RELATIONSHIPS BETWEEN DIGITAL SIGNAL PROCESSING AND CONTROL AND ESTIMATION THEORY

Alan S. Willsky*

Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

The purpose of this paper is to explore several current research directions in the fields of digital signal processing and modern control and estimation theory. We examine topics such as stability theory, linear prediction, and parameter identification, system synthesis and implementation, two-dimensional filtering, decentralized control and estimation, and image processing, in order to uncover some of the basic similarities and differences in the goals, techniques, and philosophy of the two disciplines.

Introduction

The writing of this paper was motivated by the belief that the fields
of digital signal processing and control and estimation theory possess
enough similarities <u>and</u> differences in philosophy, goals, and analytical
techniques to merit a detailed joint examination. In order to explore the
relationship between these two fields, I found it essential to concentrate
on several specific research directions to provide a focus for my investiga-
tions. The results of this study were a talk delivered during the 1976 IEEE
Arden House Workshop on Digital Signal Processing, the present paper, and a
far more comprehensive manuscript [Z1].

Although the paper consists of discussions of several specific research
directions, the primary emphasis of this paper is not on results. Rather,
I have been far more interested in understanding the goals of the research
and the methods and approach used by workers in both fields. Understanding
the goals may help us to see why the techniques used in the two disciplines
differ. Inspecting the methods and approaches may allow one to see areas
in which concepts in one field may be usefully applied in the other. In
summary, the primary goal of this study is to provide a basis for future
collaboration among researchers in both fields.

It is hoped that the above comments will help explain the spirit in
which this paper has been written. In reading this paper, the reader may
find many comments that are either partially or totally unsubstantiated.
These points have been included in keeping with the speculative nature of
the study. However, I have attempted to provide background for the specula-
tion and have limited these comments to questions which I feel represent

exciting opportunities for interaction and collaboration. Clearly these issues must be studied at a far deeper level than is possible in this initial effort. To aid others who may wish to follow up on some of the directions developed in this paper, an extensive bibliography has been included. In addition, the interested reader is referred to [Z1] in which all of these research directions are explored in substantially greater breadth and detail.

Nowhere in the paper have I made a direct attempt to define the fields of digital signal processing and control and estimation. Rather, I hope that by examining many of the issues of importance to workers in these fields, the reader will be able to piece together a picture of the disciplines and their relationship to each other. As a preface to our examination, let me mention several points concerning each field.

In digital signal processing, one of the crucial problems is the design of an implementable system meeting certain given design specifications such as an ideal frequency response. Here the emphasis often is on the word implementable, with a fair amount of attention paid to issues such as the structure of the digital filter, its complexity, in terms of architecture and computation time, the effect of finite wordlength on performance, etc. Much of this attention is motivated by the need for extremely efficient systems to perform complex signal processing tasks (e.g., the implementation of high order recursive or nonrecursive filters) at very high data rates (for example, sampling rates encountered in speech processing run on the order of 10 kHz, and higher rates are the rule in video and radar systems).

In control and estimation, the emphasis has been far less on implementation and more on developing methods for determining system design specifi-

cations for estimation or control systems. At one level these specifications are just a particular class of design guidelines which can then be used to construct an implementable digital system. However, there are major differences between the systems arising in the control context and the typical digital processing application. For one thing, the data rates for control systems are often far lower (e.g. in aircraft control systems sampling rates on the order of .1 kHz are often encountered, although much higher rates can be found in certain cases such as video-directed systems). More fundamentally, however, the signal processing to be done in a control system cannot be judged by itself, as can other signal processing systems, since it is part of a feedback loop, and the effect of the processing must be studied in the context of its closed loop effects.

Also, many modern control and estimation techniques involve the use of a state space formulation, as opposed to input-output descriptions which are usually encountered in digital signal processing applications. Some of the reasons for this difference will be made clear in the following sections, but one implication is immediately evident. The use of a state-space description implies that the system under consideration is causal. In standard feedback control problems this is clearly the case, and thus state-space formulations make a great deal of sense. As we'll see, there are digital signal processing problems involving noncausal systems or systems in which the independent variable has nothing to do with time and for which causality has no intrinsic meaning. Thus, while we will find several places in which state space concepts fit in naturally in the digital signal processing context, we'll also find others in which that is decidedly not the case.

The preceding comments were made in order to provide the reader with some insight into the perspective I have taken in writing this paper. With this as background, let us begin our examination of research topics in the two fields.

A. Design, Realization, and Implementation

In this section we investigate one subject area in which some of the differences in perspective between the two disciplines are most apparent. Specifically, we consider the question of design. However, our discussion will not deal very much with design methods but rather with the question of trying to pinpoint what researchers in the two disciplines mean by "design" and what sorts of problems their techniques are equipped to handle.

Perhaps the most obvious difference between the fields is in the type of system representations used. In digital signal processing, the emphasis is heavily on input/output descriptions, while in control and estimation the emphasis is more on state space models. The reasons for this difference stem from the different questions addressed by researchers in the two disciplines. In digital signal processing one is interested in the issue of implementation of a system with a specified input-output behavior (hence the need for an input-output description). Questions such as efficient implementation and number of bits needed to achieve the desired level of accuracy are of great importance.

On the other hand, in control and estimation theory the issue of implementation is not considered to nearly the same extent. Realization techniques do address the question of constructing a state space realization that leads to a specified input-output behavior. However, as discussed in the following subsections, such techniques do not address many of the major issues involved in implementation, and, in fact, state space realizations, when viewed as implementable algorithms, don't include some of the most important system structures that are used in digital system design.

Nevertheless, state space models do play an important role in control and esti-

mation system design. Specifically, a state space model for a given physical

system, is a necessary ingredient in the application of a number of tech-

niques for the analysis of system performance and for the design of feed-

back control or estimation systems (i.e. the specification of the desired

input-output behavior of a control or estimation system).

Thus, we see some fundamental differences between the perspectives of

researchers in the two disciplines. There also clearly exist several areas

for interaction between the fields -- to develop useful multi-input/multi-

output structures (a marriage of digital implementation and multivariable

realization concepts), to utilize state space techniques to analyze the

performance of digital filter structures, and to consider the digital imple-

mentation of state-space control and estimation system designs.

All of these issues are discussed in detail in this section.

A.1 State Space Realizations and State Space Design Techniques

The basic realization problem (for linear systems) is as follows: we are

given a (possibly time-varying) description of the input/output behavior of

a system

$$y(k) = \sum_{i=0}^{k} T(k,i)\, u(i) \tag{A.1}$$

where u and y may both be vectors. In the time-invariant case we have that

the sequence of impulse response matrices satisfies

$$T(k,i) = T(k-i,0) \triangleq T_{k-i} \tag{A.2}$$

and in this case we may be given an alternative input/output description

in the transform domain

$$Y(z) = G(z)U(z) \quad , \quad G(z) = \sum_{i=0}^{\infty} T_i z^{-i} \qquad \text{(A.3)}$$

The realization problem consists of finding a state space model

$$x(k+1) = A(k)x(k) + B(k)u(k)$$
$$y(k) = C(k)x(k) + D(k)u(k) \qquad \text{(A.4)}$$

that yields the desired input/output behavior ((A.1) or (A.3)) when $x(0) = 0$.

The realization problem has been studied in detail in the control literature, and one aspect that has received a great deal of attention is that of determining minimal realizations -- i.e. models as in (A.4) with the dimension of x as small as possible. The basic idea here is that a minimal realization has no superfluous states that either cannot be affected by inputs or do not affect the output. These concepts lead directly to the notions of controllability and observability. In the time-invariant case, one obtains a rather complete description. Specifically, we find that the system (A.3), has a finite-dimensional realization if and only if G(z) is rational with each element having no more zeroes than poles. Furthermore, any controllable and observable time-invariant realization is of minimal dimension, and any such minimal realization can be obtained from a particular one by change of basis (see, for example, [A-4,5,27,30]).

The algorithm of Ho [A-29] and that of Silverman and Meadows [A-5] provide methods for extracting minimal constant realizations from the Hankel matrix determined by the T_i (see Subsection B.3 and the references for

details of these results). Thus, if we are given a design specification or plant description in terms of a rational G(z), we can readily determine a minimal realization. On the other hand, if we are given G in the form (A.3) as opposed to in rational form, _partial_ realization algorithms must be used. We will discuss such algorithms in Subsection B.3.

State space realization algorithms can, in principle, solve certain questions related to system implementation. Specifically, the computation of a minimal realization allows us to determine the minimal number of delays required, in implementing a system with a given impulse or frequency response function, and this clearly has an impact on storage requirements. Also, one of the most important aspects of the state-space approach is that it allows one to consider multiple input/multiple output systems and time-varying systems. Since any minimal state space realization can be obtained from a given one by change of coordinates, clearly realization theory allows some flexibility in designing good digital filter structures. But it is far from the complete answer, as we will see in the next subsection. Not only is memory becoming cheaper (thus reducing the importance of minimality), but there are other implementation issues besides storage that are of importance, and one also runs into limitations in interpreting state space realizations as filter structures (see Subsection A.2).

A more important aspect of state space realizations comes from the fact that they play an extremely important part in a number of control and estimation design problems, where one uses state space realizations to _model_ the system to be controlled or the signals to be filtered. By doing this, one can bring into play extremely powerful state space techniques for compensator design [A-2, 6], decoupling of the effects of different input channels [A-7], etc. and we refer to the reader to the special issue of the IEEE Transactions

on Automatic Control [A-27] for an overview of many design methods that have been developed. These design algorithms allow one to consider a variety of extremely complicated multivariable system problems within a single framework, and this ability to handle many variables at once is at the heart of the value of state space concepts.

One important aspect of some of these techniques is that they allow one to solve quantitative optimization problems. The linear-quadratic optimal control problem is an example of this, as is the design of a Wiener filter as a steady-state Kalman filter [A-8, 27]. In this case, we begin by modelling the observed signal as the additive white noise-corrupted output of a linear state-space model (a shaping filter) driven by white noise. Having solved this realization problem, the determination of the optimal Kalman filter is reduced to solving a time-varying Riccati equation or a nonlinear algebraic Riccati equation for the steady-state (Wiener) filter. Algorithms for solving this algebraic equation essentially solve the Wiener spectral factorization problem.

In addition to providing a framework for the specification of designs, the state space framework allows one to analyze the performance characteristics of the overall system after it has been implemented. For example, the techniques described in Section C can be used to study the stability characteristics of the system. Another analytical tool used to study system performance is covariance analysis. Consider the model

$$x(k+1) = Ax(k) + w(k), \quad y(k) = Cx(k) + v(k) \tag{A.4}$$

where w and v are zero mean, independent white noises, with variances Q and R,

respectively. These noises may represent actual noise sources or the effects of small non-linearities, unmodeled phenomena, etc. A simple calculation yields an equation for the covariances $P(k)$ and $S(k)$ of $x(k)$ and $y(k)$:

$$P(k+1) = AP(k)A' + Q, \quad S(k) = CP(k)C' + R \tag{A.5}$$

If A is a stable matrix, we can evaluate the steady-state covariances P and S by solving the Lyapunov equation

$$APA' - P = -Q \tag{A.6}$$

A.2 The Implementation of Digital Systems and Filters

As discussed in [A-1], the design of digital systems consists of several parts, including the specification of the desired input/output relationship and the implementation, using finite precision arithmetic, of a system that approximates this desired behavior. From this point of view, the methods of the preceding section deal with the first issue. Realization procedures play an indirect role in these techniques in providing the state space models on which the design methods are based. But what about realizations from the point of view of system synthesis and implementation? As we shall see, state space realizations can play some role, but they are far from providing the entire solution.

A wide variety of digital filter design methods have been developed to deal with the second issue. One factor that does enter into this design question is the number of storage elements (delays) in the filter structure, and thus the issue of minimality is of some importance. Of course, in

dealing with single-input, single-output transfer functions, one can read off the order of a canonic structure and can construct several quite easily by simple inspection of the specified transfer function. The determination of the order of a canonic realization and the ability to construct several minimal realizations without much difficulty barely scratches the surface of the structures problem, however. As pointed out in [A-1], the various filter structures available may be equivalent from an input-output viewpoint if one didn't have to worry about computation time, the complexity of the digital architecture or algorithm required to implement a given structure, the effect of finite precision in representing filter coefficients, or the effects of overflow and quantization. These are the issues that motivate much of the study of various filter structures [A-1, 10,11].

Let us examine some of these issues in the context of a particularly important structure, the cascade form, obtained by factoring a transfer function H(z) as a product of lower-order transfer functions. Consider the example

$$H(z) = \frac{z^2+(b+d)z+bd}{z^2-(a+c)z+ac} = \frac{(1+bz^{-1})(1+dz^{-1})}{(1-az^{-1})(1-cz^{-1})} \tag{A.7}$$

In Figure A.1 we have realized this filter as the cascade of two first order filters. Note that the overall filter is minimal.

In Section C we consider the effects on digital filter performance of quantization and overflow on system stability. An alternative, approximate method for evaluating the effect of finite word length on system performance is to model each quantization as if it introduced noise into the system
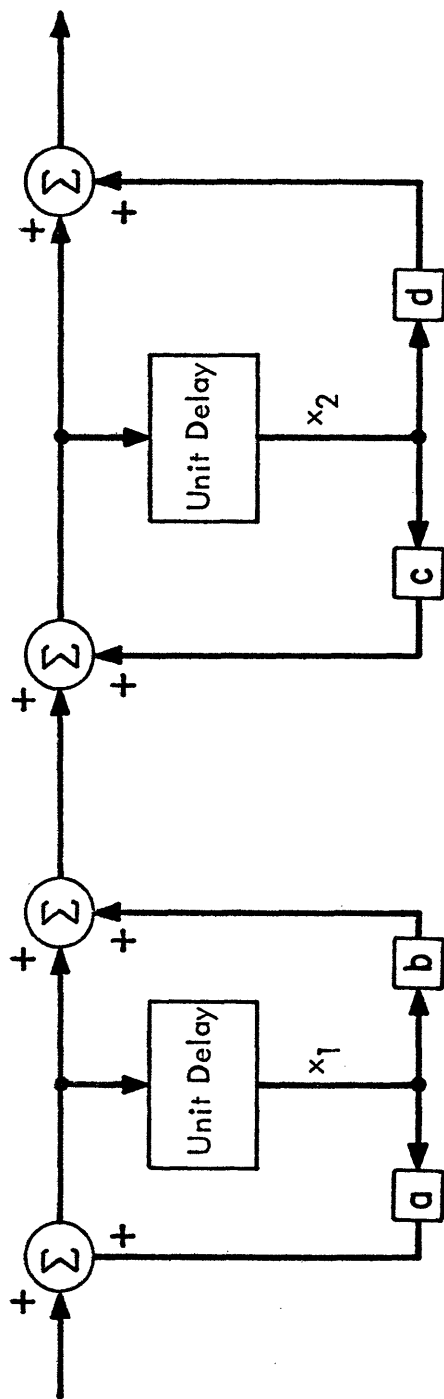
Figure A.1 A Cascade of Two First Order Filters

[A-1]. By assuming independence of these various sources--a rather strong

and sometimes unjustified assumption (as the existence of period effects,

i.e. limit cycles, indicates)--one can in principle evaluate the overall noise

power at the output, and thus obtain a measure of the size of the quantization

effects for any given filter structure. We can then use this information to

compare alternative designs (see [A-1,10,12] for detailed discussions of the

uses of such noise analysis techniques). As an example, consider the case [A-1]

of fixed-point arithmetic and roundoff quantization in which the quantization

interval q is 2^{-b}. In this case, the quantization error e introduced by a

single multiplication takes on a value between $\pm.5q$. If one makes the

assumption that e is uniformly distributed, we find that it has zero mean

and variance $q^2/12$. Then, for example, in the cascade filter of Figure A.1,

one could add one such noise source following each of the four multiplications.

We will return to this topic shortly to see how state space methods can be

used to analyze noise effects.

Another extremely important issue in filter design is the sensitivity

of filter performance to variation in coefficients. This is quite central

an issue, since one can only represent coefficients up to a finite degree

of accuracy, and hence one cannot obtain filters with arbitrary pole and

zero locations. The allowable poles and zeroes and the sensitivity to

variations in parameters depend quite significantly on the particular struc-

ture under consideration. For example, parallel and cascade structures are

often used because the perturbations in the poles are isolated from one

another.

For the remainder of this section, we wish to examine the relationship

of state space techniques and concepts to some of the questions in digital

filter design. Let us first examine the use of state space techniques to

determine filter structures. Consider the transfer function (A.7). In this

case, state space techniques yield a variety of minimal realizations of the form

$$x(k+1) = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} x(k) + \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} u(k) \qquad (A.8)$$

$$y(k) = [h_1, h_2] \, x(k) + u(k)$$

If we interpret (A.8) as an algorithm, we must compute the various products $f_{ij}x_j(k)$, $g_i u(k)$, $h_i x_i(k)$ (i, j=1,2), and perform the appropriate additions. Note that in general, there are 8 multiplications and 6 additions required.

Now consider the cascade structure of Figure A.1. Interpreting it as an algorithm (a and b multiply $x_1(k)$, c and d multiply $x_2(k)$, and we perform the required additions), we see that we require 4 multiplications and 4 additions, but this is not the most important difference between the two algorithms, since it is possible to obtain realizations (A.8) with some zero elements in (F,g,h). However, the crucial difference is the following: <u>if one interprets a state space realization as determining an algorithm of the type indicated, then there is no way that the cascade algorithm is of this type!</u> This is not to say that one cannot find a state-space description of the cascade realization. In fact

$$x(k+1) = \begin{bmatrix} a & 0 \\ (a+b) & c \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(k) \qquad (A.9)$$

$$y(k) = [(a+b), (c+d)]x(k) + u(k)$$

is such a realization.

The point made above may, at first glance, seem to be trivial, but it is not, since it points out that although any (infinite precision) algorithm can

be described dynamically in state space terms, direct interpretation of a state space description as an algorithm does not allow one to consider all possible algorithms. That is, it is relatively easy to go from an algorithm to a state-space description, but it is not at all natural or clear how to go the other way, and hindsight is needed in order to interpret a realization of the form of (A.9) as a cascade structure.

Thus, we see that state space models have limitations when one considers the issue of implementation. There are, however, several areas where interaction between the two fields may be of use. First of all, the techniques used in digital signal processing should be of use in considering the implementation of control and estimation system designs such as those mentioned in Subsection A.1. Also, recall that state space realization techniques allow one to determine minimal realizations for systems with multiple inputs and outputs. It is possible that this fact, combined with a thorough understanding of the relationship between state-space realizations and various digital system structures will lead to the development of useful filter structures for multivariable systems.

Also, as mentioned in the preceding subsection, the state space framework is particularly useful for the analysis of the properties of dynamical systems. Thus, it seems natural to ask if these techniques might be useful in the analysis of various filter structures. In Section C we discuss this question with respect to stability analysis techniques. Also, it is possible that state-space sensitivity techniques [A-9] could be useful in the study of the sensitivity of various digital filter structures, but this awaits further study.

Finally, let us examine the utility of state-space techniques in the analysis of the effect of quantization noise on filter performance. We do this by example, although it should be clear that this approach extends to arbitrary structures. Consider the cascade structure in Figure A.1 where we add quantization noise after each multiplication. A state space representation of this system can be written down by inspection:

$$x(k+1) = Fx(k) + gu(k) + \Gamma N(k)$$
$$y(k) = h'x(k) + u(k) + \Psi N(k)$$

(A.10)

where F, g, and h are given in (A.9), N(k) is the 4-dimensional noise vector whose components are the noises contributed by the multiplications by a, b, c, and d, respectively. Then $\Psi = (1,1,1,1)$, and

$$\Gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

(A.11)

If we make the usual independence assumptions concerning the components and time-behavior of N, we can directly apply the covariance analysis equations (A.5), (A.6) to determine the effect of quantization noise on x and y. Note that (A.5), (A.6) yield the effect of noise throughout the network. The utility of an approach such as this for digital network analysis needs to be examined more carefully, but it appears that it may be computationally superior to other methods, such as those that use signal flow graph techniques [A-12] or that require computing a number of partial transfer functions [A-3]. We note that Parker and Girard [A-15] used Lyapunov-type equations and analysis quite similar to our development for the evaluation of output noise power due to correlated quantization errors. In addition, similar

analyses have been undertaken by Hwang [A-17], Mullis and Roberts [A-18], and Sripad and Snyder [A-19,20]. Hwang uses Lyapunov-state space equations to study the effects of possible structure transformations and state-amplitude scalings. Mullis and Roberts have obtained some significant results for digital filter design using a framework similar to Hwang's to study what they call "minimal noise realizations (see [A-31] for further developments). Sripad and Snyder develop conditions under which quantization errors are in fact white, and they also use Lyapunov-type analysis to compare the performance of two different realizations. Within this framework, one can pose a number of other questions. For example, in the case of floating point arithmetic, the quantization error depends on the size of the signal. Can state-space procedures for analyzing "state-dependent noise" [A-16] be of value here? Questions such as these await future investigation.

In this section we have seen some of the issues involved in system design in the two fields. The issue of implementation is at the very heart of the problems considered by researchers in digital signal processing, while researchers in control and estimation have concentrated more on the development of general design procedures for state space models and methods for analyzing the characteristics of such models. We have seen that there are points of overlap and places in which techniques and concepts from one discipline may be of value in the other. State space techniques may be useful in the analysis of multivariable structures and in the analysis of sensitivity and quantization noise behavior of different structures. Such issues remain to be studied, but it is in the other direction that there is

the most to be done. The issues involved in the digital implementation of
systems specified by state space design methods remain largely unexplored.
Numerous problems abound. What is the effect of roundoff noise on closed-
loop controller performance, and how many bits must we use to achieve the
desired regulation properties [A-21,22,24,25,28]? It is well known that
"optimal" controllers and estimators require many arithmetic operations and
hence lead to low sampling rates. Can we improve overall performance by
using a simpler "suboptimal" system at a higher sampling rate [A-13]? If we
are controlling a complex system, "optimal" controllers require not only a
great deal of computation, but also the centralized processing of all infor-
mation, and the cost of relaying information to a central location may be
prohibitive. Can we devise decentralized control architectures that take
advantage both of the structure of the dynamics of the system being controlled
and the capabilities of the available types of digital processors? Here
again, if we include the cost of information transfer, "suboptimal" decen-
tralized systems may outperform the "optimal" system (see [A-14, 23, 26] for
some results and problems concerned with parallel and distributed processing
and decentralized control).

The study of problems such as these -- i.e., the interaction of imple-
mentation and architecture issues and the design of control and estimation
systems -- is still in its infancy, and it appears to offer an extremely
promising avenue for research. We note that architectural issues have
received a great deal of attention in the field of digital signal processing
[A-10,12], and this, together with the wealth of literature on digital filter

structures, indicates that there is much to be gained from future interaction and collaboration.

B. Parameter Identification, Linear Prediction, Least Squares, and Kalman Filtering

A problem of great importance in many disciplines is the determination of the parameters of a model given observations of the physical process being modeled. In control theory this problem is often called the system identification problem, and it arises in many contexts. The reader is referred to the special issue of the IEEE Transactions on Automatic Control [B-10] and to the survey paper of Åström and Eykhoff [B-11] for detailed discussions and numerous references.

Parameter identification problems also arise in several digital signal processing applications. Several examples of such problems are given in the special issue of the Proceedings of the IEEE [B-33], and one of these, the analysis, coding, and synthesis of speech, has received a great deal of attention in the past few years [B-15, 21-23]. We will use this problem as the basis for our discussion of the identification question. Our presentation is necessarily brief and intuitive, and the reader is referred to these references for details. As discussed in [B-21] a popular approach is to model a discretized speech signal $\{y(k)\}$ as the output of a linear system, which, over short enough intervals of time, can be represented by a time-invariant transfer function $G(z)$. Here the input is taken as a periodic pulse train (whose period is the pitch period) for voiced sounds (such as vowels) and as white noise for unvoiced sounds (such as the sound "sh"). In addition, a common assumption is that G is an all-pole filter, which leads to an autoregressive (AR) model

$$y(k) + a_1 y(k-1) + \ldots + a_p y(k-p) = u(k) \qquad \text{(B.1)}$$

This assumption has been justified in the literature under many conditions, although strong nasal sounds require zeros [B-21].

The problem now is to determine the coefficients $a_1, \ldots, a_p$. Having these coefficients, one is in a position to solve a number of speech analysis and communication problems. For example, one can use the model (B.1) to estimate formant frequencies and bandwidths, where the formants are the resonances of the vocal tract [B-24]. In addition, one can use the model (B.1) for efficient coding, transmission, and synthesis of speech [B-29]. The basic idea here is the following: as the model (B.1) indicates, the speech signal $y(k)$ contains highly redundant information, and a straight-forward transmission of the signal will require high channel capacity for accurate reconstruction of speech. On the other hand, one can interpret (B.1) as specifying a one-step predictor for $y(k)$ in terms of preceding values of y (assuming $u(k) = 0$). As discussed in [B-29], one often requires far fewer bits to code the prediction error u than the original signal y. Thus, one arrives at an efficient transmission scheme (linear predictive coding--LPC): given y, estimate the a_i, compute u, transmit the a_i and u. At the receiver, we then can use (B.1) to reconstruct y. An alternative interpretation of this procedure is the following: given y, estimate G, pass y through the inverse, all zero (moving average -- MA) filter $1/G(z)$, transmit the coefficients in G and the output of the inverse filter. At the receiver, we then pass the received signal through G to recover y (thus this procedure is causal and causally invertible).

The question remains as to how one estimates the a_i. The most widely used technique in the literature is linear prediction. Using the inter-

pretation of (B.1), as specifying a one-step predictor for the signal y,

we wish to choose the coefficients $a_1, \ldots, a_p$ to minimize the sum of squares

of the prediction errors $e(n) = y(n) - \hat{y}(n)$, $n \in I$. Here we assume that we are

given $y(0), \ldots, y(N-1)$, while the set I can be chosen in different manners,

and we will see in the following subsections that different choices can

lead to different results <u>and</u> to different interpretations.

Before beginning these investigations, let us carry out the minimization

required in linear prediction. Taking the first derivative with respect to

the a_i of the sum of squared errors, and setting this equal to zero, we

obtain the <u>normal equations</u>

$$\sum_{i=1}^{p} a_i c_{ik} = -c_{0k}, \qquad k = 1, \ldots, p \tag{B.2}$$

$$c_{ik} = \sum_{n \in I} y(n-i) y(n-k) \tag{B.3}$$

These equations are typical of the types of equations that arise in linear,

least-squares problems, and their efficient solution has been the topic of

many research efforts.

B.1 The Autocorrelation Method, Kalman Filtering for Stationary Process, and Fast Algorithms

Suppose we let I = all integers, where we define $y(n) = 0$ for $n < 0$,

$n \geq N$. In this case, we find that

$$c_{ij} = \sum_{n=0}^{N-1-|i-j|} y(n) y(n+|i-j|) \triangleq r(|i-j|) \tag{B.4}$$

and the normal equations become $Ta = d$, where $a' = (a_1, \ldots, a_p)$, (here the

prime denotes transpose),

$d' = (-r(1), -r(2),\ldots,-r(p))$, and T is a symmetric Toeplitz matrix
[B-37,84,91] (i.e. the ij^{th} element depends only on $|i-j|$) with $T_{ij} = c_{ij}$.
We also note [B-15] that if y is a stationary random process with autocorre-
lation

$$R(i) = E[y(n)y(n+i)] \qquad\qquad (B.5)$$

and if we want to find the one step prediction coefficients to minimize
$E[e^2(n)]$, we obtain an identical set of (Toeplitz) equations with $r(i)$
replaced by $R(i)$. This statistical point of view is extremely useful for
obtaining certain insights into the approach.

The solution of such Toeplitz equations has been the subject of a great
deal of attention in the mathematical, statistical, and engineering litera-
ture [B-3,4,15,17,18]. An efficient algorithm was proposed by Levinson
[B-17], improved upon by Durbin [B-32], and studied in the speech processing
context by several authors, including Itakura and Saito [B-23]. The method
essentially consists of solving forward and backward prediction problems
of increasing size in a recursive manner. That is, the algorithm computes
the coefficients $a(1|i,\ldots,a(i|i)$ for the best prediction of $y(n)$ based on
$y(n-1),\ldots,y(n-i)$ and the coefficients $b(1|i),\ldots,b(i|i)$ for the best pre-
diction of $y(n-i-1)$ based on $y(n-i),\ldots,y(n-1)$. The algorithm iterates
on i. As a part of this algorithm, at each stage one computes the prediction
error (for both forward and backward prediction), and thus one can determine
when to stop based on the size of this quantity. Also, at each stage we must
compute a coefficient k_i, which is known as the <u>partial correlation coefficient</u>
(see [B-15,21,23]). Finally, while the Levinson algorithm is extremely
efficient, it must be kept in mind that the computational savings reaped here
may be quite small compared to the calculation of the correlations according
to (B.4).

Let us now examine what this algorithm means from a statistical point of view. The algorithm specifies estimators of the form

$$\hat{y}(i) = - \sum_{j=1}^{i} a(j|i)y(i-j) \qquad (B.6)$$

$$\hat{y}(0) = - \sum_{j=1}^{i} b(j|i)y(j) \qquad (B.7)$$

Thus, we can think of the algorithm as providing us with the time-varying coefficients of the weighting pattern of the optimal one-step predictor and of the optimal initial time smoother. Note that these coefficients are, in general, time varying (i.e., $a(j|i) \neq a(j)$), since the mechanism of prediction is time-varying when one bases the prediction on only a finite set of data.

What does this mean as far as all-pole modeling via linear prediction goes? The answer to that is not much. In the all-pole modeling problem, we are equivalently only interested in designing a FIR filter -- i.e. a prediction filter that produces the best estimate of y(n) given the "data window" y(n-1),...,y(n-p). The coefficients of such a filter are precisely $a(1|p),...,a(p|p)$, and it doesn't matter (except from a computational point of view) that these coefficients were generated as part of a time-varying filter weighting pattern.

On the other hand, the time-varying weighting pattern interpretation is extremely important from a statistical point of view, especially if one wishes to design recursive predictors that are capable of incorporating all past measurements and not just a data window in the case when y has a

Markovian representation

$$x(k+1) = Ax(k) + w(k), \qquad y(k) = c'x(k) \qquad (B.8)$$

where x is a random n-vector, A is a constant nxn matrix, c is a constant
n-vector, and w is a zero-mean uncorrelated sequence with covariance Q.
The correlation coefficients of y can be computed by direct examination of
(B.8). We note that x and y will be stationary with

$$R(i) = c'A^i \Pi c \qquad i \geq 0 \qquad (B.9)$$

if A is stable and if Π, the covariance of x, satisfies the Lyapunov equation

$$A\Pi A' - \Pi = -Q \qquad (B.10)$$

We now wish to design an optimal predictor for recursively estimating y(n).
This is a standard estimation problem, and the solution is the Kalman filter
[B-4]:

$$\hat{x}(n) = A\hat{x}(n-1) + AK(n-1)\gamma(n-1) \ , \ \hat{y}(n) - c'x(n)$$
$$\gamma(n-1) = y(n-1) - \hat{y}(n-1) \qquad (B.11)$$

where the time-varying gain satisfies

$$K(n) = \frac{P(n|n-1)c}{c'P(n|n-1)c} \qquad (B.12)$$

Here $P(n|n-1)$ is the covariance of the prediction error $x(n) - \hat{x}(n)$,

$$P(n+1|n) = AP(n|n-1)A' + Q - \frac{AP(n|n-1)cc'P(n|n-1)A'}{c'P(n|n-1)c} \qquad (B.13)$$

Let us make a few comments about these equations. Note that the _filter_
innovations $\gamma(n)$ is precisely the prediction error e(n), and its variance
is $c'P(n|n-1)c$. Also, recall that in the all-pole framework, we could alter-

natively view the prediction filter as specifying an inverse filter, which took the y's as inputs and produced the uncorrelated sequence of prediction errors as the output. In the context of the Kalman filter, the analogous filter is the innovations representation (see representation IR-1 of [B-27]), in which we view the output of (B.11) as being $\gamma(n)$. Finally, note that one can compute the predictor coefficients $a(j|i)$ as the weighting pattern of the filter (B.11).

Examining (B.11)-(B.13), we see that the computation of the recursive filter coefficients requires the solution of the (discrete time) Riccati equation (B.13). If x is an n-vector, then (using the fact that P is symmetric), (B.13) represents $n(n+1)/2$ equations. For reasonably large values of n, this can be an extreme computational load, especially given that all that is needed for the filter is the nxm gain matrix K (when y is m-dimensional). Thus when m << n, the question of computing K without P arises quite naturally, and this issue -- in both continuous and discrete time, in stationary and in some nonstationary cases -- has been the subject of numerous papers in the recent past [B-1-5]. The underlying concepts that have led to these "fast algorithms" (at least in the stationary case) are the same as those that lead to the Levinson algorithm. For some historical and mathematical perspective on this subject, we refer the reader to [B-3,4]. In particular, the extension of the Levinson algorithm to the multivariable case is discussed in these papers (see also reference [B-18]). In this case, the matrix T in the normal equations is block-Toeplitz, and the extension to this case is decidedly nontrivial.

There are a number of deep mathematical and physical insights that can

be obtained by the examination of these fast algorithms. As discussed in [B-15,21], the Levinson algorithm involves an auxiliary partial correlation coefficient k_i, which has an interpretation as a reflection coefficient, and this fact has been utilized in speech processing, in which these coefficients specify certain parameters in an acoustic model of the speech process [B-15,21]. In addition Casti and Tse [B-20], Kailath [B-1,4] and Sidhu and Casti [B-8] have shown that the fast Kalman gain algorithms are closely related to the work of certain astrophysicists, in particular Chandrasekhar [B-19], who devised algorithms for solving finite time Wiener-Hopf equations arising in radiative transfer. Also, relationships between linear filtering and scattering theory have been brought to light in the recent papers [B-34,35]. And finally, for a good overview of some of the mathematical relationships, including some with the theory of orthogonal polynomials, we refer the reader to [B-4,42]. These ideas are of interest in that seeing these algorithms from several perspectives allows us to gain insight into their properties, potentials, and limitations.

B.2 The Covariance Method, Recursive Least Squares Identification, and Kalman Filters

Consider again the normal equations (B.2), (B.3). We now consider the range of n to be only as large as the actual data allows -- i.e., in equation (B.1) we will require that k, k-1,...,k-p all be within the range 0,...,N-1. This leads to the restriction $p \leq n \leq N-1$. Also, in this case c_{ij} is not in general a function of i-j, the matrix T is symmetric but not Toeplitz, and the fast methods of the preceding subsection don't carry over quite so nicely. Recently, however, Morf, et al. [B-30] have obtained fast

algorithms for the covariance method by exploiting the fact that, although T is not Toeplitz, it is the product of Toeplitz matrices.

Let us take a look at the covariance method from a slightly different point of view. Recall that the algorithm mentioned above and the one in the preceding subsection involve recursions on the order of the filter given a fixed set of data. Suppose now we consider a recursion for updating coefficients of a fixed order filter given more and more data. To do this, we refer to [B-11], where the covariance method is discussed. Given the data $y(0), \ldots, y(N-1)$, the covariance method attempts to fit a model of the form of (B.1) by finding a least squares fit $\hat{a}(N)$ to the equation

$$L_{N-1} a = f_{N-1} \qquad\qquad (B.14)$$

where $a' = (a_1, \ldots, a_p)$, $f'_{N-1} = (y(p), \ldots, y(N-1))$, and L has various $y(i)$ as its elements. Suppose we have $\hat{a}(N-1)$ and we now obtain the new data point $y(N)$. We would like to update our estimate in a manner more efficient than re-solving (B.14) from scratch. Following standard recursive least squares (RLS) procedures [B-11], we find that (here $\ell'(N)$ is the last row of L_N):

$$\hat{a}(N) = \hat{a}(N-1) + K(N) [y(N) - \ell'(N) a(N-1)] \triangleq \hat{a}(N-1) + K(N) r(N)$$

$$(B.15)$$

$$K(N) = \frac{P(N-1)\ell(N)}{1 + \ell'(N)P(N-1)\ell(N)} \qquad\qquad (B.16)$$

$$P(N) = (L'_N L_N)^{-1} = P(N-1) - \frac{P(N-1)\ell(N)\ell'(N)P(N-1)}{1 + \ell'(N)P(N-1)\ell(N)} \qquad (B.17)$$

Examining these equations, we see that they can be interpreted as defining a Kalman filter (see [B-12]). In fact, referring to [B-14], we

see that these are precisely the Kalman filter equations used by Melsa, et al. in speech processing. Specifically, they consider the dynamic equations

$$a(k+1) = a(k), \quad y(k) = \ell'(k)a(k) + v(k) \tag{B.18}$$

where $v(k)$ is a zero-mean, white process with variance Ψ. If Ψ is set to 1, we obtain (B.15)-(B.17). Also, in this formulation, $P(N)$ has the interpretation as the covariance of the estimation error $a-\hat{a}(N)$.

Let us note some of the properties of the recursive solution (B.15)-(B.17). Examining (B.15), we see that the increment in our estimate $\hat{a}$ is proportional to the error (innovations) in predicting the latest value of y using preceding values and our previous estimate of a. This suggests that a monitoring of the innovations $r(N)$ can be used to help detect abrupt changes in the predictor coefficients or the presence of glottal excitation in voiced sounds. In this manner one may be able to improve upon the estimation of a. Whether such a procedure would be of value is a matter for future study. Also, it is possible to make the filter more responsive to changes in the coefficients by using one of several methods available for adjusting Kalman filter [B-41]. These include exponentially age-weighting old data in favor of the more recent pieces of information or the modeling of a as a slowly-varying Markov process. In addition, the formulation (B.18) provides a method for developing an analysis system for noise-degraded speech (i.e., the case when $\Psi > 1$).

Let us now consider the computational complexity of (B.15)-(B.17). First note that one does not have to compute the correlation coefficients. However, one does have to calculate $K(N)$ at every stage, and if one solves for the gain

from the Riccati equation (B.62), one has on the order of p^2 multiplications

per stage. However, Morf, et al. [B-30] and Morf and Ljung [B-40] have

exploited the structure of the equations to obtain fast algorithms for the

direction computation of K. Combined with the fast algorithms mentioned

earlier, one now has efficient recursive procedures for the covariance method

as one increases either the order p of the predictor or the number N of data

points.

B.3 Design of a Predictor as a Stochastic Realization Problem

A problem that has attracted a great deal of attention in the control

and estimation literature is the stochastic realization problem [B-4,8-10,

13,27]. Briefly stated, the stochastic realization problem asks the

following: given a stationary Gaussian random process y with correlation

function R(n), find a Markovian representation

$$x(n+1) = Ax(n) + w(n), \quad y(n) = c'x(n) \tag{B.19}$$

where w is a zero mean white noise process with covariance Q. Referring

to (B.8)-(B.10), we see that this is equivalent to finding a factorization:

$$R(i) = c'A^i b \qquad i \geq 0 \tag{B.20}$$

$$b = Pc, \qquad APA'-P = -Q \tag{B.21}$$

Examining (B.20), (B.21), we see that the algorithm falls naturally into

two pieces: (1) find a triple (A,b,c) satisfying (B.20); (2) find P and Q

satisfying (B.21). One of the best-known studies of this problem is that of

Faurre [B-13,25]. As he pointed out, the first step of the algorithm is

simply the well-known deterministic realization problem when one is given

the "weighting pattern" R(0), R(1), R(2),... . This problem has been widely
studied in the literature [A-30, 31, B-6,9] and we will make a few comments
about this aspect of the problem in a few moments. Before discussing the
numerical aspects of the first step or the details of the second, let us
see what the first part yields in the frequency domain [B-26]. Let

$$S_y(z) = \sum_{i=-\infty}^{+\infty} R(i) z^{-i} \qquad (B.22)$$

Then, we see that the factorization (B.20) yields

$$S_y(z) = c'(zI-A)^{-1} zb + c'(z^{-1}I-A)^{-1} Ab \qquad (B.23)$$

Defining $\alpha(z) = \det(zI-A)$ and putting the two terms in (B.23) over a common
denominator, we see that the first step in the algorithm yields

$$S_y(z) = \frac{p(z)}{\alpha(z)\alpha(z^{-1})} \qquad (B.24)$$

That is, we have obtained a factorization of the denominator of S_y. If we
can also factor the numerator we will have determined the desired transfer
function $\beta(z)/\alpha(z)$, which, when driven by white noise, yields the spectrum
$S_y(z)$. It is clear from (B.19) that it is this second part of the spectral
factorization that is accomplished by the second step of the stochastic
realization algorithm. Finally, note that the model obtained contains <u>both</u>
poles and zeroes.

There are several methods for performing the second step of the
algorithm. Faurre [B-13] showed that (B.21) could be solved for values
of P inside a given range, and he identified the smallest such covariance,
P_*, as that arising from an <u>innovations representation</u> of y -- i.e., a steady-

state Kalman filter. Thus to complete the second step we can either solve

an algebraic Riccati equation or can use the "fast algorithms", as described

in Subsection B.1 to compute the time-varying Kalman gain. Letting the

transients die out, we then obtain the desired steady-state filter. Although

this approach involves solving for the entire gain time history, this proce-

dure may be faster than direct solution of the algebraic Riccati equation.

Let us now turn to the numerical aspects of the first stage -- i.e. the

computation of the factorization (B.20). The algorithms of Rissanen [B-8]

and Ho [A-29] are based on the examination of the Hankel matrix

$$
H_N = \begin{bmatrix} R(0) & R(1) & R(2) & \cdots & R(N-1) \\ R(1) & R(2) & R(3) & \cdots & R(N) \\ \vdots & & & & \\ R(N-1) & R(N) & R(N+1) & \cdots & R(2N-2) \end{bmatrix} \tag{B.25}
$$

It is well-known [B-36] (see also Subsection A.1) that R admits a factoriza-

tion (B.20) if and only if there is some integer n such that

$$
\text{rank } H_N \le n \quad \text{for all } N \tag{B.26}
$$

Ho's original algorithm yielded a minimal realization (i.e. dim A in (B.20)

is as small as possible) if a bound n was known in advance. A far more

critical question (from a practical point of view) is the partial realization

question. Here we take into account that we only have available a finite

number of correlations R(0), R(1), ..., R(N-1), and one would like to obtain

the minimal factorization that matches these. One can use Ho's algorithm

for this, but it is not recursive -- i.e. if we incorporate R(N), we must

re-solve the whole problem. Fortunately, Rissanen [B-8] and Dickinson, et al.

[B-6] have developed efficient, recursive procedures (the latter of which is based on the Berlekamp-Massey algorithm [B-7], which was developed for the scalar case). We note that these algorithms essentially solve the Pade approximation problem, and we refer the reader to the references for details.

Thus, efficient algorithms exist for spectral factorization and one would expect good results <u>if the process y truly has a Markovian representation and if one has the exact values of the correlations</u>. This points out a conceptual difference between linear prediction and the above stochastic realization procedure. In linear prediction, no pretense is made about exactly matching a model. All that is wanted is a least-squares fit, and thus one would expect this procedure to be relatively robust when one uses a finite record of real data to generate an estimate of the correlation function which is then used in the linear prediction procedure. On the other hand, it can easily be seen that an infinitesimal perturbation of H_N in (B.25) can make it have full rank. In this case, the partial realization procedures -- which in essence are looking to match a model exactly -- will yield a system of extremely high dimension. Thus, it appears that these algorithms are inherently sensitive to errors in estimates of the correlation coefficients. In addition, if y has no Markovian representation, the linear prediction approach may yield a useful fit, while the partial realization procedure may very well run astray as it tries to fit the data "too closely".

Does this mean that the above procedure is of no use in identifying parameters in a speech model? The answer to that is perhaps not. What is needed is a modification of the first step of the stochastic realization algorithm. As the version described here stands, it is too sensitive and

in fact, DeJong [B-37] has shown that these methods are numerically unstable in that the inexact minimal realization supplied by these algorithms, as implemented on a finite wordlength computer, may not be a "numerical neighbor" of the sequence $\{R(i)\}$ that is to be factored. By rephrasing the algorithm in terms of ϵ-rank -- the least rank of all systems within an "ϵ-neighborhood" of the given sequence -- DeJong obtains a slower algorithm that is similar to Rissanen's but is numerically stable. This approach is extremely appealing for two reasons: (1) We can, within this framework, seek minimal realizations in the ϵ-neighborhood of a sequence $\{R(i)\}$ that itself is not realizable by a finite dimensional system; (2) We can seek the "nearest" reduced-order realization of given dimension of a given system.

In addition to the work of DeJong, a number of other methods have been proposed for "approximate" Pade approximations [B-31,38,39]. One interesting possibility is the all-pole approximations -- i.e., we perform linear prediction on the R(i). This would require computing the correlation of the R(i)! (Note that an all-pole assumption here would not necessarily lead to an all-pole model in (B.19)).

One of our goals in this section has been to point out a number of similarities between the goals and techniques of the two disciplines. We have also seen some of the differences, but others have not been discussed. In particular, in this section we have treated identification for identification's sake. As pointed out in [B-11] in control system design, identification is often simply a means toward the goal of efficient control. Thus, in many control applications, the value of identification is not measured by

the accuracy of the parameter estimates, but rather by the performance of the overall system. In addition, in control there are several types of identification problems, since one has the opportunity to excite the system through inputs. Different problems arise if the system is operating open loop, in a time-invariant closed-loop mode, or in an adaptive closed loop mode. We refer the reader to [B-10,12] for more on this subject and for further references. In addition, in many on-line control problems the number of parameters to be identified is not very large -- four or five. In fact, one of the key problems in practical adaptive control is the careful choosing of which few parameters to identify.

On the digital filtering side, one is often interested in the accuracy of the parameter estimates. This is of importance, for example, if one is attempting to design an all-pole filter that matches a given impulse response in a least squares sense, or if one is attempting to estimate formants from an all-pole speech model. On the other hand, for linear predictive coding, the accuracy of the parameters may be of secondary interest, while the primary concern is more efficient coding of speech data. In this case, accuracy is of importance only in so far as it makes the coding scheme more efficient. Also, in the speech problem, we are usually dealing with many unknown parameters -- between twelve and sixteen [B-21].

With regard to the speech problem, we note that linear prediction has proven to be a particularly appropriate tool for a variety of reasons, ranging from the fact that the all-pole model is often a realistic one to the property that the linear prediction procedure tends to match the spectral envelope of the data [B-15]. In this section we have described a number of related

identification concepts (see [Z-1] for more), some of which may be useful in solving problems in speech analysis, such as enhancement of noise-degraded speech. We have also pointed out a number of questions concerning some of these methods, such as the need for detailed numerical analyses of the many "fast" algorithms, and the necessity of further analysis and experimentation to assess whether any of these techniques can improve upon the performance achievable in speech processing using linear prediction.

C. Stability Analysis

In the field of digital signal processing, stability issues arise when one considers the consequences of finite word length in digital filters. On the one hand, a digital filter necessarily has finite range, and thus overflows can occur, while on the other, one is inevitably faced with the problem of numerical quantization -- roundoff or truncation. Since the filter has finite range, the question of the state of the filter growing without bound is irrelevant. However, the nonlinearities in the filter, introduced by whatever form of finite arithemtic is used, can cause zero-input limit cycles and can also lead to discrepancies between the ideal and actual response of the filter to certain inputs. Following the discussions in [C-3, 9], the typical situation can be described as follows

$$x(n+1) = A\chi(n) + Bu(n), \qquad y(n) = Cx(n) \tag{C.1}$$

$$\chi(n) = N(x(n))$$

where N is a nonlinear, memoryless function that accounts for the effects of overflow and quantization. If one assumes that the associated linear system (i.e., N = identity) is designed to meet certain specifications, one would like to know how the nonlinearity N affects overall performance. For example, filter designers are interested in determining bounds on the magnitudes of limit cycles and in finding out how many bits one needs to keep the magnitudes of such oscillations within tolerable limits.

On the other side, a typical feedback control system is described by

$$y = G_1(e), \qquad e = u - G_2(y) \tag{C.2}$$

where the input u, the output y, and the error e are functions of time, and G_1 and G_2 represent the dynamics of the forward and feedback paths, respectively. In control theory one is interested either in the analysis or the synthesis of such systems. In the synthesis problem we are given an open loop system G_1 and are asked to design a feedback system G_2 such that the overall system has certain desirable stability properties. In the case of stability analysis, one may be interested either in the driven or the undriven characteristics. In the driven case the problem involves determining if bounded inputs lead to bounded outputs and if small changes in u lead to small changes in the y. In the undriven case, we are interested in seeing if the system response decays, remains bounded, or diverges when the only perturbing influences are initial conditions.

It is clear that the problems of interest to researchers in both disciplines have a good deal in common, and, as we shall see, workers in each area have obtained results by drawing from very similar bags of mathematical tricks. However, there are differences between the methods used and results obtained in the two areas. In the analysis of digital filters the work has been characterized by the study of systems containing quite specific non-linearities. In addition, much of the work has dealt with specific filter structure. In particular, second-order filters have received a great deal of attention [C-2,3,7,9], since more complex filters can be built out of seres - parallel interconnections of such sections. Also, the class of wave digital filters [C-5,6] have been studied in some detail. Studies in

these areas have yielded extremely detailed descriptions of regions of
stability in parameter space and numerous upper and lower bounds on limit
cycle magnitudes (see [C-3,4,13,24-26]).

In control theory, on the other hand, the recent trend has been in the
development of rather general theories, concepts, and techniques for sta-
bility analysis. A number of rather powerful mathematical techniques have
been developed, but there has not been as much attention paid to obtaining
tight bounds for specific problems. In addition, problems involving limit
cycles have not received nearly as much attention in recent years as issues
such as bounded-input, bounded-output stability and global asymptotic sta-
bility (although there clearly is a relationship between these issues and
limit cycles).

C.1 The Use of Lyapunov Theory

The technique of constructing Lyapunov functions to prove the stability
of dynamical systems has been used by researchers in both fields (see [C-22]
for details and further discussions). Consider a system with state $x(k)$ and
with equilibrium point $x = 0$. A Lyapunov function $V(x)$ for this system is
a scalar function for which $V(0) = 0$ and which is nonincreasing along system
trajectories (i.e. $V(x(k))$ is nonincreasing as a function of time).

If this function has some additional properties, we can prove stability
or instability. Basically, we think of V as an "energy" function. One then
obtains results depending upon how energy behaves along trajectories. In-
tuitively, if V is everywhere positive except at $x = 0$ and $V(x(k))$ decreases
monotonically, the system dissipates energy and is stable. On the other
hand, if $V(x_0) < 0$ for some x_0, then the system cannot be asymptotically

stable, since the nonincreasing nature of $V(x(k))$ guarantees that the system can't approach the zero energy state if started at x_0. One advantage of Lyapunov-type results is that the hypotheses for results such as those just mentioned can be checked without the construction of explicit solutions to difference or differential equations. However, the major problem with the theory is the difficulty in finding Lyapunov functions in general.

With respect to the limit cycle problem, Willson [C-2,8] has utilized Lyapunov functions to determine conditions under which second order digital filters will not have overflow limit cycles and will respond to "small" inputs in a manner that is asymptotically close to the ideal response. Parker and Hess [C-13] and Johnson and Lack [C-25,26] have used Lyapunov functions to obtain bounds on the magnitude of limit cylces. In each of these the Lyapunov function used was a quadratic form which in fact proved asymptotic stability for the ideal linear system. In Willson's work [C-8], he was able to show that his results were in some sense tight by constructing counterexamples when his condition was violated. In [C-13,25,26] the bounds are not as good as others that have been found, and, as Parker and Hess state, this may be due to the difficulty of determining which quadratic Lyapunov function to use. As pointed out by Claasen, et al., [C-3], it appears to be difficult to find appropriate Lyapunov functions for the discontinuous nonlinearities that characterize quantization.

There is a class of digital filters -- wave digital filters (WDF) [C-5,6] -- for which Lyapunov techniques are particularly useful. Such filters have been developed by Fettweis so that they possess many of the pro-

perties of classical analog filters. Motivated by these analogies, Fett-
weis [C-5] defines the notion of "instantaneous pseudopower", which is a
particular quadratic form in the state of the WDF. By defining the notion
of "pseudopassivity" of such a filter, Fettweis introduces the notion of
dissipativeness. With this framework, the pseudopower becomes a natural
candidate for a Lyapunov function, and in [C-6], Fettweis and Meerkötter
are able to apply standard Lyapunov arguments to obtain conditions on the
arithmetic used that guarantee the asymptotic stability of pseudopassive
WDF's. The introduction of the concept of dissipativeness in the study of
stability is an often-used idea (see the note of Deoser [C-14]), and a
number of important stability results have as their basis some notion of
passivity. We also note that the use of passivity concepts and the tools
of Lyapunov theory appear to be of some value in the development of new
digital filter structures that behave well in the presence of quantization
[C-7].

Lyapunov concepts have found numerous applications in control theory.
The construction of quadratic Lyapunov equations for linear systems is well
understood and is described in detail in [C-22]. The key result in this
area is the following. Consider the discrete-time system

$$x(k+1) = Ax(k) \tag{C.3}$$

This system is asymptotically stable if and only if for any positive definite
matrix L, the solution Q of the (discrete) Lyapunov equation

$$A'QA - Q = -L \tag{C.4}$$

is also positive definite. In this case the function x'Qx is a Lyapunov

function that proves the asymptotic stability of (C.3). Note that this result provides a variety of choices for Lyapunov functions (we can choose any $L > 0$ in (C.4)). Parker and Hess [C-13] obtain their bounds by choosing $L = I$ (here (C.3) represents the ideal linear model). Tighter bounds might be possible with other choices of L, but, as they mention, it is not at all clear how one would go about finding a "better" choice (other than by trial and error).

In addition to their direct use in specific applications, one of the most important uses of Lyapunov concepts is as an intermediate step in the development of other more explicit results. For example, the stability of optimal linear regulators with quadratic criteria and of optimal linear estimators can be proven by constructing particularly natural quadratic Lyapunov functions [B-41, C-32]. A further use of Lyapunov theory has been to provide a framework for the development of many more explicit stability criteria. Examples of these are a number of the frequency domain stability criteria that have been developed in the last 10 to 15 years (see [C-10-12, 15,16,20,21]). These results are the subject of the next subsection.

C.2 Frequency Domain Criteria, Passivity, and Lyapunov Functions

We have already mentioned that the notion of passivity is of importance in stability theory and have seen that Fettweis and Meerkötter have been able to utilize passivity notions to study certain digital filters via Lyapunov techniques. The relationship between passivity, Lyapunov functions, and many of the frequency domain criteria of stability theory is quite deep, and in this subsection we wish to illustrate some of these

ideas. We refer the reader to the work of J.C. Willems [C-19,23,30], in particular, for a detailed development. The general development in these references is beyond the scope of this paper, but we will indicate some of the basic ideas for a discrete-time system, denoted by the symbol G, with input u and output y. In this case, one can define input/output (I/O) stability as

$$\sum_{i=1}^{\infty} u_i^2 < \infty \quad => \quad \sum_{i=1}^{\infty} y_i^2 < \infty \qquad (C.5)$$

i.e., if the input has finite energy, so does the output. If we can make the stronger statement

$$\left(\sum_{i=1}^{\infty} y_i^2 \right)^{1/2} \leq K \left(\sum_{i=1}^{\infty} u_i^2 \right)^{1/2} \qquad (C.6)$$

we call K the I/O gain. A system is called <u>passive</u> if (<u>strictly passive</u> if there is an ε>0 such that)

$$\sum_{i=1}^{N} u_i y_i \geq 0 \left(\geq \varepsilon \sum_{i=1}^{N} u_i^2 \right), \quad \text{for all } u_i, N \qquad (C.7)$$

The motivation for the definition (C.7) stems from the interpretation of the left-hand side of (C.7) as the total energy input to the system. Thus a passive system always requires a positive amount of energy to be fed into it. This notion has extremely strong ties to the usual notions of passivity and dissipativeness for electrical networks and is, in fact, a natural generalization of these concepts [C-30,34].

Having this framework, one can derive important results on the stability and passivity of feedback interconnections of passive systems (see [C-30]),

much like the results of Fettweis for his pseudopassive blocks. As out-
lined by Willems in [C-30], there are three basic stability principles.
The first involves the interconnection of passive systems as mentioned
above, while the second is the small loop gain theorem (stability arises
if the gain around the loop is less than unity -- a result used in the
digital filter context in [C-31]). The third result involves notions of
passivity and of sector nonlinearities. A nonlinearity is inside (strictly
inside) the sector [a,b], if its graph is bounded by (strictly contained
within) the lines y = ax and y = bx. Thus, the effective gain of this non-
linearity is between a and b. As an example, the operation of roundoff
is inside the sector [0,2] (see [C-3,9] for the sector characteristics of
other quantizers). To indicate how sector nonlinearity conditions can be
used, consider (C.2) with G_1 specified by a stable discrete time transfer
function G(z), and G_2 a memoryless nonlinearity, f, assumed to be inside
the sector [0,k]. In this case, the general sector I/O stability theorem
reduces to showing that $(G_1 + \frac{1}{k})$ is a passive system, and, as developed
in [C-19,30], this will be the case if and only if $G(z) + \frac{1}{k}$ is positive
real.

$$Re(G(e^{j\omega})) + \frac{1}{k} \geq 0 \qquad \omega \, \varepsilon \, [0, \, 2\pi) \qquad\qquad (C.8)$$

which is Tsypkin's stability criterion [C-33].

A variant of this type of result involves the use of multipliers [A-19]
in which one modifies the feedback system of (C.2) by inserting a second
system in the forward path and its inverse in the feedback path. One can
then apply the basic stability results to the modified G_1 and G_2. In this

manner one obtains Popov's stability condition [C-16] and the discrete-time analog due to Tsypkin [C-12, 20]: suppose f is nondecreasing and is strictly inside the sector [0,k]. Then the feedback system is finite gain I/O stable if there exists an $\alpha \geq 0$ such that

$$\text{Re}[(1 + \alpha(1 - e^{-j\omega})) \, G(e^{j\omega})] + \frac{1}{k} \geq 0 \qquad \forall \, \omega \, \varepsilon \, [0, \, 2\pi] \qquad (C.9)$$

Claasen, et al. [C-9] have obtained direct analogs of (C.8) and (C.9) for the absence of limit cycles of period N:

$$\text{Re}[G(e^{j2\pi\ell/N}] + \frac{1}{k} > 0, \qquad \ell = 0, \, 1, \, \ldots, \, N-1 \qquad (C.10)$$

or the existence of $\alpha_p \geq 0$ such that

$$\text{Re}\{[1 + \sum_{p=1}^{N-1} \alpha_p (1 - e^{j2\pi\ell p/N})]G(e^{j2\pi\ell/N})\} + \frac{1}{k} > 0 \qquad (C.11)$$

Here f is inside the sector [0,k] and is also nondecreasing in the case of (C.11). The proofs given in [C-9] rely heavily on the passivity relations (C.10), (C.11) and an application of Parseval's theorem in order to contradict the existence of a limit cycle of period N. This last step involves the assumed periodicity in a crucial way, but the application of Parseval and the use of the positive real relationship (C.10) is very reminiscent of stability arguments in feedback control theory [C-19]. In the proof of (C.11) the monotonicity of f is used in conjunction with a version of the rearrangement inequality [C-18,19] which has also been used to study stability of feedback control systems.

As mentioned at the end of the preceding subsection, many frequency domain results can be derived with Lyapunov-type arguments. We have also

seen in this subsection that many of these results can be derived via passivity arguments. Clearly the two are related, and the crucial result that leads to this relationship is the Kalman-Yacubovich-Popov lemma [C-27,28,30], which relates the positive realness of certain transfer functions to the existence of solutions to particular matrix equalities and inequalities. Kalman [C-28] utilized this result to obtain a Lyapunov-type proof of the Popov criterion, and Szego [C-27] used a discrete-time version to obtain a Lyapunov-theoretic proof of Tsypkin's criterion. We also note that the positive real lemma plays a crucial role in several other problem areas including the stochastic realization and spectral factorization problem [B-13] and the study of algebraic Riccati equation [C-29].

Finally, we note that many of these passivity-Lyapunov results have instability counterparts (e.g., see [C-1,17]). Such results may be useful in developing sufficient conditions for the existence of limit cycles.

In this section we have considered some of the aspects of stability theory that point out the relationship among the techniques, goals, and results of researchers in both disciplines. As we have seen, many of the results in the two disciplines involve the use of very similar mathematical tools. On the other hand, the perspectives and goals of researchers in the two fields are somewhat different. The development of a mutual understanding of these perspectives and goals can only benefit researchers in both fields and is in fact absolutely crucial for the successful study of certain problems. For example, in the implementation of digital control

systems one must come to grips with problems introduced by quantization.
Digital controller limit cycles at frequencies near the resonances of the
plant being controlled can lead to serious problems. In addition, the use
of a digital filter in a feedback control loop creates new quantization
analysis problems. Finite arithmetic limit cycles can occur only in recur-
sive (infinite impulse response) filters. However, if a nonrecursive (finite
impulse response) filter is used in a feedback control system, quantization
errors it produces can lead to limit cycles of the closed-loop system [C-31].
How can one analyze this situation, and how does one take quantization
effects into account in digital control system design? Questions such as
these await future investigation.

D. <u>Multiparameter Systems, Distributed Processes, and Random Fields</u>

A growing interest has developed over the past few years into problems involving signals and systems that depend on more than one independent variable. In this section we consider several problem areas involving multiparameter signals and systems in order to examine some of the key issues that arise. For an up-to-date view of some of the research in this area, we refer the reader to the recent special issue of the <u>Proceedings of the IEEE</u> [D-49].

D.1 <u>Two Dimensional Systems and Filters</u>

In analogy with the 1-D case, a 2-D linear shift invariant (LSI) system can be described by a convolution of the input $x(m,n)$ and the unit impulse response $L(m,n)$. Alternatively, taking 2-D z-transforms, we obtain

$$Y(z_1,z_2) = H(z_1,z_2)X(z_1,z_2) \tag{D.1}$$

Of special interest are the rational system functions, $H=A/B$, which arise from 2-D difference equations such as

$$\sum_{(k,\ell)\epsilon I_2} b(k,\ell)y(m-k,n-\ell) = \sum_{(k,\ell)\epsilon I_1} a(k,\ell)x(m-k,n-\ell) \tag{D.2}$$

Here I_1, I_2 are finite sets of pairs of integers.

Let us first discuss the problem of recursion. Given the equation (D.2), we want to use it to calculate the next output given previous outputs and the input. Unlike the 1-D case, in which the index n has the interpretation of time, in the 2-D case, in general, it is not clear what "next" or "previous" mean. In fact, just given (D.2) it is not clear that there is any definition of next or previous that will allow us to compute y(m,n) recursively. Dudgeon [D-1], Pistor [D-12], and Ekstrom and Woods [D-28] have studied this problem in great detail. Let us consider one of the most important special cases of (D.2) in which A=1 and b has its support as indicated in Figure D.1. We then have

$$y(m,n) = - \frac{1}{b(0,0)} \sum_{\substack{k=0 \\ (k,\ell) \neq (0,0)}}^{M} \sum_{\ell=0}^{N} b(k,\ell)y(m-k,n-\ell) + \frac{1}{b(0,0)} x(m,n) \qquad (D.3)$$

Note that from (D.3) and the figure, it is evident that we must store values of y(k,ℓ) for (k,ℓ) to the south and west of the domain over which we wish to calculate y. If this domain is infinite in either direction, the required storage is also infinite. In fact the required storage grows linearly as we increase the domain in either direction (see [D-1] for details). Thus storage requirements in 2-D depend on far more than the order (M,N) of the filter.

We also find that the storage requirements depend on the sequencing of the recursion. Several directions of recursion are indicated in Figure D.2. Each recursion calls for its own sequence of data accessing and discarding. The N and E recursions appear to have particularly
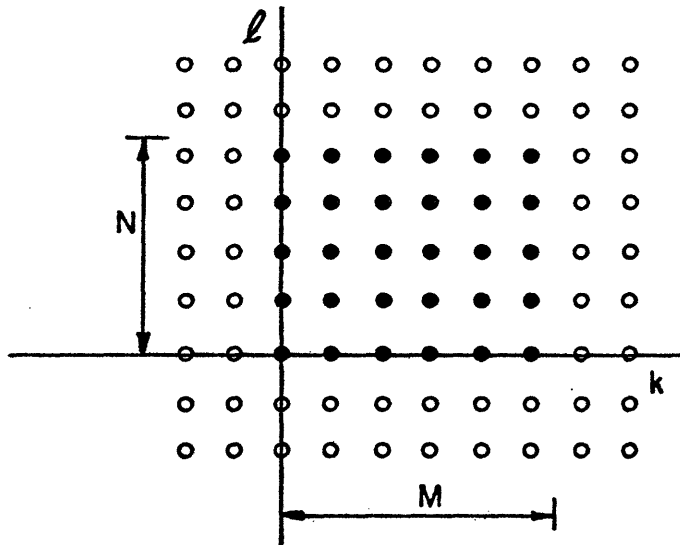
Figure D.1   Support of a First Quadrant or "Northeast" (NE) Function.
(Possible nonzero locations are indicated by solid dots.)

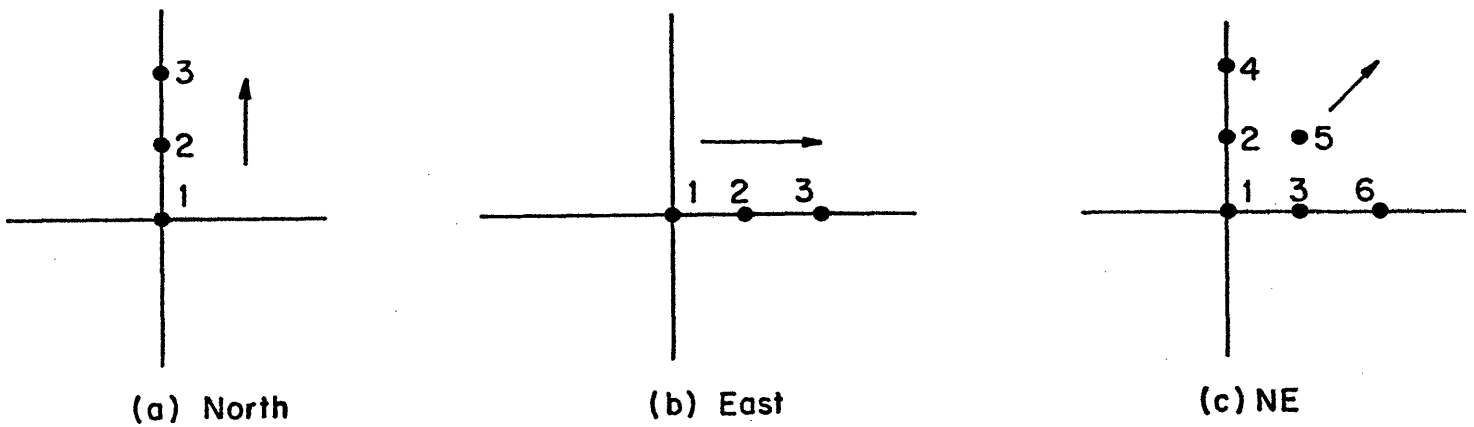(a) North   (b) East   (c) NE

Figure D.2   Several Possible Directions of Recursion for (D.9)

simple sequencing rules, but the data must be processed serially. On

the other hand, the NE recursion has a more complex sequencing but

leads to the possibility of parallel computation, since, for example,

points 4, 5, and 6 can be calculated simultaneously. The possible directions

for recursion and potential use of parallel computation can be determined

with the aid of a conceptual device--the precedence relation [D-26],

which partially orders points with the rule $(m,n) \prec (\ell,k)$ if $y(m,n)$

must be calculated before we can calculate $y(\ell,k)$.

Let us now return to the question of recursibility. Clearly the

picture is symmetric--i.e., we can have NW, SE, and SW recursions, with

$b(k,\ell)$ restricted to be a function on the corresponding quadrant. However,

as shown by Dudgeon [D-1], this by no means exhausts the possibilities

for recursion. In addition to the one quadrant functions, we can obtain

recursive difference equations with $b(k,\ell)$'s that are one-sided [D-1].

In this case the support of b is as in Figure D.3, and we can calculate

$y(m,n)$ column by column, recursing to the north and then shifting to the

next column to the east.

Let us make another connection with 1-D processing. Suppose that

one of the two indices, say m, has the interpretation as time. Then

one might think of $y(m,n)$ and $x(m,n)$ as (1-D) spatially distributed

processes that evolve in time. Temporal causality might then correspond

to the support of b in Figure D.3 being modified by deleting the points

on the positive n axis, yielding a "strictly" one-sided function. In

this case, one could define the "state" of the system, and it is clear

that this "state" will be finite dimensional only if the range of n is

bounded, which is precisely when the required storage for the 2-D
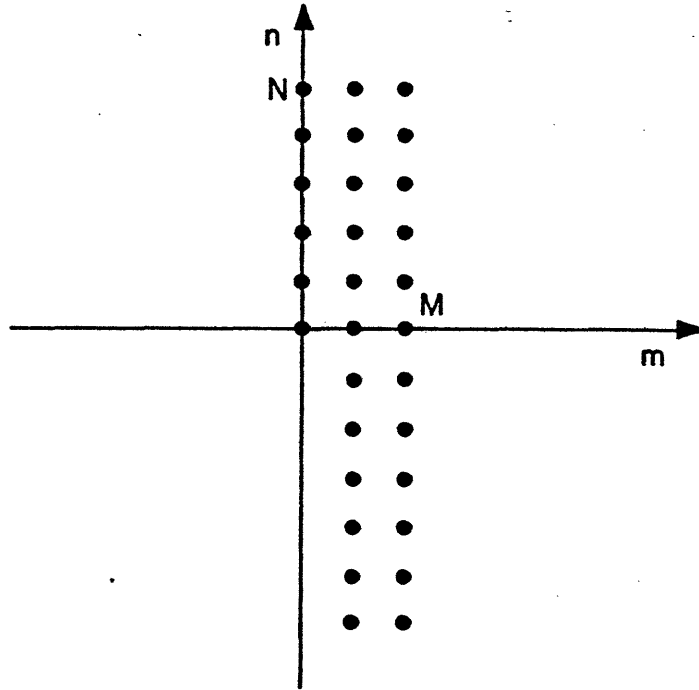
recursion is finite.

Figure D.3   Support of a One-Sided Function

As mentioned earlier, the ability to solve a 2-D difference equation recursively leads directly to the definition of a partial order on the part of the 2-D grid over which we wish to solve the equation. Given this precedence relation, one then has some freedom in deciding how to sequence the calculations. Specifically, if we think of a sequence of calculations as determining a total order on the part of the 2-D grid of interest, all we require is that this total order be compatible with the precedence relation. Once we have such a total order, we can either view this as transforming 1-D filters into 2-D filters or vice versa [D-15]. One widely used order is the line-scan (D-1,5,15]:

$$(i,j) < (\ell,k) \text{ if } i<\ell \text{ or } i=\ell \text{ and } j<k$$

Assuming we are interested only in lines of finite extent, we can readily see one of the problems with this order and with orders in general. If we attempt to process the order input data with a 1-D LSI system, the resulting 2-D processing is not SI, essentially because our ordering has placed the last point on one line "next" to the first point on the next.

We close our discussion of 2-D orders and precedence relations by noting that these very same issues arise naturally in certain feedback control problems. Ho and Chu [D-21] consider optimal control problems in which one has a set of decision makers who base their decisions on certain observed data. Ho and Chu define a precedence relation

among decisions: $j \prec i$ if the decision of j affects the observation of

i. They assume that this is a partial order--i.e. that if $j \prec i$, we

cannot have $i \prec j$ (this is precisely the condition needed for recursibility

of 2D filters). Then, under a "partially nested information condition"--

if $j \prec i$, then i's observation includes knowledge of j's observation--they

solve an optimal control problem. Witsenhausen [D-22] has also studied

this partial order and has pointed out that if one totally orders the

set of decision makers in a way compatible with the precedence relation,

one can then define the state evolution of the system. Hence we see

that there may be many possible sets of states corresponding to different

compatible total orders (just as storage requirements vary with the choice

of recursion).

In the preceding discussion we have seen that the presence of a

partial order as opposed to the usual 1-D total order leads to some

complications. New difficulties are also encountered in the study of

the stability of recursive 2-D filters [D-1,13]. As in the 1-D case,

the stability of the filter depends on the direction of recursion,

and there are many more possibilities in 2-D. In addition, although

there are analogs of results such as those in 1-D that test to see

if all poles are inside the unit circle [D-1,9], the required calculations

are far more complex. This increase in complexity also arises in the

related problem of the stabilization of a given 2-D system function

while keeping the magnitude of the frequency response unchanged. In

1-D this is done easily by shifting those poles that lie outside the

unit circle, but this cannot be done that easily in 2-D, since we cannot

factor 2-D polynomials.*

Another stabilization approach in 1-D is spectral factorization--
i.e. we write a given rational H(z) as the product of two pieces,
$H_E(z)$ and $H_W(z)$, where H_E has all its poles inside the unit circle (and
hence is stable if used to process inputs in the eastern direction) and
H_W has all its poles outside the unit circle (stable to the west).
Thus, in 2-D, one is tempted to seek factorizations into four stable
quadrant filters [D-12] or into two stable half-plane filters [D-1,16,28]
much like the 1-D case. Such techniques have been developed using
2-D cepstral analysis, and we refer the reader to the references. We
do note that the lack of a fundamental theorem of algebra does mean
that the factors in these factorizations will not in general have
finite order denominators.

A final stabilization procedure is based on the guaranteed stabi-
lity in 1-D of least squares inverses. The least squares inverse (LSI)
is obtained using exactly the methodology one brings into play in
performing linear prediction of speech. Given the denominator B and
its inverse transform b, one seeks a finite extent impulse response p
that approximates the convolutional inverse of b by choosing the
coefficients in p to minimize the sum of the squares of the difference
between b*p and the unit impulse. In 1-D, one has the guarantee that
p is minimum phase (i.e. that the all pole model 1/p is stable). In
[D-13] Shanks, et al., conjectured that this minimum phase property holds
in 2-D. Under this assumption, they proposed the use of a double least
squares inverse to stabilize and unstable denominator of a NE filter.

_____

*This is often referred to as the "absence of the fundamental theorem
 of algebra" for multivariable polynomials (see, for example, [D-1]).

Using this design procedure, numerous 2-D filters have been designed.
Unfortunately, Genin and Kamp [D-53] have recently shown that this
conjecture is false in 2-D if one constrains oneself to quarter-plane
filters (although it is true in certain restricted cases [D-50]). On
the other hand, Marzetta [D-43] has shown that the desired minimum
phase property does hold if the least squares inverse problem is posed
in terms of half-plane filters. We will return to this point again
later.

As in the 1-D case, a critical question in the design of 2-D IIR
filters is the existence of limit cycles and the effect of roundoff
noise on the filter output. The results in [D-10] on the existence of
horizontal, vertical, and noninteracting diagonal limit cycles parallel
1-D results. Open questions involve the intriguing question of whether
one can extend any of the other techniques discussed in Section C.
Do the passivity-Tsypskin-positive real-frequency domain results of
Claasen, et al., [C-9] and others extend to the 2-D case? What about
the Lyapunov techniques of Willson [C-2]? Of course in this case one
would need 2-D state space models and a 2-D Lyapunov theory, which in
itself might be of interest in providing a method to test for stability
of 2-D LSI systems even with perfect arithmetic.

The analysis of roundoff noise in 2-D filters can be carried out
much as for 1-D filters, but another open question concerns the exten-
sion of the covariance noise analysis method described in Section A
for 1-D roundoff analysis. Again one would need a state space model
in order to consider this question.

D.2 Two-Dimensional State Space Models

In addition to 1-D state space descriptions for recursively
ordered 2-D systems [D-26], some work has been done on the analysis
of 2-D state space models. Roesser [D-27] considers the NE model

$$v(i+1,j) = A_1 v(i,j) + A_2 h(i,j) + B_1 x(i,j)$$
$$h(i,j+1) = A_3 v(i,j) + A_4 h(i,j) + B_2 x(i,j) \qquad (D.4)$$
$$y(i,j) = C_1 v(i,j) + C_2 h(i,j) + Dx(i,j)$$

here x is the input, y is the output, and v and h together play the
role of a "state" variable, "camping" vertical and horizontal information,
respectively. Given this model, Roesser considers several issues,
including a variation of constants formula to solve (D.4), a 2-D
version of the Cayley-Hamilton Theorem, which in turn is used to obtain
an efficient method for computing the transition matrix, and the notions
of controllability and observability. In obtaining his algorithm for
recursively computing the transition matrix via the Cayley-Hamilton
theorem, Roesser used the notion of 2-D eigenvalues in a crucial manner,
and in the usual nonfactorizable case the calculation of zeroes of a
characteristic polynomial is extremely difficult. This not only
complicates his transition matrix algorithm, but it makes stability tests
more difficult, as we have already mentioned. Furthermore, the model
(D.4) is limited to quadrant-causal systems. This is perfectly reasonable
for the study of quadrant-recursive filters, but its value for the
analysis of other 2-D signals is unclear. For example, Roesser mentions

the possibility of a 2-D filtering theory, where (D.4) plays the role

of a "spatial shaping filter." As Ekstrom and Woods [D-28] point out,

one cannot obtain arbitrary spectra from a NE shaping filter. Hence,

one may need two such filters, as well as a method for modelling the

spectra of the signal field. Finally, we note that Roesser's "state"

$(v(i,j),h(i,j))$ might better be termed a "local state" [D-24]. As we

saw earlier, in recursively solving 2-D equations, the required amount

of storage in general depends on the size of the arrays of interest,

while dimensions of v and h correspond to the order of the system, as

in (D.3).

Issues of this type have also been considered by Fornasini and

Marchesini [D-24]. They consider local NE state space descriptions of

the form

$$x(m+1,n+1) = A_0 x(m,n) + A_1 x(m+1,n) + A_2 x(m,n+1) + Bu(m,n)$$

$$y(m,n) = Cx(m,n) \tag{D.5}$$

They show that a NE IIR filter can be realized as in (D.5) if and

only if the transform of the impulse response is rational. The "if"

part of this result involves the construction of a realization that

is a generalization of the 1-D "standard controllable form." Having

such realizations, attention naturally focusses on minimality. This

leads directly to the notions of (local) controllability and observability,

with finite rank conditions for these properties being developed in a

manner analogous to that of Roesser. The main minimality result of

Marchesini and Fornasini is that minimality implies local controllability
and observability but that local controllability and observability do not
imply minimality.

Attasi [D-23] has studied a special case of (D.5), in which

$$A_0 = -A_1 A_2 = -A_2 A_1 \qquad\qquad (D.6)$$

In this case, the system transfer function is separable $(H(z_1,z_2)=H_1(z_1)H_2(z_2))$,
and, as shown in [D-25], this is the only case in which the global
state is finite dimensional. As any FIR filter can be realized by
(D.5), (D.6), any stable impulse response can be approximated arbitrarily
closely by a system of this form. This, of course, is neither startling
nor necessarily very useful, since the dimension of the resulting state-
space system may be extremely large. Having this framework, Attasi defines
dual notions of local controllability and observability and derives
conditions somewhat simpler than in [D-24,27] because of the assumed
separability. Attasi also considers minimal realizations, obtains a
state space decomposition result and minimal realization algorithm much
like those in 1-D, and shows that minimality implies controllability and
observability. He also proves the converse of this last result, but
this is only true if one looks for the minimal realization in the
class of models satisfying (D.6). We refer the reader to [D-24] for an
example illustrating these points.

Undoubtedly the major contribution of Attasi's work is that he
did something with his models, and we will discuss his filtering
results in the next subsection. He also developed a 2-D Lyapunov
equation, which he used in a generalization of the "invariance principle,"

[D-45]. The exact implication of this result for 2-D stability theory

and its potential utility in such areas as limit cycle analysis remain

as questions for further work. Attasi also considers systems as in

(D.5), (D.6) which are driven by white noise. Again he obtains a 2-D

Lyapunov equation for the state covariance, and this result may be of

some value in performing roundoff noise analysis for 2-D filters. In

addition, he develops a stochastic realization theory that exactly

parallels the 1-D case with one rather surprising exception, in that,

unlike the 1-D case, in the 2-D case the stochastic realization is

essentially unique. This is due primarily to the additional constraints

imposed by the fact that we use a single quadrant shaping filter.

Another novel feature of Attasi's development is the necessity for

using non-square factors--i.e. to perform the required factorization

$$S(z_1, z_2) = H(z_1, z_2)H'(z_1^{-1}, z_2^{-1}) \tag{D.7}$$

where H is NE causal and of the form (D.5), (D.6), one must consider

rectangular factors. For example, if y is a scalar process, then

H in general must be 1xm, and, in fact, the aforementioned uniqueness

result fixes the value of m.

Recently, Morf, et al., [D-31] have made several noteworthy

contributions to 2-D state space theory. They consider the properties

of polynomial and rational matrices in two variables in order to

generalize the scalar 2-D polynomial results of Bose [D-30] and the

matrix 1-D polynomial results of Rosenbrock [D-32] and Wolovich [D-33].

The concepts of local controllability and observability for the Roesser

model are explored in [D-31], and the authors point out that these

conditions neither imply nor are implied by the minimality of the

realization (this is done with several instructive examples). To obtain

notions of controllability and observability that are equivalent to

to minimality, Morf, et al., generalize the approach of Rosenbrock,

and this leads to the notions of modal controllability and observability

and a related concept of minimality. In this setting the existence of

minimal realizations becomes a difficult. problem, and one may not even

exist if we restrict ourselves to systems with real parameters. In

related work, Sontag [D-29] has also found realizations of lower

dimension than those proposed by Fornasini and Marchesini, and he has

shown that minimal realizations need not be unique up to a change of

basis. All of these facts indicate that the 2-D state space model is

an extremely complex one and offers some extremely difficult mathematical

and conceptual problems. It remains to be seen whether any of these state

models and realization theories can provide a useful framework for

solving 2-D analysis and synthesis problems.

D.3 Image Processing, Random Fields, and Space-Time Processes

Digital processing of images for data compression, noise removal,

or enhancement is one of the major areas of applications of 2-D digital

signal processing techniques. In addition, image processing has spurred

a great deal of work in the analysis of spatially-distributed stochastic

variables--random fields. Let $g(i,j)$ denote the image radiant energy

as a function of two discrete spatial variables, where, for the time being, we will assume that the system is free of noise. The image results from an image formation process that transforms the original radiant energy $f(i,j)$ into the observed image. A general model that is often used for the image formation process is

$$g(i,j) = \sum_{k,\ell=1}^{N} h(i,j,k,\ell) f(k,\ell) \qquad i,j=1,\ldots,N \qquad (D.8)$$

where h is the point-spread function (PSF), which models the smoothing and blur that take place in the image formation process [D-4,19,46]. Note that one important case of (D.8) is the shift-invariant case, in which h depends only on $i-k$ and $j-\ell$. In this case (D.8) is a 2-D convolution.

In addition to the image formation process, one must take into account the process of image recording and storing. Several noise-corrupted nonlinear image models have been developed [D-19,46] for this; however, as discussed in [D-46], often one may be able to justify the use of an additive noise model

$$q(i,j) = g(i,j) + v(i,j) \qquad (D.9)$$

where v is an additive noise process. We now turn our attention to the analysis of this model.

At various points in this development, it will be more convenient to view f,g,q, and v as vectors by performing a scan (lexicographic)

ordering, in which case we write q=Hf + v, where H is an $N^2 \times N^2$ matrix

formed from the PSF. Examination of (D.8) yields that H is an NxN

matrix of NxN blocks $\{H_{ij}\}$, where the (m,n) element of H_{ij} is

h(i,m,j,n). If the imaging system is shift-invariant, it is readily

seen that H is block Toeplitz, and, in fact, each of the blocks is

itself a Toeplitz matrix. Note also that if h is separable, then

$$h(i,j,m,n) = h_1(i,m)h_2(j,n) \text{ and } H = A_1 \otimes A_2$$

where $\otimes$ denotes the tensor or Kronecker product, and A_i is an

NxN matrix obtainable from h_i.

It is evident from the preceding development that probabilistic

and statistical methods must play some role in image processing. In

this context, f,g,v, and perhaps h are random fields. For now we

consider such a random field s(i,j) to be characterized by its mean

$\bar{s}(i,j)$ and its covariance

$$r(i,j,m,n) = E\{[s(i,j)-\bar{s}(i,j)][s(m,n)-\bar{s}(m,n)]\} \tag{D.10}$$

The field will be called (wide-sense) stationary if

$$r(i,j,m,n) = r(i-m,j-n) \tag{D.11}$$

Note that if s is ordered lexicographically, then its covariance R

is the N^2 x N^2 matrix obtained from r in the same manner that H is

obtained from the PSF h. We also observe that R is block Toeplitz

with Toeplitz blocks if s is stationary.

One important problem in image processing is the efficient repre-

sentation of images for storage or transmission [D-46,47]. One well-

known method for obtaining a less redundant representation of an image

is the Karhunen-Loeve transform [D-46], which involves the diagonalization

of R. However, in general, this transform involves exorbitant amounts

of computation. There are, however, several special cases in which this

transform can be calculated efficiently. One of these, motivated by

similar analysis performed by Hunt [D-46] and Andrews and Hunt [D-19],

is quite instructive. Suppose that s is stationary, and that any

particular pixel is uncorrelated with ones some distance d away. Then

the block Toeplitz covariance matrix is nonzero only near the

main diagonal (and the same can be said for each of the blocks).

We now modify R and its blocks, to make R block circulant with circulant

blocks. A block circulant matrix is block Toeplitz with each row a

cyclic shift to the right of the preceding one, where the last block

on the right of one row becomes the first block on the left in the

next row. This merely means replacing some of the zeroes in R with

nonzero entries. Intuitively, imagining the image as a flat array, we

have connected opposite edges, first to create a cylinder, and then a

torus. The reason for making this approximation is that the matrix

of eigenvectors of R_c, the circulant modification of R, can be computed

efficiently using the fast Fourier transform (see [D-19, Z-1]), and

thus the Karhunen-Loeve expansion can be performed quickly,.

As discussed in Section B, one of the most widely used coding or compression schemes for 1-D time series, such as speech, is linear prediction, in which we design a one-step predictor or inverse whitening filter for the time series. This method has several appealing features in 1-D--it is efficient (if one uses the Levinson algorithm), it leads to recursive coding and decoding algorithms, and it yields excellent performance. In 2-D the situation is not as clear. What direction do we predict in and what old data do we use to do the prections? At this time, answers to these questions are beginning to be formed. Genin and Kamp [D-53] have shown that NE predictors need not be minimum phase, and Marzetta [D-43] has provided an argument for why this is in fact the case. Specifically, in 1-D, we are guaranteed that the optimal predictor for $y(n)$ based on $y(n-1)$, $y(n-2),\ldots,y(n-r)$ is necessarily minimum phase; however, if we skip some points in the past--e.g., if we predict $y(n)$ based on $y(n-1),y(n-2)$, and $y(n-4)$-- the optimal predictor may not be minimum phase. Marzetta points out that NE predictors do skip points. For example, consider the predictor of $y(m,n)$ based on $y(m-1,n)$, $y(m,n-1)$, and $y(m-1,n-1)$. If we totally order the points in the plane in a fashion compatible with the partial order for calculating points recursively to the NE, then $(m-1,n)$, $(m,n-1)$, and $(m-1,n-1)$ will never be the three immediate predecessors of (m,n). Thus, just as in 1-D, there is no reason to expect the optimal predictor to be minimum phase. Marzetta then points out that if we don't skip points--i.e. if we use a full half-plane predictor-- we do get the minimum phase properties, Levinson-type algorithms

involving reflection coefficients, etc. Note that this predictor is primarily of conceptual interest, since the predictor involves the incorporation of an entire, infinite extent column before any points in the preceding column may be included. We refer the reader to [D-43] for details and for practical, suboptimal methods which also have the minimum phase property. We also refer the reader to [D-51] for another generalization of the Levinson algorithm to 2-D.

We now turn our attention to the problem of restoring blurred and noise-corrupted images [D-4,19,46]. A number of nonrecursive methods have been developed for the removal of blur and for noise rejection--inverse filtering, 2-D minimum mean-square error (MMSE) Wiener filtering, etc., and we refer the reader to the survey [D-46] for more on these methods and for further references. We merely point out here that techniques such as the Wiener filter have some difficulties and limitations as image processing systems. To a great extent this is due to the fact that the MMSE criterion is not particularly well-suited to the way in which the human visual system works [D-20]. In particular, the Wiener filter is overly concerned with noise suppression. In addition, in order to make the filter computationally feasible, one often assumes stationarity. This in turn leads to a filter that is insensitive to abrupt changes--i.e. it tends to smooth edges and reduce contrast. On the other hand, in high contrast regions, the human visual system will readily accept more noise in order to obtain greater resolution. Several schemes have been proposed that are aimed at trading-off between the potentially high-resolution, poor noise performance of the inverse filter and the lower-resolution, good noise performance of the Wiener

filter. One of these is the constrained least squares filter, [D-4,19].

Several other observations can be made concerning the processing systems mentioned so far. As mentioned earlier they are nonrecursive and in principle require the block processing of the entire image or substantial sections of the image. Hence the computational burden of these schemes can be quite high. In 1-D, one find that recursive methods are often preferable to nonrecursive ones because of their computational advantages. As discussed in [D-19] the 1-D Kalman filter offers great computational savings over nonrecursive methods, and an appealing question is the extension of such filters to 2-D. Anyone familiar with 1-D Kalman filtering theory realizes that the design of the filter relies heavily on a dynamic representation of the received signal. Hence, to develop such techniques in 2-D, we need a more complex model of an image than that provided by the mean and covariance. The need for the use of such models is an obvious drawback to this approach, but the potential gains in computational efficiency represent a distinct advantage.

One approach to recursive processing of images involves the 1-D processing of the scan-ordered image (see subsection D.1). This work has been developed by Nahi, Silverman, and their colleagues [D-2,5,7,34]. Suppose we have an image $f(m,n)$ (assumed to be zero mean for convenience) with stationary covariance $r(k,\ell)$, and we observe $q=f+v$ where the additive noise v is, for simplicity, assumed to be zero mean and white, with variance R. We now take the scan ordering of the NxN grid on which q, f, and v are defined. Let us use the same symbols to denote

the resulting 1-D processes. We then have

$$E[f(k)f(\ell)] = S(k,\ell) \qquad k,\ell=1,\ldots,N^2 \qquad (D.12)$$

where $S(k,\ell)$ can be calculated from knowledge of $r(m,n)$. Note that

the scanned image $f(k)$ is <u>not</u> stationary due to the abrupt change that

occurs when the scanner reaches the end of one line and begins the

next. We wish to use Kalman filtering techniques in order to suppress

the noise. In order to do this, we need a state space model for f.

Unfortunately, as pointed out in [D-5], $S(k,\ell)$ does not have the

required separability that is needed in order for such a realization

to exist. Hence, some sort of approximation is needed, and several

have been developed. The simplest of these involves finding a

stationary approximation $R(k)$ to (D.12), much as Manry and Aggarwal

found shift-invariant approximations to the shift-varying scanning

filters they studied in [D-15]. Having $R(k)$, one can then use some

realization procedure to find a Markov model that realizes or approximates

the given correlation function.

We can now obtain an image restoration scheme by direct appli-

cation of Kalman filtering. Several comments are in order. We first

note that the filter has an artificial causality--only the points below

and to the left on the same line affect the estimate of a given pixel.

This can be partially removed by the use of smoothing. With the model

we have developed, this can be done efficiently with two Kalman filters,

scanning in opposite directions and starting at opposite ends of the

image. The resulting estimate still has difficulties because of the

effects at the ends of lines. In this case, one can remove some of these difficulties by transposing the image and performing the same type of processing again—we then have NE, NW, SE, and SW Kalman filters.

The recursive methods discussed so far have assumed that there is no blurring due to a nontrivial PSF. If there is such blurring, essentially we must develop a scan-ordered, 1-D dynamical model for the effect of the blur and then incorporate this model into our Kalman filter. The simplest example of this—motion blur along the direction of the scan—was considered by Aboutalib and Silverman [D-7] (see [D-54] for consideration of more general blurs). Again this system offers computational advantages over nonrecursive schemes, but the restoration system may be very sensitive to errors in the knowledge of the PSF.

The previous technique did not directly use a 2-D recursive model for the image. The first work along this line was that of Habibi [D-6] who considered a 2-D, recursive, auto-regressive shaping filter

$$x(k+1,\ell+1) = \rho_2 x(k+1,\ell) + \rho_1 x(k,\ell+1) - \rho_1\rho_2 x(k,\ell) + \sqrt{(1-\rho_1^2)(1-\rho_2^2)}\, w(k,\ell)$$

$$(D.13)$$

where $w(k,\ell)$ is a white, zero mean, unit variance process. Assuming measurements of the form $y=x+v$, Habibi then developed a suboptimal estimator to estimate $x(k+1,\ell+1)$ based on $\{y(m,n) \,|\, m \leq k, \, n \leq \ell\}$. The suboptimality of Habibi's estimator arises essentially because x is

only the local state, and one needs to estimate the global state for optimal filtering. The most complete study of optimal 2-D Kalman filtering has been performed by Woods and Radewan [D-41]. We assume that we have a one-sided causal dynamic model for the random field

$$x(m,n) = \sum_{k=1}^{M} \sum_{\ell=-M}^{+M} b(k,\ell)x(m-k,n-\ell)$$

$$+ \sum_{\ell=1}^{M} b(0,\ell)x(m,n-\ell) + w(m,n)$$

(D.14)

Suppose we want to estimate $x(m,n)$ given all values of $q=x+v$ in the past, where past is defined relative to the direction of recursion in (D.14). Woods and Radewan point out that this can be done optimally with an extremely high dimensional Kalman filter to estimate the global state of the system, which in this case has dimension on the order of MN (M=order of the filter, N=width of the image).

Optimal line-by-line Kalman filtering for images has also been considered by Attasi [D-23] using a stochastic version of the model discussed in Subsection D.2. Specifically the image is assumed to be generated by a separable vector analog of the model used by Habibi [D-6]

$$x(i,j) = F_1 x(i-1,j) + F_2 x(i,j-1) - F_1 F_2 x(i-1,j-1) + w(i-1,j-1)$$

(D.15)

$$q(i,j) = f(i,j) + v(i,j) = Hx(i,j) + v(i,j)$$

We wish to obtain the optimal estimate $\hat{x}(m,n)$ of $x(m,n)$ given $q(i,j)$ for $i \leq m$ and all j. The optimal estimate in this case consists essentially

of two 1-D operations. Suppose we have $\hat{x}(m-1,n)$ for all n. We first predict ahead one line to obtain

$$\bar{x}(m,n) = F_1\hat{x}(m-1,n), \text{ for all } n \tag{D.16}$$

Note that each of these estimates is calculated independently. We now observe the new line of measurements $q(m,n)$ for all n, and we create the error process and the error measurement

$$e(m,n) = x(m,n) - \bar{x}(m,n) \tag{D.17}$$

$$y(m,n) = q(m,n) - H\bar{x}(m,n) = He(m,n) + v(m,n) \tag{D.18}$$

Thus we have a 1-D estimation problem--estimate $e(m,n)$ for all n, given $y(m,n)$ for all n. Attasi shows that one can obtain a finite dimensional 1-D realization for $e(m,n)$ as a function of n. Hence, this estimation problem reduces to the usual 1-D smoothing problem. The solution consists of two 1-D Kalman filters starting at opposite ends of the line. Furthermore, the optimal smoother can again be implemented with two filters of the type devised by Attasi--one sweeping the columns in order of increasing m, and the other in order of decreasing m. This is reminiscent of the decomposition of zero phase filters into two half-plane filters [D-12,28].

The method of proof used by Attasi involves the taking of z-transforms along the n direction and the treatment of m as a time variable.

Essentially we regard the 2-D system as a high-dimensional (infinite if the domain of n is unbounded) 1-D system, where we use a spatial transform "along" the 1-D state vector in order to simplify the calculations. The key step in Attasi's development is a derivation of a set of Riccati equations, parametrized by the transform variable z, for the power spectral density $S_m(z)$ of $e(m,n)$ considered as a function of n. One can then factor these spectra to obtain the 1-D realizations of the e's.

Methods which transform the data in one direction in order to simplify or to study recursion in the other have also been used in several other image processing schemes. For example, a method very similar to Attasi's was used in [D-52]. In addition, Jain and Angel [D-11] have considered fields described by a nearest neighbor, interpolative equation [D-3]

$$x(m,n) = \alpha_1 [x(m,n+1) + x(m,n-1)] + \alpha_2 [x(m+1,n) + x(m-1,n)] + w(m,n)$$

$$(D.19)$$

Following [D-11], let us consider the vector scan process--i.e. we process an entire line of observed data, $y = x + v$, at a time. Defining the resulting 1-D vector processes x_m, y_m, w_m, and v_m, we can write (D.19] as

$$x_{m+1} = Q x_m - x_{m-1} + w_m \qquad\qquad (D.20)$$

where Q is a symmetric, tridiagonal, Toeplitz matrix. Jain and Angel

point out that the diagonalization of Q, $\Lambda = M'QM$, can be performed

with the aid of the FFT without any approximation. Thus, if we

define the transformed quantities $\bar{x}_m$, $\bar{y}_m$, etc., $(\bar{x}_m = M'x_m)$ we obtain

a set of N decoupled estimation problems, indexed by j (which indexes

the components of the transformed vectors):

$$\bar{x}_{m+1,j} = \lambda_j \bar{x}_{m,j} - \bar{x}_{m-1,j} + \bar{w}_{m,j} \qquad (D.21)$$

$$\bar{y}_{m,j} = \bar{x}_{m,j} + \bar{v}_{m,j} \qquad (D.22)$$

Each of these problems can be solved using a Kalman filter, and we obtain

an extremely efficient implementation--transform the observations,

solve the low-dimensional decoupled estimation problems (perhaps in

parallel), and transform back.

As we have seen, optimal 2-D Kalman filtering algorithms require

large amounts of storage and computation. Thus, the study of suboptimal

estimators that require less computation is of importance. One suboptimal

filter developed in [D-41] is the reduced update Kalman filter. Examining

the optimal filter of Woods and Radewan, we see that the predict cycle

is computationally straightforward--one simply uses the recursion (D.14)

assuming no noise and using preceding estimates. The measurement update

part of the optimal filter, on the other hand, involves updating the

estimates of all of the components of the state. Assuming N>>M, we

expect that a given pixel is most correlated only with a small percentage

of the elements of the state vector. Therefore, it seems reasonable only to update the estimates of those components of the state that are within a certain distance of the point being processed--i.e., we constrain many of the gain elements to be zero and essentially allow only "near neighbor updates."

We have now surveyed a number of nonrecursive and recursive estimation methods. The recursive techniques come with many of the same criticisms that were made concerning nonrecursive filters. They require detailed models of the image statistics and image formation process, and they are essentially based on the MMSE criterion. Hence, they in general will sacrific resolution in favor of noise suppression. In addition, these recursive techniques necessarily affect the image because of the assumed model structure. Some of the recursive techniques allow the inclusion of image blur, while in other cases the extensions to include blur have yet to be developed. Also, we have seen that in some cases optimal Kalman filtering is extremely complex, and suboptimal, but intuitively appealing, recursive filter structures must be used. In other cases we have observed that the use of the structure of the assumed model can lead to extremely efficient optimal estimation algorithms (with the aid of transform techniques). In addition, although work in this area has been limited in extent [D-7,34], the recursive techniques are directly amenable to the analysis of space-varying and nonstationary models. Thus, in spite of the many qualifications, we find enough positive attributes to warrant continued study of recursive techniques for image restoration.

One important area for future work involves the reliance on a priori
information. As mentioned earlier, one often can assume knowledge of
the PSF or can determine it by observing known test scenes through the
imaging system. In other cases, we may not have such information and
must estimate the PSF as well as the image. Thus one important question
concerns the robustness of these techniques in the face of modelling
errors. As mentioned in Section A, techniques do exist for the
sensitivity analysis of 1-D state-space models and 1-D Kalman filters.
Can we extend these methods to the 2-D case, and how well do the 2-D
algorithms perform? In addition, methods abound in 1-D for on-line
parameter identification and adaptive estimation in the presence of
unknown parameters. Can we apply these methods with any success to
the 2-D problem?

A second area of concern is the resolution-noise suppression
tradeoff. As mentioned earlier, the human visual system is willing
to accept more noise in certain regions, such as edges, in order to
improve resolution. Thus, in relatively slowly varying regions of
the image, we would like to remove noise, while where there are abrupt
scene changes or other high frequency fluctuations of interest, we
would prefer to forego noise suppression in favor of resolution [D-4].
In this context an important problem is the detection of edges or
boundaries between different regions in an image. We also note that
in many applications the determination of the boundaries themselves may
be the key issue [D-35]. In recent years a variety of techniques have
been developed for detecting and recognizing various types of boundaries

in 2-D data (as an example, see [D-48]). In 1-D, a variety of recursive

techniques have been developed for the estimation and detection of abrupt

changes in signals [D-44]. These techniques have been successfully applied

in a wide variety of applications, and an important question then is

the extension of methods such as these to the detection of boundaries

in images (see [D-8,35] for some work along these lines).

Throughout this subsection we have seen several examples of 2-D

signal processing problems in which good use is made of the transformation

of the signals obtained by considering them to be 1-D vector time signals,

in which the other independent spatial variable is used to index components

of the vectors. There are, of course, many problems in which the

processes to be studied truly are space-time processes [Z-1], and in

many of these the use of 2-D concepts can often be of great value. One

of the best examples of this type arises in problems of seismic signal

processing [D-17,18,42] in which we observe the time response of the

earth using a spatial array of sensors. Other applications plus several

specific problem formulations are discussed in [Z-1]. In addition, [Z-1]

contains a brief discussion of results and formulations that utilize both

1-D and 2-D tools of stochastic calculus and martingale theory [D-36,37].

Such techniques are in their infancy and work continues to determine their

utility and limitations. We note only that the problem of the lack of

a natural total order in 2-D causes difficulties in extending 1-D

stochastic calculus concepts to 2-D. This is not surprising, given the

several complications that we have already discussed.

Given the several examples described earlier in this section, it is our contention that there is potentially much to be gained by utilizing both the perspective of 2-D signals and systems as well as that of 1-D space-time systems in studying problems of either type. As a final example of how this concept might be used, consider the study of large interconnected systems. In this case we let the spatial variable index subsystem variables which may be vector quantities themselves. A general linear model then is the recursive 2-D model

$$x(k+1,i) = \sum_j A_{ij} x(k,j) + \sum_j B_{ij} u(k,j) + w(k,j) \qquad (D.23)$$

$$y(k,i) = \sum_j C_{ij} x(k,j) + v(k,j) \qquad (D.24)$$

Much as in the analysis of 2-D Kalman filters [D-41], the off-line analysis of such systems (solution of Lyapunov or Riccati equations, for example), as well as the on-line implementation of centralized optimal controllers or estimators, may become prohibitively complex. Indeed the analogy extends farther, as the "nearest neighbor" constrained filter of Woods-Radewan involves precisely the same philosophy as is used in many decentralized control and estimation problems [A-26, D-38].

Let us note that there may be many other useful insights to be drawn from this type of analogy. For example, if the model (D.23), (D.24) falls into the class considered by Attasi, then the optimal

centralized Kalman filter can be efficiently implemented using Attasi's line-by-line optimal filter, which involves dynamics in the transmission of information among subsystems (Attasi's 2 filters "along" each "line"). As a second example, in the case in which (D.23), (D.24) are spatially invariant, Melzer and Kuo [D-39] and Chu [D-40] made good use of the structure by taking spatial transforms in studying centralized and decentralized optimal controllers. Similar analysis is contained in [Z-1] for the case of a finite string with circular symmetry. Much as in the case of block circulant approximations [D-19], this allows us to use FFT techniques to reduce the complexity of the on and off-line calculations for centralized controllers in a manner very similar to that of Jain and Angel [D-11]. In addition, the use of spatial windowing techniques [A-1] to obtain nearest neighbor decentralized control algorithms may allow us to develop useful designs for such circularly symmetric systems.

Concluding Remarks

In this paper we have examined a number of broad research areas that have attracted workers in two disciplines--digital signal processing and control and estimation theory. The goal of this examination has been the acquisition of some perspective on relationships among the questions asked, methods used, and general philosophies adopted by researchers in these disciplines. Upon undertaking this study it was my feeling that such perspective would be extremely valuable in promoting collaboration and interaction among researchers in the two fields. Upon concluding this study, I think that my initial feeling has been thoroughly substantiated. Not only are there numerous examples of questions in one discipline that can benefit from the point of view of the other, but also we have found a number of new issues that naturally arose from combining the two points of view.

Each of the disciplines has its own distinct character, and clearly these will and should be maintained. On the other hand, each discipline can gain from understanding the other. State space methods have their limitations, such as in specifying useful digital algorithms and structures. On the other hand, state space methods provide extremely powerful computer-aided algorithms for noise analysis, optimal design specification, etc. State space ideas also allow one to consider multivariable and time-varying systems. All of these aspects of state space theory may prove of value to people involved in digital signal processing. On the other side, researchers in digital filtering have answered many crucial questions related to turning design specifications into implementable designs. The deep understanding that workers

in digital signal processing have concerning the problems of digital implementation is something that researchers in control and estimation would do well to gain. Thus it seems clear that a mutual understanding will prove beneficial to all concerned.

Numerous questions have been raised and speculation on various possibilities has been made throughout this paper. Whether any of these issues has a useful answer is a question for the future. It is my feeling that many of them do, and it is my hope that others will think so as well.

Acknowledgements

References

A-1.   A.V. Oppenheim and R.W. Schafer, Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.

A-2.   C.-T. Chen, Introduction to Linear System Theory, Holt, Rinehart and Winston, Inc., New York, 1970.

A-3.   R. Yarlagadda, "Network Synthesis--A State-Space Approach," IEEE Trans. Circ. Th., Vol. CT-19, May 1972, pp. 227-232.

A-4.   R.E. Kalman, "Mathematical Description of Linear Dynamical Systems", SIAM J. Contr., Vol. 1, 1963, pp. 152-192.

A-5.   L.M. Silverman and H.E. Meadows, "Equivalence and Synthesis of Time-Variable Linear Systems," Proc. 4th Allerton Conf. Circ. and Sys. Th., 1966, pp. 776-784.

A-6.   J.C. Willems and S.K. Mitter, "Controllability, Observability, Pole Allocation, and State Reconstruction," IEEE Trans. Aut. Contr., Vol. AC-16, No. 6, December 1971, pp. 582-595.

A-7.   A.S. Morse and W.M. Wonham, "Status of Noninteracting Control," IEEE Trans. Auto. Contr., Vol. AC-16, No. 6, December 1971, pp. 568-581.

A-8.   I.B. Rhodes, "A Tutorial Introduction to Estimation and Filtering", IEEE Trans. Aut. Contr., Vol. AC-16, No. 6, December 1971, pp. 688-706.

A-9.   J.B. Cruz, Jr., ed., Feedback Systems, McGraw-Hill, New York, 1972.

A-10.   R.E. Crochiere, "Digital Network Theory and Its Application to the Analysis and Design of Digital Filters," Ph.D. Dissertation, M.I.T., Cambridge, Mass., April 1974.

A-11.   L.R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing, Prentice Hall Inc., Englewood Cliffs, New Jersey, 1975.

A-12.   R.E. Crochiere and A.V. Oppenheim, "Analysis of Linear Digital Networks", Proc. IEEE, Vol. 63, No. 4, April 1975, pp. 581-595.

A-13.   G.K. Roberts, "Consideration of Computer Limitations in Implementing On-Line Controls", M.I.T. ESL Rept. ESL-R-665, Cambridge, Mass., June 1976.

A-14.   NASA Langley Workshop on Realizing the Potentials of Micropro-
        cessors for Aircraft Avionics Systems, NASA Langley Research
        Center, Hampton, Virginia, Jan. 22, 1976.

A-15.   S.R. Parker and P.E. Girard, "Correlated Noise Due to Roundoff
        in Fixed Point Digital Filters," IEEE Trans. Circ. and Sys.,
        Vol. CAS-23, No. 4, April 1976, pp. 204-211.

A-16.   D.L. Kleinman, "On the Stability of Linear Stochastic Systems,"
        IEEE Trans. Aut. Contr., Vol. AC-14, August 1969, pp. 429-430.

A-17.   S.Y. Hwang, "Roundoff Noise in State-Space Digital Filtering:
        A General Analysis", IEEE Trans. on Acous., Speech, and Sig. Proc.,
        Vol. ASSP-24, June 1976, pp. 256-262.

A-18.   C.T. Mullis and R.A. Roberts, "Synthesis of Minimum Roundoff Noise
        Fixed Point Digital Filters," IEEE Trans. on Circ. and Sys.,
        Vol. CAS-23, No. 9, Sept. 1976, pp. 551-562.

A-19.   A. Sripad and D.L. Snyder, "Steady-State, Roundoff-Error Analysis
        in Digital Filters Realized with Fixed-Point Arithmetic," Proc.
        Allerton Conf., Urbana, Illinois, Oct. 1976.

A-20.   A. Sripad and D.L. Snyder, "A Necessary and Sufficient Condition
        for Quantization Errors to be Uniform and White," IEEE Trans.
        Acoust., Speech and Sig. Proc., Vol. ASSP-25, No. 5, Oct. 1977,
        pp. 442-448.

A-21.   R.E. Rink and H.Y. Chong, "Performance of State Regulator Systems
        with Floating-Point Computation," submitted to IEEE Trans. Aut.
        Control.

A-22.   R.E. Rink, "Optimal Utilization of Fixed-Capacity Channels in
        Feedback Control," Automatica, Vol. 9, 1973, pp. 251-255.

A-23.   N.R. Sandell, "Information Flow in Decentralized Systems,"
        in Directions in Large Scale Systems, Y.-C. Ho and S.K. Mitter,
        eds., Plenum Press, New York, 1976.

A-24.   J.B. Knowles and R. Edwards, "Effect of a Finite Word-Length
        Computer in a Sampled-Data Feedback System," Proc. IEE, Vol. 112,
        1965, pp. 1197-1207.

A-25.   J.B. Slaughter, "Quantization Errors in Digital Control Systems,"
        IEEE Trans. Aut. Control, Vol. AC-9, 1964, pp. 70-74.

A-26.   N.R. Sandell, P. Varaiya, and M. Athans, "A Survey of Decentralized
        Control Methods for Large Scale Systems," IEEE Trans. Aut. Control,
        April 1978, to appear.

A-27.   M. Athans, ed., "Special Issue on Linear-Quadratic-Gaussian
        Problem," IEEE Trans. Aut. Control, Vol. AC-16, No. 6, Dec. 1971.

A-28.   R. Curry, et al., "State Estimation with Coarsely Quantized,
        High-Data-Rate Measurements," IEEE Trans. Aerospace and Elec.
        Sys., Vol. AES-11, 1975, pp. 613-620.

A-29.   B.L. Ho and R.E. Kalman, "Effective Construction of Linear State
        Variable Models from Input-Output Functions", Regelungstechnik,
        Vol. 14, 1966, pp. 545-548.

A-30.   L. Silverman, "Realization of Linear Dynamical Systems", IEEE
        Trans. Aut. Control, Vol. AC-16, Dec. 1971, pp. 554-567.

A-31.   L.B. Jackson, A.G. Lindgren, and Y. Kim, "Synthesis of State-
        Space Digital Filters with Low Roundoff Noise and Coefficient
        Sensitivity," Proc. IEEE Internat. Symp. Circ. and Sys.,
        Phoenix, Arizona, 1977.

B-1.   T. Kailath, "Some New Algorithms for Recursive Estimation in
        Constant Linear Systems," IEEE Trans. Inf. Th., Vol. IT-19,
        No. 6, November 1973, pp. 750-760.

B-2.   A. Lindquist, "A New Algorithm for Optimal Filtering of Discrete-
        Time Stationary Processes," SIAM J. Control, Vol. 12, No. 4,
        Nov. 1974, pp. 736-746.

B-3.   A. Lindquist, "On Fredholm Integral Equations, Toeplitz Equations
        and Kalman-Bucy Filtering," Applied Math. and Opt., Vol. 1, No. 4,
        pp. 355-373.

B-4.   T. Kailath, "A View of Three Decades of Linear Filtering Theory,"
        IEEE Trans. Inf. Th., Vol. IT-20, No. 2, March 1974, pp. 146-181.

B-5.   J. Rissanen, "A Fast Algorithm for Optimum Linear Predictors",
        IEEE Trans. Aut. Contr., Vol. AC-18, No. 5, October 1973, p. 555.

B-6.   B.W. Dickinson, M. Morf, and T. Kailath, "A Minimal Realization
        Algorithm for Matrix Sequences," IEEE Trans. Aut. Contr., Vol. AC-19,
        No. 1, February 1974, pp. 31-38.

B-7.   J.L. Massey, "Shift Register Synthesis and BCH Decoding," IEEE
        Trans. Inf. Th., Vol. IT-15, January 1969, pp. 122-127.

B-8.   J. Rissanen, "Recursive Identification of Linear Systems", SIAM
        J. Contr., Vol. 9, No. 3, 1971, pp. 420-430.

B-9.   J. Rissanen and T. Kailath, "Partial Realization of Random Systems",
        Automatica, Vol. 8, No. 4, July 1972, pp. 389-396.

B-10.   T. Kailath, D.Q. Mayne, and R.K. Mehra, eds., "Special Issue on
        System Identification and Time-Series Analysis," IEEE Trans. Aut.
        Contr., Vol. AC-19, No. 6, December 1974.

B-11.   K.J. Astrom and P. Eykhoff, "System Identification--A Survey,"
        Automatica, Vol. 7, 1971, pp. 123-162.

B-12.   T. Soderstrom, L. Ljung, and I. Gustavsson, "A Comparative Study
        of Recursive Identification Methods," Report 7427, Lund Institute
        of Technology, Dept. of Aut. Contr., Lund, Sweden, December 1974.

B-13.   P. Faurre, "Representation of Stochastic Processes," Ph.D. Disser-
        tation, Stanford University, 1967.

B-14.   J.D. Gibson, J.L. Melsa, and S.K. Jones, "Digital Speech Analysis
        Using Sequential Estimation Techniques," IEEE Trans. Acous., Speech
        and Sig. Proc., Vol. ASSP-23, August 1975, pp. 362-369.

B-15.   J. Makhoul, "Linear Prediction: A Tutorial Review", Proc. IEEE, Vol. 63, No. 4, April 1975, pp. 561-580.

B-16.   F. Itakura and S. Saito, "A Statistical Method for Estimation of Speech Spectral Density and Formant Frequencies," Electron. Commun. Japan, Vol. 53-A, No. 1, 1970, pp. 36-43.

B-17.   N. Levinson, "The Wiener RMS Error in Filter Design and Prediction," App. B in N. Weiner, Extrapolation, Interpolation and Smoothing of Stationary Time Series, M.I.T. Press, Cambridge, Mass., 1942.

B-18.   R.A. Wiggins and E.A. Robinson, "Recursive Solutions to the Multi-channel Filtering Problem," J. Geophys. Res., Vol. 70, 1965, pp. 1885-1891.

B-19.   S. Chandrasekhar, Radiative Transfer, Oxford University Press, 1950.

B-20.   J. Casti and E. Tse, "Optimal Linear Filtering Theory and Radiative Transfer: Comparisons and Interconnections," J. Math. Anal. and Appl., Vol. 40, October 1972, pp. 45-54.

B-21.   J.D. Markel and A.H. Gray, Jr., Linear Prediction of Speech, Springer-Verlag, New York, 1976.

B-22.   B.S. Atal, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave," J. Acoust. Soc. Amer., Vol. 47, 1970, p. 65.

B-23.   F. Itakura and S. Saito, "Speech Analysis-Synthesis Based on the Partial Autocorrelation Coefficient," Acoust. Soc. of Japan Meeting, 1969.

B-24.   J.D. Markel, "Basic Formant and F_0 Parameter Extraction from a Digital Inverse Filter Formulation," IEEE Trans. Audio and Electroacoust., Vol. AU-21, 1973, pp. 154-160.

B-25.   P. Faurre and J.P. Marmorat, "Un Algorithme de Realization Stochastique," C.R. Acad. Sci. Paris, Ser. A, Vol. 268, 1969, pp. 978-981.

B-26.   B.W. Dickinson, T. Kailath, and M. Morf, "Canonical Matrix Fraction and State-Space Descriptions for Deterministic and Stochastic Linear Systems," IEEE Trans. Aut. Contr., Vol. AC-19, No. 6, December 1974, pp. 656-667.

B-27.   M. Gevers and T. Kailath, "An Innovations Approach to Least-Squares Estimation - Part VI, Discrete-Time Innovations Representations and Recursive Estimation," IEEE Trans. Aut. Contr., Vol. AC-18, December 1973, pp. 588-600.

B-28. M. Morf, T. Kailath, and B. Dickinson, "General Speech Models and Linear Estimation Theory," in _Speech Recognition_, Academic Press, Inc., New York, 1975.

B-29. B.S. Atal and M.R. Schroeder, "Adaptive Predictive Coding of Speech Signals," _Bell Sys. Tech. J._, 1970, pp. 1973-1986.

B-30. M. Morf, B. Dickinson, T. Kailath, and A. Vieira, "Efficient Solutions of Covariance Equations for Linear Prediction," _IEEE Trans. Acoust., Speech, and Sig. Proc._, Vol. ASSP-25, No. 5, Oct. 1977, pp. 429-433.

B-31. R. Parthasarathy and H. Singh, "On Suboptimal Linear System Reduction," _Proc. IEEE_, Vol. 63, November 1975, pp. 1610-1611.

B-32. J. Durbin, "The Fitting of Time-Series Models," _Rev. Inst. Internat. Statist._, Vol. 28, 1960, pp. 233-244.

B-33. A.V. Oppenheim, ed., "Special Issue on Digital Signal Processing," _Proc. IEEE_, Vol. 63, No. 4, April 1975.

B-34. L. Ljung, T. Kailath, and B. Friedlander, "Scattering Theory and Linear Least Squares Estimation--Part I: Continuous Time Problems," _Proc. IEEE_, Vol. 64, No. 1, January 1976, pp. 131-139.

B-35. B. Friedlander, T. Kailath, and L. Ljung, "Scattering Theory and Linear Least Squares Estimation--Part II: Discrete-Time Problems," _J. Franklin Inst._, January 1976.

B-36. R.E. Kalman, P.L. Falb, and M.A. Arbib, _Topics in Mathematical System Theory_, McGraw-Hill, New York, 1969.

B-37. L.S. DeJong, "Numerical Aspects of Realization Algorithms in Linear Systems Theory," Ph.D. Dissertation, Technical University of Eindhoven, The Netherlands, July 1975.

B-38. R.N. McDonough and W.H. Huggins, "Best Least-Squares Representation of Signals by Exponentials," _IEEE Trans. Aut. Contr._, Vol. AC-13, No. 4, August 1968, pp. 408-412.

B-39. C.S. Burrus and T.W. Parks, "Time Domain Design of Recursive Digital Filters," _IEEE Trans. Audio and Electroacoustics_, Vol. AU-18, No. 2, June 1970, pp. 137-141.

B-40. M. Morf, L. Ljung, and T. Kailath, "Fast Algorithms for Recursive Identification," _IEEE Trans. Aut. Control_, to appear; also, _Proc. 1976 IEEE Conf. on Decision and Control_, Clearwater Beach, Florida, Dec. 1976.

B-41.   A.H. Jazwinski, <u>Stochastic Processes and Filtering Theory</u>,
        Academic Press, New York, 1970.

B-42.   G. Szegö, <u>Orthogonal Polynomials</u>, Amer. Math. Soc., New York, 1959.

C-1.   E. Noldus, "On the Instability of Nonlinear Systems," IEEE
        Trans. Aut. Control, Vol. AC-18, No. 4, August 1973, pp. 404-405.

C-2.   A.N. Willson, Jr., "Limit Cycles Due to Adder Overflow in Digital
        Filters," IEEE Trans. on Circ. Th., Vol. CT-19, No. 4, July 1972,
        pp. 342-346.

C-3.   T.A.C.M. Claasen, W.F.G. Meclenbräuker, and J.B.H. Peek, "Effects
        of Quantization and Overflow in Recursive Digital Filters," IEEE
        Trans. Acoust., Speech, and Sig. Proc., Vol. ASSP-24, No. 6,
        Dec. 1976, pp. 517-529.

C-4.   I.W. Sandberg and J.F. Kaiser, "A Bound on Limit Cycles in Fixed-
        Point Implementation of Digital Filters," IEEE Trans. on Audio
        and Electroacoustics, Vol. AU-20, No. 2, June 1972, pp. 110-112.

C-5.   A. Fettweis, "Pseudopassivity, Sensitivity, and Stability of
        Wave Digital Filters," IEEE Trans. Circ. Th., Vol. CT-19, Nov.
        1972, pp. 668-673.

C-6.   A. Fettweis and K. Meerkötter, "Suppression of Parasite Oscilla-
        tions in Wave Digital Filters," IEEE Trans. on Circ. and Sys.,
        Vol. CAS-22, No. 3, March 1975, pp. 239-246.

C-7.   K. Meerkötter and W. Wegener, "A New Second-Order Digital Filter
        Without Parasitic Oscillations," Archiv für Elektronik und Übertrag.,
        Vol. 29, July-Aug. 1975, pp. 312-314.

C-8.   A.N. Willson, Jr., "Some Effects of Quantization and Adder Over-
        flow on the Forced Response of Digital Filters," Bell Sys. Tech. J.,
        Vol. 51, No. 4, April 1972, pp. 863-887.

C-9.   T. Claasen, W.F.G. Mecklenbräuker, and J.B.H. Peek, "Frequency
        Domain Criteria for the Absence of Zero-Input Limit Cycles in
        Nonlinear Discrete-Time Systems with Applications to Digital
        Filters," IEEE Trans. on Circ. and Sys., Vol. CAS-22, No. 3,
        March 1975, pp. 232-239.

C-10.   Ya.Z. Tsypkin, "Frequency Criteria for the Absolute Stability of
        Nonlinear Sampled Data Systems," Automat. Remote Contr., Vol. 25,
        March 1964, pp. 261-267.

C-11.   Ya.Z. Tsypkin, "Absolute Stability of a Class of Nonlinear Auto-
        matic Sampled Data Systems," Automat. Remote Contr., Vol. 25,
        July 1964, pp. 918-923.

C-12.   Ya.Z. Tsypkin, "A Criterion for Absolute Stability of Automatic
        Pulse-Systems with Monotonic Characteristics of the Nonlinear
        Element," Sov. Phys. Dokl., Vol. 9, October 1964, pp. 263-266.

C-13. S.R. Parker and S.F. Hess, "Limit Cycle Oscillations in Digital Filters," IEEE Trans. Circ. Th., Vol. CT-18, Nov. 1971, pp. 687-697.

C-14. C.A. Desoer, "On the Relation Between Pseudo-Passivity and Hyperstability," IEEE Trans. on Circ. and Sys., Vol. CAS-22, No. 11, November 1975, pp. 897-898.

C-15. V.M. Popov, Hyperstability of Control Systems, Springer-Verlag, New York, 1973.

C-16. R.W. Brockett and J.L. Willems, "Frequency Domain Stability Criteria: Parts I and II," IEEE Trans. on Auto. Contr., Vol. AC-10, 1965, pp. 255-261 and 401-413.

C-17. R.W. Brockett and H.B. Lee, "Frequency Domain Instability Criteria for Time-Varying and Nonlinear Systems," Proc. IEEE, Vol. 55, 1967, pp. 604-619.

C-18. J.C. Willems and R.W. Brockett, "Some New Rearrangement Inequalities Having Application in Stability Analysis," IEEE Trans. Aut. Contr., Vol. AC-13, No. 5, October 1968, pp. 539-549.

C-19. J.C. Willems, The Analysis of Feedback Systems, The M.I.T. Press, Cambridge, Mass., 1971.

C-20. R.W. Brockett, "The Status of Stability Theory for Deterministic Systems," IEEE Trans. Aut. Contr., Vol. AC-11, July 1966, pp. 596-606.

C-21. G. Zames, "On the Input-Output Stability of Nonlinear Time-Varying Feedback Systems, Part I and II," IEEE Trans. Aut. Contr., Vol. AC-11, April 1966, pp. 228-238 and July 1966, pp. 465-477.

C-22. J.L. Willems, Stability Theory of Dynamical Systems, John Wiley and Sons, Inc., New York, 1970.

C-23. J.C. Willems, "Dissipative Dynamical Systems Part I: General Theory," and "Part II: Linear Systems with Quadratic Supply Rates," Archive for Rational Mechanics and Analysis, Vol. 45, No. 5, 1972, pp. 321-343.

C-24. S.R. Parker and S. Yakowitz, "A General Method for Calculating Quantization Error Bounds Due to Roundoff in Multivariable Digital Filters," IEEE Trans. Circ. and Sys., Vol. CAS-22, No. 6, June 1975, pp. 570-572.

C-25. G.W. Johnson, "Upper Bound on Dynamic Quantization Error in Digital Control Systems Via the Direct Method of Liapunov," IEEE Trans. Aut. Contr., Vol. AC-10, No. 4, October 1965, pp. 439-448.

C-26. G.N.T. Lack and G.W. Johnson, "Comments on 'Upper Bound on Dynamic Quantization Error in Digital Control Systems Via the Direct Method of Liapunov,'" IEEE Trans. Aut. Contr., Vol. AC-11, April 1966, pp. 331-334.

C-27. G.P. Szego, "On the Absolute Stability of Sampled-Data Control Systems," Proc. Nat. Acad. Sci., Vol. 256, No. 49, 1963.

C-28. R.E. Kalman, "Lyapunov Functions for the Problem of Lur'e in Automatic Control", Proc. Nat. Acad. Sci., Vol. 49, 1963, pp. 201-205.

C-29. J.C. Willems, "Least Squares Stationary Optimal Control and the Albebraic Riccati Equation," IEEE Trans. Aut. Control, Vol. AC-16, December, 1971, pp. 621-634.

C-30. J.C. Willems, "Mechanisms for the Stability and Instability in Feedback Systems," Proc. IEEE, Vol. 64, No. 1, January 1976, pp. 24-35.

C-31. A. Fettweis and K. Meerkotter, "On Parasite Oscillations in Digital Filters under Looped Conditions," IEEE Trans. on Circ. and Sys., Vol. CAS-24, No. 9, Sept. 1977, pp. 475-481.

C-32. P. Dorato and A.H. Levis, "Optimal Linear Regulations: The Discrete-Time Case," IEEE Trans. Aut. Control, Vol. AC-16, December 1971, pp. 613-620.

C-33. Ya.Z. Tsypkin, "Fundamentals of the Theory of Non-linear Pulse Control Systems," Proc. of 2nd IFAC, Basel, Switzerland, 1963, pp. 172-180.

C-34. B.D.O. Anderson and R.W. Newcomb, "Linear Passive Networks: Functional Theory," Proc. IEEE, Vol. 64, No. 1, January 1976, pp. 72-88.

D-1. R.M. Mersereau and D.E. Dudgeon, "Two-Dimensional Digital Filtering," Proc. IEEE, Vol. 63, No. 4, April 1975, pp. 610-623.

D-2. S.R. Powell and L.M. Silverman, "Modeling of Two-Dimensional Covariance Functions with Application to Image Restoration," IEEE Trans. Aut. Contr., Vol. AC-19, No. 1, Feb. 1974, pp. 8-13.

D-3. J.W. Woods, "Two-Dimensional Discrete Markovian Fields," IEEE Trans. Inf. Th., Vol. IT-18, No. 2, March 1972, pp. 232-240.

D-4. M.M. Sondhi, "Image Restoration: The Removal of Spatially Invariant Degradations," Proc. IEEE, Vol. 60, No. 7, July 1972, pp. 842-853.

D-5. N.E. Nahi, "Role of Recursive Estimation in Statistical Image Enhancement," Proc. IEEE, Vol. 60, No. 7, July 1972, pp. 872-877.

D-6. A. Habibi, "Two-Dimensional Bayesian Estimate of Images," Proc. IEEE, Vol. 60, No. 7, July 1972, pp. 878-883.

D-7. A.O. Aboutalib and L.M. Silverman, "Restoration of Motion Degraded Images," IEEE Trans. Circ. and Sys., Vol. CAS-22, No. 3, March 1975, pp. 278-286.

D-8. N.E. Nahi and A. Habibi, "Decision-Directed Recursive Image Enhancement," IEEE Trans. Circ. and Sys., Vol. CAS-22, No. 3, March 1975, pp. 286-293.

D-9. D.D. Siljak, "Stability Criteria for Two-Variable Polynomials," IEEE Trans. Circ. and Sys., Vol. CAS-22, No. 3, March 1975, pp. 185-189.

D-10. G.A. Maria and M.M. Fahmy, "Limit Cycle Oscillations in a Cascade of Two-Dimensional Digital Filters," IEEE Trans. Circ. and Sys., Vol. CAS-22, No. 10, October 1975, pp. 826-830.

D-11. A.K. Jain and E. Angel, "Image Restoration, Modelling, and Reduction of Dimensionality," IEEE Trans. Comp., Vol. C-23, No. 5, May 1974, pp. 470-476.

D-12. P. Pistor, "Stability Criterion for Recursive Filters," IBM J. Res. Devel., Vol. 18, No. 1, Jan. 1974, pp. 59-71.

D-13. J.L. Shanks, S. Treitel, and J.H. Justice, "Stability and Synthesis of Two-Dimensional Recursive Filters," IEEE Trans. Aud. Electr., Vol. AU-20, No. 2, June 1972, pp. 115-128.

D-14. J.M. Costa and A.N. Venetsanopoulos, "Design of Circularly Symmetric Two-Dimensional Recursive Filters," IEEE Trans. Acous., Speech, and Sig. Proc., Vol. ASSP-22, No. 6, Dec. 1974, pp. 432-443.

D-15.   M.T. Manry and J.K. Aggarwal, "Picture Processing Using One-Dimensional Implementations of Discrete Planar Filters," *IEEE Trans. Acoust., Speech and Sig. Proc.*, Vol. ASSP-22, No. 3, June 1974, pp. 164-173.

D-16.   P. Whittle, "Stochastic Processes in Several Dimensions," *Bull. Int. Stat. Inst.*, Vol. 40, 1963, pp. 974-994.

D-17.   R.L. Sengbush and M.R. Foster, "Design and Application of Optimal Velocity Filters in Seismic Exploration," *IEEE Trans. Comp.*, Vol. C-21, No. 7, July 1972, pp. 648-654.

D-18.   J.M. Davis, "Velocity Analysis: An Application of Deterministic Estimation to Reflection Seismology," *IEEE Trans. Comp.*, Vol. C-21, No. 7, July 1972, pp. 730-734.

D-19.   H.C. Andrews and B.R. Hunt, *Digital Image Restoration*, Prentice Hall, Inc., Englewood Cliffs, N.J., 1977.

D-20.   T.G. Stockham, "Image Processing in the Context of a Visual Model," *Proc. IEEE*, Vol. 60, No. 7, July 1972, pp. 828-842.

D-21.   Y.C. Ho and K.C. Chu, "Team Decision Theory and Information Structures in Optimal Control Problems: Part I," *IEEE Trans. Aut. Contr.*, Vol. AC-17, No. 1, Feb. 1972, pp. 15-22.

D-22.   H.S. Witsenhausen, "Some Remarks on the Concept of State," in *Directions in Large Scale Systems*, Y.-C. Ho and S.K. Mitter, eds., Plenum Press, New York, 1976.

D-23.   S. Attasi, "Modelling and Recursive Estimation for Double Indexed Sequences," in *System Identification: Advances and Case Studies*, R.K. Mehra and D.G. Lainiotis, eds., Academic Press, New York, 1976.

D-24.   E. Fornasini and G. Marchesini, "State-Space Realization Theory of Two-Dimensional Filters," *IEEE Trans. Aut. Contr.*, Vol. AC-21, No. 4, Aug. 1976, pp. 484-492.

D-25.   M. Fliess, "Series Reconnaissables, Rationnelles et Algebriques," *Bull. Sci. Math.*, Vol. 94, 1970, pp. 231-239.

D-26.   D.S.K. Chan, "A Novel Framework for the Description of Realization Structures for 1-D and 2-D Digital Filters, *Proc. Electronics and Aerospace Systems Conf.*, Sept. 1976, Arlington, Virginia; also, Ph.D. Dissertation, M.I.T. Dept. of Elec. Eng. and Comp. Sci., in progress.

D-27.   R.P. Roesser, "A Discrete State-Space Model for Linear Image Processing," *IEEE Trans. Aut. Control*, Vol. AC-20, No. 1, February 1975, pp. 1-10.

D-28.   M.P. Ekstrom and J.W. Woods, "Two-Dimensional Spectral Factoriza-
        tion with Applications in Recursive Digital Filtering," IEEE
        Trans. Acoustics, Speech, and Sig. Proc., Vol. ASSP-24, No. 2,
        April 1976, pp. 115-128.

D-29.   E.D. Sontag, "On Linear Systems and Noncommutative Rings," Mathe-
        matical Systems Theory, Vol. 9, No. 4, 1976, pp. 327-344.

D-30.   N.K. Bose, "Problems and Progress in Multidimensional Systems
        Theory," Proc. IEEE, Vol. 65, No. 6, June 1977, pp. 824-840.

D-31.   M. Morf, et al., "New Results in 2-D Systems Theory, Part I:
        2-D Polynomial Matrices, Factorization, and Coprimeness," and
        "Part II: 2-D State Space Models - Realization and the Notions
        of Controllability, Observability, and Minimality," Proc. IEEE,
        Vol. 65, No. 6, June 1977, pp. 861-872, 945-961.

D-32.   H.H. Rosenbrock, State-Space and Multivariable Theory, J. Wiley,
        New York, 1970.

D-33.   W.A. Wolovich, Linear Multivariable Systems, Springer-Verlag,
        New York, 1974.

D-34.   M.S. Murphy and L.M. Silverman, "Image Model Representation and
        Line-by-Line Recursive Restoration," 1976 IEEE Conf. on Dec. and
        Control, Clearwater Beach, Florida, Dec. 1976.

D-35.   N.E. Nahi and S. Lopez-Mora, "Estimation of Object Boundaries in
        Noisy Images", 1976 IEEE Conf. on Dec. and Control, Clearwater
        Beach, Florida, Dec. 1976.

D-36.   E. Wong and M. Zakai, "Martingales and Stochastic Integrals for
        Processes with a Multidimensional Parameter", Z. Wahrscheinlich-
        keitstheorie, Vol. 29, 1974, pp. 109-122.

D-37.   A.S. Willsky and P.Y. Kam, "Recursive Estimation of Signals Sub-
        ject to Random Propagation Delays," submitted to Appl. Math. and
        Optimization.

D-38.   D.P. Looze, P.K. Houpt, N.R. Sandell, and M. Athans, "On Decen-
        tralized Estimation and Control with Application to Freeway Ramp
        Metering," IEEE Trans. Aut. Control, April 1978, to appear.

D-39.   S.M. Melzer and B.C. Kuo, "Optimal Regulation of Systems Described
        by a Countably Infinite Number of Objects," Automatica, Vol. 7,
        1971, pp. 359-366.

D-40.   K.C. Chu, "Optimal Decentralized Regulation for a String of
        Coupled Systems," IEEE Trans. Aut. Contr., Vol. AC-19, June 1974,
        pp. 243-246.

D-41.   J.W. Woods and C.H. Radewan, "Kalman Filtering in Two-Dimensions," _IEEE Trans. Inf. Th._, Vol. IT-23, No. 4, July 1977, pp. 473-482.

D-42.   L.C. Wood and S. Treitel, "Seismic Signal Processing," _Proc. IEEE_, Vol. 63, No. 4, April 1975, pp. 649-661.

D-43.   T.L. Marzetta, "A Linear Prediction Approach to Two-Dimensional Spectral Factorization and Spectral Estimation," Ph.D. Dissertation, M.I.T., Department of Elec. Eng. and Comp. Sci., Jan. 1978.

D-44.   A.S. Willsky, "A Survey of Failure Detection Methods in Linear Dynamic Systems," _Automatica_, Vol. 12, Nov. 1976, pp. 601-611.

D-45.   J.P. LaSalle, "An Invariance Principle in the Theory of Stability, Differential Equations and Dynamical Systems," in _Differential Equations and Dynamical Systems_, J.K. Hale and J.P. Lasalle, eds., Academic Press, 1967, pp. 277-286.

D-46.   B.R. Hunt, "Digital Image Processing," _Proc. IEEE_, Vol. 63, No. 4, April 1975, pp. 693-708.

D-47.   A. Habibi and G.S. Robinson, "A Survey of Digital Picture Coding," _Computer_, Vol. 7, No. 5, May 1974, pp. 22-34.

D-48.   R. Bajcsy and M. Tavakoli, "Computer Recognition of Roads from Satellite Pictures," _IEEE Trans. on Sys., Man, and Cyb._, Vol. SMC-6, No. 9, Sept. 1976, pp. 623-637.

D-49.   N.K. Bose, "Special Issue on Multidimensional Systems," _Proc. IEEE_, Vol. 65, No. 6, June 1977.

D-50.   E.I. Jury, V.R. Kolavennu, and B.D.O. Anderson, "Stabilization of Certain Two-Dimensional Recursive Digital Filters," _Proc. IEEE_, Vol. 65, No. 6, June 1977, pp. 887-892.

D-51.   J.H. Justice, "A Levinson-Type Algorithm for Two-Dimensional Wiener Filtering Using Bivariate Szegö Polynomials," _Proc. IEEE_, Vol. 65, No. 6, June 1977, pp. 882-886.

D-52.   F.C. Schoute, M.F.ter Horst, and J.C. Willems, "Hierarchic Recursive Image Enhancement," _IEEE Trans. Circ. and Sys._, Vol. CAS-24, No. 2, Feb. 1977, pp. 67-78.

D-53.   Y. Genin and Y. Kamp, "Two-Dimensional Stability and Orthogonal Polynomials on the Hypercircle," _Proc. IEEE_, Vol. 65, No. 6, June 1977, pp. 873-881.

D-54.   A. O. Aboutalib, M. S. Murphy, and L. M. Silverman, "Digital Restoration of Images Degraded by General Motion Blurs," _IEEE Trans. Aut. Control_, Vol. AC-22, No. 3, June 1977, pp. 294-302.

Z-1.   A.S. Willsky, <u>Digital Signal Processing and Control and Estima-tion Theory--Points of Tangency, Areas of Intersection, and Parallel Directions</u>, The M.I.T. Press, Cambridge, Mass., to appear.