975.

178. J. Pearl, "Basis-restricted transformations and performance measures for spectral representations", IEEE Trans. Inf. Theory, Vol. IT-17, pp. 751-752, 1971.

179. J. Pearl, H.C. Andrews and W.K. Pratt, "Performance measures for transform data coding", IEEE Trans. Communic., Vol. COMM-20, pp. 411-415, 1972.

180. W.K. Pratt, "Generalized Wiener filtering computing techniques", IEEE Trans. Comput., Vol. C-21, pp. 630-642, 1972.

181. R.J. Lechner, "Harmonical analysis of switching functions", in *Recent Developments in Switching Theory*, Ed., A. Mukhopadhyay, pp. 122-228, New York: Academic Press, 1974.

182. E.C. Posner, "Combinatorial structures in planetary reconnaisance", in *Error Correcting Codes*, J.B. Mann, Ed., pp. 15-46, New York: J. Wiley and Sons, 1968.

183. M. Koenig and J.P. Zolesio, "Problemes numerique relatifs aux fonctions de Walsh", C.R. Acad. Sci. Paris, Vol. 271, pp. 1074-1077, 1970.

184. N. Ahmed and K. Rao, *Orthogonal Transforms for Digital Signal Processing*. New York: Springer Verlag, 1975.

185. J. Lukasiewicz, *Selected Works*. Ed., I.L. Borkowski, Amsterdam: North Holland, 1970.

186. I.S. Reed, "A class of multiple error-correcting codes and the decoding scheme", Trans. IRE, PGIT-4, pp. 38-49, 1954.

187. D.E. Mueller, "Application of Boolean algebra to switching circuit design and to error detection", IRE Trans. Electronic Comput., Vol. EC-3, pp. 6-12, 1954.

188. W.W. Peterson and E.J. Weldon, Jr., *Error Correcting Codes*. Cambridge, Mass.: MIT Press, 1972.

189. I. Ninomya, "A theory of the coordinate representation of switching functions", IEEE Trans. Electronic Comput., Vol. EC-12, pp. 152, 1963.

190. R.J. Lechner, "Transformations among switching function canonical forms", IEEE Trans. Electronic Comput., Vol. EC-12, pp. 129-130, 1963.

191. R.J. Lechner, "Invariant properties of Hadamard transforms under affine groups", Proc. 1970 Symp. Applic. Walsh Functions, Washington, D.C., pp. 88-94, 1970.

192. S.L. Hurst, "The application of Chow parameters and Rademacher-Walsh matrices in the synthesis of binary functions", Comput. J., Vol. 16, pp. 165-173, 1973.

193. C.R. Edwards, "The application of Rademacher-Walsh transforms to Boolean function classification and threshold logic synthesis", IEEE Trans. Comput., Vol. C-24, pp. 48-62, 1975.

194. M. Davio, A. Thayse and G. Bioul, "Symbolic computation of Fourier transforms of Boolean functions", Philips Res. Repts., Vol. 27, pp. 386-403, 1972.

195. J.P. Deschamps and A. Thayse, "On a theory of discrete functions, Part I: The lattice structure of discrete functions", Philips Res. Repts., Vol. 28, pp. 397-423, 1973.

196. A. Thayse and J.P. Deschamps, "On a theory of discrete functions, Part II: The ring and field structures of discrete functions", Philips Res. Repts., Vol. 28, pp. 424-465, 1973.

197. J.P. Deschamps and A. Thayse, "Applications of discrete functions, Part I: Transient analysis of combinational networks", Philips Res. Repts., Vol. 28, pp. 497-529, 1973.

198. A. Thayse and J.P. Deschamps, "Applications of discrete functions, Part II: Analysis of asynchronous switching networks", Philips Res. Repts., Vol. 29, pp. 429-452, 1974.

199. S.W. Golomb, "On the classification of Boolean functions", IRE Trans. Circ. Theory, Vol. CT-6, pp. 176-186, 1959.

200. M.L. Dertouzos, *Threshold Logic: A Synthesis Approach*. Cambridge, Mass.: MIT Press, 1965.

201. S.L. Hurst, "Digital summation threshold logic gates: a new circuit element", Proc. IEE (London), Vol. 120, pp. 1301-1307, 1973.

202. B.T. Polyak and Yu. A. Schreider, "The application of Walsh functions in approximate calculus", Voprosy Teor. Matem. Mashin., Vol. II, pp. 174-190, 1962. In English translation of J.E. Gibbs.

203. R. Okkes and I. Pryke, "The compression of particle data count by use of the fast Hadamard transform", Proc. 1975 Symp. Theory and Applic. Walsh and Other Non-sinus. Functions, Hatfield, England, 1975.

204. C.K. Yuen, "Analysis of Walsh transform using integration by parts", SIAM J. Math. Anal., Vol. 4, pp. 574-584, 1973.

205. H.C. Andrews, *Computer Techniques in Image Processing*. New York: Academic Press, 1972.

206. H.F. Harmuth and R. de Buda, "Conversion of sequence-limited signals into frequency-limited signals and vice-versa", IEEE Trans. Inf. Theory, Vol. IT-17, pp. 343-344, 1971.

207. K.H. Siemens and R. Kitai, "Digital Walsh-Fourier analysis of periodic waveforms", IEEE Trans. on Instrum. and Measurement, Vol. IM-18, pp. 316-321, 1969.

208. K.H. Siemens and R. Kitai, "Walsh series to Fourier series conversion", Proc. 1972 Symp. Applic. Walsh Functions, Washington, D.C., pp. 295-297, 1972.

209. K.H. Siemens and R. Kitai, "A non-recursive equation for the Fourier transform of a Walsh function", IEEE Trans. Electromagnetic Compat., Vol. EMC-15, pp. 81-83, 1973.

210. R. Kitai, "Walsh to Fourier spectral conversion for periodic waveforms", IEEE Trans. Electromagnetic Compat., Vol. EMC-17, pp. 266-269, 1975.

211. R. Kitai, "Synthesis of periodic sinusoids from Walsh waves", IEEE Trans. Instrum. and Measurement, Vol. IM-24, pp. 313-317, 1975.

212. W.M. Walmsley and W.A. Evans, "The simulation and design of a frequency response analyzer using intermediate domain processing", Proc. 1975 Symp. Theory and Applic. Walsh and Other Non-sinus. Functions, Hatfield, England, 1975.

213. Y. Emsellem and G. de Marsily, "Restitution automatique des permeabilites d'une nappe, le probleme inverse et la deconvolution", La Houille Blanche, No. 8, pp. 861-868, 1969.

214. Y. Emsellem and G. de Marsily, "An automatic solution of the inverse problem", Water Resources, Vol. 7, pp. 1264-1283, 1971.

215. G. de Marsily and D. Poitrinal, "A deconvolution algorithm by projection on Walsh functions", Proc. 1975 Symp. Theory and Applic. Walsh and Other Non-sinus. Functions, Hatfield, England, 1975.

216. V.A. Romanov and V.A. Semeran, "Algorithm for identifying the dynamic characteristics of objects by means of orthogonal Walsh functions", Automatika and Telemekhanika, No. 4, pp. 110-117, 1973.

217. C.F. Chen and C.H. Hsiao, "Walsh series analysis in optimal control", Int. J. Contr., Vol. 21, pp. 881-897, 1975.

218. C.F. Chen and C.H. Hsiao, "Design of piecewise constant gains for optimal control via Walsh functions", IEEE Trans. Automatic Control, Vol. AC-20, pp. 596-603, 1975.

219. C.F. Chen, Y.T. Tsay and T.T. Wu, "Walsh operational matrices for fractional calculus and their applications to distributed systems", Proc. 1976 IEEE Int. Symp. Circ. and Syst., Munich, 1976.

220. N. Gopalsami and B.L. Deekskatulu, "Comments on Design of piecewise constant gains for optimal control via Walsh functions", IEEE Trans. Automatic Control, Vol. AC-21, pp. 634-635, 1976. With author's reply (C.F. Chen), see above, pp. 635-636.

221. H. Burkhardt, "Realization of optimal control algorithms in the sequency domain of the Walsh transformation", Proc. 1975 Int. Symp. on Electromagnetic Compat., Montreux, Switzerland, 1976.

222. H. Burkhardt, "The use of Walsh functions in optimal control theory", Proc. 1975 Symp. Theory and Applic. Walsh and Other Non-sinus. Functions, Hatfield, England, 1975.

223. H.C. Andrews and C.L. Patterson, "Outer product expansions and their uses in digital signal processing", IEEE Trans. Comput., Vol. C-25, pp. 140-148, 1976.

224. A.G. Tescher and H.C. Andrews, "The role of adaptive phase coding in two- and three-dimensional Fourier and Walsh image compression", in *Applications of Walsh Functions and Sequency Theory*, Eds., H. Schreiber and G.F. Sandy, pp. 26-65, New York: IEEE Press, 1974.

225. P. Noll, "Adapative orthogonal transform coding system for speech signals", Proc. 1976 IEEE Int. Symp. on Information Theory, Ronneby, Sweden, 1976.

226. T.S. Huang, J.W. Burnett and A.G. Deczky, "The importance of phase in image processing filtering", Proc. 1976 IEEE Int. Symp. Circ. and Systems, Munich, 1976.

227. K. Shibata, "Block waveform coding of image signals by orthogonal transformation", Proc. 1973 Symp. Applic. Walsh Functions, Washington, D.C., pp. 137-143, 1973.

228. S.C. Knauer, "Real-time video compression algorithm for Hadamard transform processing", IEEE Trans. Electromagnetic Compat., Vol. EMC-18, pp. 28-36, 1976.

229. A.R. Elliott, "Voice processing using Walsh-Hadamard functions", in *Applications of Walsh Functions and Sequency Theory*, Eds., H. Schreiber and G.F. Sandy, pp. 77-93, New York: IEEE Press, 1974.

230. H.H. Schreiber, "Theory and applications of sequency multiplexing", in *Applications of Walsh Functions and Sequency Theory*, Eds., H. Schreiber and G.F. Sandy, pp. 316-344, New York: IEEE Press, 1974.

231. R. Barett, "The use of Walsh and other non-sinusoidal functions in communications signal multiplexing", in *Applications of Walsh Functions and Sequency Theory*, Eds., H. Schreiber and G.F. Sandy, New York: IEEE Press, 1974.

232. H. Hubner, "Signal multiplexing in the correlation and matrix domain", Proc. 1975 Symp. Theory and Applic. Walsh and Other Non-sinus. Functions, Hatfield, England, 1975.

233. H.C. Andrews, *Picture Processing and Digital Filtering*. Berlin: Springer Verlag, 1975.

234. J.W. Carl, "On the use of Walsh functions in man-made and biological pattern recognition systems", in *Applications of Walsh Functions and Sequency Theory*, Eds., H. Schreiber and G.F. Sandy, pp. 9-25, New York: IEEE Press, 1974.

235. D.A. Linkens and Z.B. Temel, "The use of Walsh transforms in the analysis of gastro intestinal signals", Proc. 1975 Symp. Theory and Applic. Walsh and Other Non-sinus. Functions, Hatfield, England, 1975.

236. L. Kirvida, "Texture measurement for the automatic classification of imagery", IEEE Trans. Electromagnetic Compat., Vol. EMC-18, pp. 38-42, 1976.

237. A.S. French and E.G. Butz, "The use of Walsh functions in the Wiener analysis of non-linear systems", IEEE Trans. Comput., Vol. C-23, pp. 225-231, 1974.

238. M.S. Corington, "Solution of differential and integral equations with Walsh functions", IEEE Trans. Circ. Theory, Vol. CT-20, pp. 470-476, 1973.

239. H.F. Harmuth, Y.K. Hong and J. Kamal, "Electromagnetic Walsh waves: transmitters, receivers and their applications", in *Applications of Walsh Functions and Sequency Theory*, Eds., H. Schreiber and G.F. Sandy, pp. 275-296, New York: IEEE Press, 1974.

240. H.F. Harmuth, "Non-sinusoidal electromagnetic waves as carriers in high performance radar", Proc. 1975 Symp. Theory and Applic. Walsh and Other Non-Sinus. Functions, Hatfield, England, 1975.

241. H.F. Harmuth, "Antennas for non-sinusoidal electromagnetic waves", Proc. IEEE 1976 Int. Symp. on Electromagnetic Compat., Washington, D.C., pp. 254-257, 1976.

242. H.D. Helms, "Fast Fourier transform method for computing difference equations and simulating filters", IEEE Trans. Audio and Electroacoust., Vol. AU-15, pp. 85-90, 1967.

243. G.D. Bergland, "A guided tour of the fast Fourier transform", IEEE Spectrum, Vol. 6, pp. 41-52, 1969.

244. E.I. Jury and Ya. Z. Tsypkin, "On the theory of discrete systems", Automatica, Vol. 7, pp. 89-107, 1971.

245. J.W. Cooley, P.A. Lewis and P.D. Welch, "The fast Fourier transform algorithm and its applications", IBM Res. Paper RC-1743, 1967.

246. I. Stewart, *Galois Theory*. London: Chapman and Hall, 1973.

247. A. Weil, "Number of solutions of equations in a finite field", Bull. Amer. Math. Soc., Vol. 55, pp. 497-508, 1949.

248. K. Ireland and M.I. Rosen, *Elements of Number Theory: Including an Introduction to Equations over Finite Fields*. Tarrytown-on-Hudson, N.Y.: Bogden and Quigley, 1972.

249. D.E. Knuth, *The Art of Computer Programming, Vol. 2: Seminumerical Algorithms*. Reading, Mass.: Addison-Wesley, 1969.

250. I.J. Good, "The relation between two fast Fourier transforms", IEEE Trans. Comput., Vol. C-20, pp. 310-317, 1971.

251. J.M. Pollard, "The fast Fourier transform in a finite field", Math. Comput., Vol. 25, pp. 365-374, 1971.

252. R.C. Agarwal and C.S. Burrus, "Fast convolution using Fermat number transforms with applications to digital filtering", IEEE Trans. Acoust. Speech and Signal Processing, Vol. ASSP-22, pp. 89-97, 1974.

253. R.C. Agarwal and C.S. Burrus, "Fast one-dimensional digital convolution by multi-dimensional techniques", IEEE Trans. Acoust., Speech and Signal Processing, Vol. ASSP-22, pp. 1-10, 1974.

254. J.D. Bruie, "Fast convolutions with finite field transforms", IEEE Trans. Acoust., Speech and Signal Processing, Vol. ASSP-23, p. 240, 1975.

255. I.S. Reed and T.K. Truong, "The use of finite fields to compute convolutions", IEEE Trans. Inf. Theory, Vol. IT-21, pp. 208-213, 1975.

256. I.S. Reed and T.K. Truong, "Complex integer convolutions over a direct sum of Galois fields", IEEE Trans. Inf. Theory, Vol. IT-21, pp. 657-661, 1975.

257. J. Justesson, "On the complexity of decoding Reed-Solomon codes", IEEE Trans. Inf. Theory, Vol. IT-22, pp. 237-248, 1976.

258. R.G. Gallager, *Information Theory and Reliable Communications*. New York: Wiley, 1968.

259. C.Y. Lee, "Some properties of non-binary error-correcting codes", IRE Trans. Inf. Theory, Vol. IT-4, pp. 77-82, 1958.

260. J.L. Massey and M.K. Sain, "Codes, automata, and continuous systems: explicit interconnections", IEEE Trans. Automatic Control, Vol. AC-12, pp. 644-650, 1967.

261. D.A. Huffman, "A linear circuit view-point on error correcting codes", IRE Trans. Inf. Theory, Vol. IT-2, pp. 20-28, 1956.

262. J. Richalet, "Operational calculus for finite rings", IEEE Trans. Circ. Theory, Vol. CT-12, pp. 558-570, 1965.

263. R.G. Faradzhev, "Analytical procedures for linear sequential machines", Eng. Cyber., Vol. 5, pp. 78-86, 1965.

264. M.B. Nathanson, "Derivatives of binary sequencies", SIAM J. Appl. Math., Vol. 21, pp. 407-412, 1971.

265. M.B. Nathanson, "Difference operators and periodic sequences over finite modules", Acta Math. Acad. Sci. Hungar., to appear.

266. S.B. Akers, "On the theory of Boolean functions", SIAM J. Appl. Math., Vol. 7, pp. 487-498, 1959.

267. A. Thayse, "Boolean differential calculus", Philips Res. Repts., Vol. 26, pp. 229-246, 1971.

268. I.S. Reed, "Boolean difference calculus and fault finding", SIAM J. Appl. Math., Vol. 24, pp. 134-143, 1973.

269. S. Rudeanu, *Boolean Functions and Equations*. Amsterdam: North Holland, 1974.

270. A. Thayse, "Difference operators and extended truth vectors for discrete functions", Discrete Math., Vol. 14, pp. 171-202, 1976.

271. A. Thayse, "Differential calculus for functions from (GF(p)) into GF(p)", Philips Res. Repts., Vol. 29, pp. 560-586, 1974.

272. B. Benjauthrit and I.S. Reed, "Galois switching functions and their applications", IEEE Trans. Comput., Vol. C-25, pp. 78-86, 1976.

273. M.P. Remler, "Pattern recognition by convolution polynomial", IEEE Trans. Comput., Vol. C-23, pp. 528-530, 1974.

274. S.H. Chang and T. Joseph, "On ordering of a class of generalized Walsh functions", Proc. 1972 Symp. Applic. Walsh Functions, Washington, D.C., pp. 337-343, 1972.

275. A.M. Trahtman and V.A. Trahtman, "The frequency of Walsh functions", Radio Eng., Vol. 28, pp. 56-58, 1973.

276. S. Cohn-Sfetcu and S.T. Nicholas, "On the concepts of frequency and sequency", unpublished report.

277. J. Shekel, "Instantaneous frequency", Proc. IRE, Vol. 41, p. 548, 1953.

278. D.A. Linden, "A note concerning instant frequency", Proc. IRE, Vol. 40, p. 1970, 1958.

279. L.V. Ahlfors, *Complex Analysis*. New York: McGraw-Hill, 1966.

280. N. Ahmed and R.B. Schultz, "Position spectrum coefficients", IEEE Trans. Audio and Electro Acoust., Vol. AU-19, pp. 326-327, 1971.

281. L.J. Griffiths, "Rapid measurement of digital instantaneous frequency", IEEE Trans. Acoust., Speech and Signal Processing, Vol. ASSP-23, pp. 207-221, 1975.

282. P.J. Nicholson, *Algebraic Theory of the Finite Fourier Transform*. Doctoral Dissertation, Stanford University, Stanford, Calif., 1969.

283. K.W. Henderson, "Some notes on Walsh functions", IEEE Trans. Electronic Comput., Vol. EC-13, pp. 50-52, 1964.

284. M. Cohn, "Walsh functions, sequency and Gray codes", SIAM J. Appl. Math., Vol. 21, pp. 442-447, 1971.

285. G.H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and non-linear least squares problems whose variables separate", SIAM J. Numer. Anal., Vol. 10, pp. 413-432, 1973.

286. M.R. Osbourne, "Some special non-linear least squares problems", SIAM J. Numer. Anal., Vol. 12, pp. 571-592, 1975.

287. R. Bellman, R. Kalaba and J. Lockot, *Numerical Inversion of the Laplace Transform*. New York: Elsevier, 1966.

288. R. Piessens, "A new numerical method for the inversion of the Laplace transform", J. Inst. Math. Applic., Vol. 10, pp. 185-197, 1972.

289. I.N. Sneddon, *Fourier Transforms*. New York: McGraw Hill, 1951.

290. I.N. Sneddon, *The Use of Integral Transforms*. New York: McGraw Hill, 1972.

291. D.W. Widder, *An Introduction to Transform Theory*. New York: Academic Press, 1971.

292. W.L. Perry, "Finite Mellin convolution equations", SIAM J. Math. Anal., Vol. 4, pp. 238-243, 1973.

293. W.L. Perry, "Inversion of general integral transforms", SIAM J. Math. Anal., Vol. 5, pp. 764-773, 1974.

294. D.G. Gardner, J.C. Gardner, G. Lausch and W.W. Meinke, "Method for the analysis of multicomponent exponential decays", J. Chem. Phys., Vol. 31, pp. 978-986, 1959.

295. J. Schlesinger, "Fit to experimental data with exponential functions using the fast Fourier transform", Nucl. Instrum. and Methods, Vol. 106, pp. 503-508, 1973.

296. M.R. Smith and S. Cohn-Sfetcu, "Comments on 'Fit to experimental data with exponential functions using the fast Fourier transform' ", Nucl. Instrum. and Methods, Vol. 114, pp. 171-173, 1974.

297. S. Cohn-Sfetcu, M.R. Smith and S.T. Nichols, "On the representation of signals by basis kernels with product argument", Proc. IEEE, Vol. 63, pp. 326-327, 1975.

298. G.M. Robbins and T.S. Huang, "Inverse filtering for linear shift-variant imaging systems", Proc. IEEE, Vol. 60, pp. 862-871, 1972.

299. A.V. Oppenheim, "Generalized superposition", Inf. and Control, Vol. 11, pp. 528-536, 1967.

300. A.V. Oppenheim, R.W. Schafer and Th.G. Stockham, "Nonlinear filtering for multiplied and convolved signals", Proc. IEEE, Vol. 56, pp. 1264-1291, 1968.

301. M.P. Ekstrom, "A Monte-Carlo approach to numerical deconvolution", Proc. 1976 IEEE Int. Conf. Acoust., Speech and Signal Processing, Philadelphia, Pa., pp. 372-375, 1976.

302. K.L. Peacock and S. Treitel, "Predictive deconvolution: theory and practice", Geophysics, Vol. 34, pp. 155-169, 1969.

303. D.P. Skinner and D.G. Childers, "Real time composite signal decomposition", IEEE Trans. Acoust., Speech and Signal Processing, Vol. ASSP-24, pp. 267-270, 1976.

304. Th. G. Stockham, T.M. Cannon and R.R. Ingebretsen, "Blind deconvolution through digital signal processing", Proc. IEEE, Vol. 63, pp. 678-692, 1975.

305. E.O. Brigham, *The Fast Fourier Transform*. Englewood Cliffs, N.J.: Prentice Hall, 1974.

306. H.F. Silverman and A.E. Pearson, "On deconvolution using the discrete Fourier transform", IEEE Trans. Audio and Electro Acoust., Vol. AU-21, pp. 112-118, 1973.

307. M.P. Ekstrom, "A spectral characterization of the ill-conditioning in numerical deconvolutions", IEEE Trans. Audio and Electro Acoust., Vol. AU-21, pp. 344-348, 1973.

308. J.E. Welchel, Jr. and M.F. Guinn, "FFT organizations for high speed digital filtering", IEEE Trans. Audio and Electro Acoust., Vol. AU-18, pp. 159-168, 1970.

309. M.R. Smith and S. Cohn-Sfetcu, "Decomposition of multicomponent exponential decays by spectral analytic techniques", to appear in Nov. 1976 issue of Technometrics.

310. P. Mancini and A. Pilo, "A computer program for multiexponential fitting by the pealing method", Comput. Biomed. Res., Vol. 3, pp. 1-14, 1970.

311. C. Lanczos, *Applied Analysis*. Englewood Cliffs, N.J.: Prentice Hall, 1956.

312. A. Lemaitor and J. Melange, "An efficient method for multi-exponential fitting with a computer", Comput. Biomed. Res., Vol. 4, pp. 555-560, 1971.

313. R.N. McDonough and W.M. Huggins, "Best least squares representation of signals by exponentials", IEEE Trans. Automatic Control, Vol. AC-13, pp. 408-412, 1968.

314. W.H. Kautz, "The approximation problem", IRE Trans. Circ. Theory, Vol. CT-1, pp. 29-39, 1954.

315. M.R. Smith and S. Cohn-Sfetcu, "On the design and use of digital filters in certain experimental situations", J. Phys. E.: Sci. Instrum, Vol. 8, pp. 1-8, 1975.

316. M.R. Smith and H.A. Buckmaster, "On the analysis of noisy exponential decays", J. Mag. Res., Vol. 17, pp. 29-33, 1975.

317. W.S. Moore and T. Yalcin, "The experimental measurement of exponential time constants in the presence of noise", J. Mag. Res., Vol. 11, pp. 50-57, 1973.

318. M.L. Lion, "Spline fit made easy", IEEE Trans. Comput., Vol. C-25, pp. 522-527, 1976.

319. H.Z. Cummins and E.R. Pike, Eds., *Photon Correlation and Light Beating Spectroscopy*. New York: Plenum, 1974.

320. S. Cohn-Sfetcu, "Linear system estimates via intermediate domain processing", Proc. 14th Annual Allerton Conf. on Circ. and System Theory, Urbana, 1976.

321. P.D. Welch, "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms", IEEE Trans. Audio and Electro Acoust., Vol. AU-15, pp. 70-73, 1967.

322. L.P. Dennis, *Estimation of Frequency Response Functions for Linear Systems Using the FFT*. Ph.D. Thesis, The University of Calgary, Calgary, Alta., 1973.