

LINEAR AND NONLINEAR DISCRETE FILTERING
FOR CONTINUOUS SYSTEMS

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
MARVIN A. NEEDLER
1971



This is to certify that the

thesis entitled

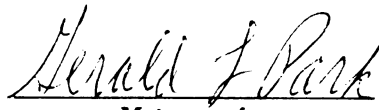
LINEAR AND NONLINEAR DISCRETE FILTERING
FOR CONTINUOUS SYSTEMS

presented by

Marvin A. Needler

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Systems Science


Major professor

Date February 12, 1971

ABSTRACT

LINEAR AND NONLINEAR DISCRETE FILTERING FOR CONTINUOUS SYSTEMS

by

Marvin A. Needler

The subject of this thesis is the synthesis of discrete-time filters for real-time estimation of the state of continuous-time dynamic systems.

The dynamic system is modeled by a set of stochastic differential equations and the stochastic disturbance is modeled as a Brownian motion process. The filtering process is assumed to utilize on-line digital computation and thus a discrete-time recursive filter is developed to estimate the state of the system based on sampled-data measurements.

Starting with the equations for propagation of the probability density function of the state conditioned on previous measurements, the filtering equations are developed both for linear and nonlinear filtering.

In the linear filtering case, two new methods are suggested for reducing computation and increasing accuracy. These methods are illustrated in the digital simulation of state estimation of a steam turbo-generator system. In addition, the methods of divergence compensation are reviewed.

In the nonlinear filtering case, a second-order filter is suggested, and it is used to estimate the state of a reentry body. An original

adaptive filter is mechanized to reduce filter error and is shown to improve the estimation accuracy of the linear and nonlinear filter in reentry estimation. The adaptive filter adapts the filter gain to the statistics of the measurement estimation errors without requiring significant computation and without altering the state covariance computation. Based on the reentry estimation problem, the adaptive filter produces an improved state estimate that approaches the optimum estimate for the given filtering problem.

LINEAR AND NONLINEAR DISCRETE FILTERING FOR CONTINUOUS SYSTEMS

By

Marvin A. Needler

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering and Systems Science

1971

ACKNOWLEDGEMENTS

I would like to express my gratitude to my major professor, Dr. Gerald L. Park, for the guidance and inspiration that he provided throughout my doctoral program. In addition; I am deeply grateful to the other guidance committee members, Dr. Richard C. Dubes, Dr. J. Sutherland Frame, Dr. William L. Kilmer, and Dr. John B. Kreer, for their invaluable contributions to my doctoral program. A special note should be made of the joint effort provided by Dr. Frame in developing Section 3-2 of this thesis.

Support was granted by Consumers' Power Company Cooperative Research Project at Michigan State University for one summer and by the National Science Foundation for one year. This support is gratefully acknowledged.

In addition, I would like to thank Mr. Howard Bell, who introduced me to filtering theory. Also, I am indebted to Mrs. Ann Darrenkamp for typing the final manuscript. And finally, the original manuscript was typed by my wife, Priscilla, whose support made this thesis possible.

TABLE OF CONTENTS

Chapter I	Introduction	1
Chapter II	Analysis and Development of Filtering Theory	6
2-1.	The System Model	6
2-2.	Extrapolation of the Conditional Density Function and Its Moments	8
2-3.	Update of the Conditional Density Function and Its Moments	10
2-4.	The System Disturbance Model	11
Chapter III	Discrete Filtering of a Linear Continuous System	15
3-1.	Linear Filtering	16
3-2.	State and Covariance Extrapolation	18
3-3.	An Example of State Estimation of a Steam Turbo-Generator System	24
3-4.	The Problem of Filter Divergence	28
Chapter IV	Approximate Filtering of Nonlinear Systems	33
4-1.	Nonlinear Filtering	33
4-2.	An Example of a Nonlinear Filtering Problem	37
4-3.	Adaptive Filtering for Nonlinear Systems	44
4-4.	A Discussion and an Example of Adaptive Nonlinear Filtering	49

Chapter V	Conclusions	57
5-1.	Thesis Review	57
5-2.	Results and Conclusions	58
5-3.	Suggestions for Further Research	60
LIST OF REFERENCES		61
APPENDIX A		69
APPENDIX B		72
APPENDIX C		74

LIST OF FIGURES

Figure 3-1.	The Block Diagram of a Steam Turbo-Generator System.	27
Figure 3-2.	The RMS Estimation Error of Frequency Deviation for the Steam Turbo-Generator.	29
Figure 4-1.	Altitude Estimation Error for the First-Order Filter of the Reentry Estimation Problem.	42
Figure 4-2.	Altitude Estimation Error for the Second-Order Filter of the Reentry Estimation Problem.	43
Figure 4-3.	Altitude Estimation Error for the First-Order Adaptive Filter of the Reentry Estimation Problem.	51
Figure 4-4.	Altitude Estimation Error for the Second-Order Adaptive Filter of the Reentry Estimation Problem.	53
Figure 4-5.	RMS Altitude Estimation Error for Four Similar Filters.	54

I INTRODUCTION

The subject of this thesis is the synthesis of discrete-time filters for real-time estimation of the state of continuous-time dynamic systems.

Historically, the theory of estimation dates back to the least-squares orbit estimation of Gauss [GAU1] in 1809. More recently, Wiener [WIE1] and Kolmogorov [KOL1] derived solutions for linear least-squares estimation. Wiener studied continuous-time systems using spectral methods to obtain minimum-variance unbiased estimates of the scalar state by observing a signal where both the state and the observation are jointly wide-sense stationary and ergodic processes. Kolmogorov used a recursive orthonormalization procedure to solve the analogous discrete-time problem.

In 1960, Kalman [KAL1] published a recursive solution for the minimum-variance unbiased estimate of the state of the Markov vector process of a discrete-time system where the system is disturbed by a random vector and the state is observed through a non-invertible transformation. Kalman and Bucy [KAL2, KAL3, BUC1] and Stratonovich [STR1] extended this work to the continuous-time filter and incorporated the problem of observations contaminated with additive random noise. In addition to the filtering problem (estimating concurrently in time to the observations), solutions were developed for the prediction problem

(estimating into the future) and the smoothing problem (estimating after the time of the observations).

During the decade following the derivation of the Kalman-Bucy filter, much effort was involved in further refinement and development of filtering theory. Much of the motivation comes from optimal control theory where it was found that a simple deterministic model would no longer suffice for a realistic model of a system that is contaminated by noise and whose state can not be directly observed [BEL1, SAR1]. The Kalman-Bucy filter was modified to serve in aero-space applications for guidance, navigation, and orbit and reentry estimation [SMI1, PUR1]. In addition, applications in industrial process control, electric power systems and system identification have been developed [SAR1, CHA1]. More recently, applications in urban and freeway traffic flow have received attention.

In a large class of optimal control problems, namely, for linear systems with quadratic performance index and with Gaussian noise, the optimal control law is a feedback law, which is a computed gain times the state [BRY1, LEE1]. In this case, it can be shown that the combined optimal estimation and control scheme is the optimal estimator and the optimal controller found independently--this principle is known as the separation principle [JOS1, POT1, WON1]. In the nonlinear case and/or the non-quadratic performance index case, there is no known separation theorem; nonetheless, the separation principle provides a reasonable, approximate design approach.

Other approaches to the optimal control problem with inaccessible states have been developed. In particular, the methods proposed by

Pearson [PEA1] and Ferguson [FER1] obviate the need for a filter. However, it should be noted that these particular methods are restricted to optimal linear regulators and do not consider the case of noisy measurements or system disturbances.

Linear filtering has been rather completely exploited. Since the appearance of Kalman's original paper using the orthogonal properties of random vectors to minimize the error criterion, other methods such as least squares, maximum likelihood, minimum variance, calculus of variations, and the maximum principle [BRY1, ATH1, HO1] have been used to prove the form of the optimal filter. Other problems such as non-zero mean, cross-correlated, and auto-correlated system and measurement noise have been considered [BRY2]. Filter stability and convergence have also been given considerable attention. In addition, the Kalman-Bucy filter has been extended to the case of nonlinear systems, called the extended Kalman filter, first-order filter, or the quasilinear filter [FRI1, MOW1, SCH2, HO2].

The problems remaining in linear filtering are those that involve mechanizing a filter that uses a suitable model, i.e., does not have detrimental modeling errors, and that is feasible in terms of computational requirements.

Nonlinear filtering theory is not as completely developed as linear filtering theory due to the inherently more complicated structure of nonlinear systems where superposition does not hold. By using a probabilistic approach, it is possible to derive the equations governing the evolution of the conditional probability density function of the state conditioned on the sequence of observations. It has been shown that, in general, the optimal nonlinear filter requires an infinite

number of moments to realize the conditional probability density function [KJS1, KJS2]. However, a finite approximation can be made that may be satisfactory [KJS3, BUC2].

Another problem in nonlinear filtering is finding a global filtering theory that will assure filter convergence. No general convergence theory has been found that is equivalent to that in linear filtering theory. Indeed, the problem is so difficult that proofs of filter convergence are restricted to the scalar case without system disturbances [BIR1].

An approach is suggested in this thesis that improves the convergence properties of both the linear and nonlinear filter whenever modeling errors are prevalent. Previous work [DEN1, PIN1, SMI1] has been concerned with the problem of filter divergence. The idea of adaptive filtering has been suggested by Jazwinski and Dennis [JAZ2, DEN2], whereby an artificial system disturbance is constructed to prevent filter divergence. In this thesis a new approach is suggested to adapt the estimates to the measured statistics of the filtering problem; namely, the filter gain matrix is corrected to provide consistent filter statistics and measurement statistics.

In many problems, it is necessary to estimate the state of a vector-valued continuous-time system. Often, it is not possible to measure continuously, and even if continuous-time measurements are available, only sampled data can be used to allow for the computation time that is necessary even in high-speed digital computers. Thus, the structure that emerges is a discrete-time filter for the continuous-time system. This structure has some of the aspects of both continuous

problems and discrete problems. For the most part, the discrete-filter problem is a sub-problem of the discrete-filter continuous-system problem, whereas the continuous filter problem must account for the continuous observation process.

One important problem involved in a continuous-time system is that of modeling the random processes. Assuming the process to be continuous-valued and composed of stationary independent increments leads naturally to the Brownian motion process. It shall be assumed that this process has a Gaussian distribution function with zero mean. In order to solve differential equations involving the Brownian motion process, the Ito stochastic integral is employed.

To summarize the previous discussion, the purpose of this thesis is to study the synthesis of filters. The discrete-time filter, continuous-time dynamic system is emphasized. In Chapter 1, the general filtering equations are developed, and some results are derived for simulation of the stochastic processes. In Chapter 2, the linear filter equations are given, some new computational methods are presented for reducing computation and increasing accuracy, these methods are applied to a simulation example and finally, the problem of modeling errors is discussed. In Chapter 4, a nonlinear filter is derived and used in a simulation example. Finally, a new statistical method is developed that adapts the nonlinear filter to modeling errors to provide consistent elements for the nonlinear filter.

II. Analysis and Development of Filtering Theory

Underlying the subject of filtering theory is the theory of probability, random variables, and stochastic processes. The central principles of stochastic processes that are needed here are reviewed in Appendix A. (See [FEL1, PAP1, DOO1], for example.) In this chapter, the equations of filtering theory are developed for discrete-time filtering of continuous-time, discrete-parameter, dynamic systems.

2-1. The System Model

A large class of dynamic systems can be modeled by the vector-valued stochastic differential equation,

$$dx = f(x,t)dt + G(x,t)d\beta \quad (2-1-1)$$

or its formal white-noise equivalent,

$$\dot{x} = f(x,t) + G(x,t)w \quad (2-1-2)$$

where x is an n -random vector, β is an m -vector Brownian motion process, f is a vector function, and G is a matrix function. The properties of the Brownian motion process, its relation to the white-noise process $w(t)$, and its physical significance are given in [DOO1].

The existence and uniqueness of the solution to (2-1-1) over an interval T , namely,

$$x = \int_T f(x,t)dt + \int_T G(x,t)d\beta \quad (2-1-3)$$

(where the integrals are interpreted in Appendix A) can be proven under the conditions that f and G are Lipschitz in x . (See [IT01 or D001].)

The Linear System Dynamic Model

In the case that f and G are not state dependent, the system model is linear,

$$dx = A(t)x dt + B(t)d\beta, \quad (2-1-4)$$

or

$$\dot{x} = A(t)x + B(t)w \quad (2-1-5)$$

where A is an $n \times n$ matrix, B is an $n \times m$ matrix, and β is a Brownian motion process. The linear model is of great importance for several reasons, one of which is that the solution can be written uniquely as

$$x(t) = \Phi(t, t_0)x(t_0) + \int_{t_0}^t \Phi(t, \tau)B(\tau)d\beta(\tau) \quad (2-1-6)$$

where $\Phi(t, t_0)$ is the transition matrix transforming the unforced response over the interval $[t_0, t]$ and the forced response is a stochastic vector integral. If $x(t_0)$ is taken as a Gaussian vector independent of β , then $x(t)$ is a Gaussian vector, and the covariance matrix of the state $P(t, s)$ can be written as

$$\begin{aligned} P(t, s) &= E \{x(t)x'(s)\} \\ &= \Phi(t, t_0)P(t_0)\Phi'(s, t_0) + \\ &\quad \int_{t_0}^{\min(t, s)} \Phi(t, \tau)B(\tau)Q(\tau)B'(\tau)\Phi'(s, \tau)d\tau \end{aligned} \quad (2-1-7)$$

where

$$E \{\beta(s)\beta'(t)\} = \int_{t_0}^{\min(t, s)} Q(\tau)d\tau. \quad (2-1-8)$$

The Observation Model

The general form of the continuous observation model is the stochastic differential equation

$$dy = h(x,t)dt + dv \quad (2-1-9)$$

where y is an r -dimensional observed quantity, h is a vector function and v is a Brownian motion vector:

$$E \{v(t)v'(s)\} = \int_{t_0}^{\min(t,s)} R(\tau)d\tau.$$

In the case of discrete-time measurements, the analogous model of (2-1-9) is

$$y(t_k) = h(x(t_k), t_k) + v(t_k). \quad (2-1-10)$$

The additive noise v is a white-noise sequence with covariance

$$E\{v(t_k)v'(t_\ell)\} = R(t_k) \delta_{k\ell} \quad (2-1-11)$$

where $\delta_{k\ell}$ is the Kronecker delta function: $\delta_{k\ell} = 1$ for $k = \ell$ and $\delta_{k\ell} = 0$ for $k \neq \ell$. In the case that h is not state dependent, (2-1-10) may be written as

$$y(t_k) = M(t_k)x(t_k) + v(t_k) \quad (2-1-12)$$

where M is an $r \times n$ matrix.

For the linear observation model and system model, a necessary condition for estimating the state of the system is that the system be observable [KAL2]. No equivalent theory exists for the general nonlinear problem.

2-2. Extrapolation of the Conditional Density Function and Its Moments

The Markov, Brownian motion process $x(t)$ satisfying the Ito stochastic differential equation (2-1-1) has a probability density

function $p(x,t)$ that is conditioned on the previous observations, $Y(t_k) = \{y(t_1), \dots, y(t_k)\}$. Between measurements this conditional probability density function, called the transitional probability density function, $p(x,t|Y(t_k))$, is given by the Kolmogorov forward equation,

$$\frac{dp(x,t)}{dt} = - \sum_{i=1}^n \frac{\partial [p(x,t)f_i]}{\partial x_i} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 [p(x,t)[GQG']_{ij}]}{\partial x_i \partial x_j} \quad (2-2-1)$$

where f_i is the i th component of f and $[GQG']_{ij}$ is the ij th element of GQG' . For a derivation and discussion of this expression and diffusion processes in general, see [FEL1, ITO1, DYN1, ITO2, MCK1]. The initial condition for (2-2-1) is the value at the prior measurement, $p(x,t_k | Y(t_k))$.

By the use of Ito's Theorem A-3, the exact equations for the extrapolation of the moments of $p(x,t)$ between measurements can be determined. First, define the expected value of ϕ in (A-11) as

$$\hat{\phi}(x(t)|t_k) = E \{ \phi(x) | Y(t_k) \} = \int \phi(x) p(x,t | Y(t_k)) dx \quad (2-2-2)$$

Then the probability density, $p(x,t|Y(t_k))$ satisfies (2-2-1), and the extrapolation of $\hat{\phi}$ is

$$\begin{aligned} d\hat{\phi}(x(t)|t_k) = E\{\phi_x' \cdot f | Y(t_k)\} + \frac{1}{2} \text{tr} E \{ GQG' \phi_{xx} | Y(t_k) \} dt, \\ t_k \leq t < t_{k+1} \end{aligned} \quad (2-2-3)$$

from (A-11) where each subscript denotes the partial derivative with respect to x . The conditional moments are determined as follows:

for $\phi(x) = x$ in (2-2-3), the conditional mean $\hat{x}(t|t_k)$, defined as $E\{x(t)|t_k\}$ is found from the solution of the differential equation,

$$\frac{d\hat{x}(t|t_k)}{dt} = E\{f(x,t)|Y(t_k)\} = \hat{f}(t|t_k),$$

$$t_k \leq t < t_{k+1}. \quad (2-2-4)$$

For $\phi(x) = x x'$, the covariance matrix $P(t|t_k)$, defined as

$$P(t|t_k) = E\{[x - \hat{x}(t|t_k)][x - \hat{x}(t|t_k)]'\} \quad (2-2-5)$$

is found from the solution of the matrix differential equation,

$$\begin{aligned} \frac{dP(t|t_k)}{dt} = & E\{xf' - \hat{x}(t|t_k)\hat{f}(t|t_k)\} \\ & + E\{fx' - \hat{f}(t|t_k)\hat{x}'(t|t_k)\} \\ & + E\{QG'\}, \quad t_k \leq t < t_{k+1} \end{aligned} \quad (2-2-6)$$

Higher-order moments could be developed but will not be derived here. The solution of (2-2-4) and (2-2-6) can only be approximated for the general nonlinear estimation problem since an infinite number of moments are needed to solve for $p(x,t)$ [KUS1, KUS2].

2-3. Update of the Conditional Density Function and Its Moments

In the discrete-time filter, the other part of the filtering problem is to compute the change in the conditional probability density function when additional information is available in the form of a new measurement. This aspect of the problem is discussed in various reports, [COX1, JAZ1].

By the use of Bayes' rule, the probability density function immediately after the measurement can be written as

$$p[x(t_k)|Y(t_k)] = \frac{p[y(t_k)|x(t_k)]p[x(t_k)|Y(t_{k-1})]}{p[y(t_k)|Y(t_{k-1})]} \quad (2-3-1)$$

where

$$p[y(t_k)|Y(t_{k-1})] = \int p[y(t_k)|x(t_k)]p[x(t_k)|Y(t_{k-1})]dx(t_k) \quad (2-3-2)$$

and

$$\begin{aligned} p[y(t_k)|x(t_k)] &= p[y(t_k) - h(x(t_k), t_k)] \\ &= p[v(t_k)] \end{aligned} \quad (2-3-3)$$

which is a Gaussian distribution of zero mean and covariance given by (2-1-11).

The expected value, $\hat{x}(t_k|t_k)$, is the conditional mean of the probability density of (2-3-1). Multiplying by x and integrating over x yields

$$\hat{x}(t_k|t_k) = \frac{E\{x(t_k)p[y(t_k)|x(t_k)]|t_k^-\}}{E\{p[y(t_k)|x(t_k)]|t_k^-\}} \quad (2-3-4)$$

where the superscript minus indicates a value just prior to the measurement. The conditional covariance matrix $P(t_k|t_k)$ is obtained from the second moment of (2-3-1) and by expanding (2-2-5),

$$P(t_k|t_k) = \frac{E\{x(t_k)x'(t_k)p[y(t_k)|x(t_k)]|t_k^-\}}{E\{p[y(t_k)|x(t_k)]|t_k^-\}} \quad (2-3-5)$$

The higher-order moments are not needed for this work and are thus omitted. As was the case for the conditional moment extrapolations, the updated moments depend on an approximation to solve for the general first and second moments (2-3-4) and (2-3-5) since these require the solution of the conditional probability density. The approximate solutions are given in Section 4-1.

2-4. The System Disturbance Model

In the case that the state model is of the form (2-1-5), then the state covariance matrix can be written as

$$P(t) = \Phi(t, t_0)P(t_0)\Phi'(t, t_0) + \int_{t_0}^t \Phi(t, \tau)B(\tau)Q(\tau)B'(\tau)\Phi'(t, \tau)d\tau$$

(2-4-1)

In this case, a discrete-time equation may be used to model (2-1-5):

$$x(t_{k+1}) = \Phi(t_{k+1}, t_k)x(t_k) + u(t_k) \quad (2-4-2)$$

This discrete representation can be justified by defining $u(t_k)$ as a white noise sequence with covariance matrix,

$$E\{u(t_k) u'(t_j)\} = V(t_k) \delta_{kj} \quad (2-4-3)$$

where

$$V(t_k) = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau)B(\tau)Q(\tau)B'(\tau)\Phi'(t_{k+1}, \tau)d\tau \quad (2-4-4)$$

A numerical method of solving (2-4-4) is given in Section 3-2. For purposes of digital computation, the simulation of (2-4-4) requires a multivariate Gaussian distribution generator. The equations for generating this distribution are given in Appendix B.

Although the above approach is straightforward, it avoids the difficult paradoxes encountered by Bryson and Ho [BRY1] and Bierman [BIE1] in transforming from the continuous-time to the discrete-time model.

Another method of approximating (2-1-5) is to write the discrete-time equivalent as

$$x(t + \Delta t) \doteq (I + A(t)\Delta t)x(t) + \sqrt{\Delta t} B(t)u(t_k) \quad (2-4-5)$$

where Δt is small and $u(t_k)$ is a white-noise sequence with covariance matrix $Q(t_k)$, the same as the cov $\{w(t)\}$. The contribution of the disturbance in (2-4-4) is

$$\begin{aligned}
 V(t) &= \int_{t_k}^{t_{k+1}} \delta(t_{k+1}, \tau) B(\tau) Q(\tau) B'(\tau) \delta'(t_{k+1}, \tau) d\tau \\
 &\doteq B(t_k) Q(t_k) B'(t_k) \Delta t
 \end{aligned} \tag{2-4-6}$$

for $t_{k+1} - t_k = \Delta t > 0$.

The contribution of the disturbance in (2-4-5) is

$$\begin{aligned}
 V(t_k) &= E\{\sqrt{\Delta t} B(t) u(t_k) u'(t_k) B'(t) \sqrt{\Delta t}\} \\
 &= B(t_k) Q(t_k) B'(t_k) \Delta t
 \end{aligned} \tag{2-4-7}$$

and therefore, for small enough value of Δt , (2-4-5) is equivalent at each Δt interval to its continuous-time counterpart (2-1-5). This approach may be considered where a sample function is desired with a larger number of samples than in (2-4-2) and where the computation of (2-4-4) is not necessary.

For a more general utilization of the property of discrete-time modeling of continuous-time processes, consider (2-1-1),

$$dx = f(x, t)dt + G(x, t)d\beta.$$

The discrete-time model is written as

$$x(t + \Delta t) = x(t) + f(x(t), t)\Delta t + \sqrt{\Delta t} G(t)u(t_k) \tag{2-4-8}$$

where $u(t_k)$ is a white-noise sequence with covariance matrix $Q(t_k)$. The contribution of the disturbance in (2-4-8) is the same as in (2-4-7).

If the covariance matrix of the β process is given by (2-1-8), then by Theorem A-2,

$$E\left\{\int_{\Delta t} G(x, t)d\beta\right\} = 0 \tag{2-4-9}$$

and

$$\begin{aligned}
 E\left\{\int_{\Delta t} G(x,t)d\beta\left[\int_{\Delta t} G(x,t)d\beta\right]'\right\} &= \\
 &= \int_{\Delta t} G(t,x)Q(t)G'(t,x)dt \quad (2-4-10) \\
 &\doteq G(t,x)Q(t)G'(t,x) \Delta t
 \end{aligned}$$

for small Δt (i.e., the interval Δt must be small with respect to the time rate of change of G and Q). Once again, the simulation generator of this Gaussian multivariable disturbance is given in Appendix B.

III. Discrete Filtering of a Linear Continuous System

Since the conception of linear filtering theory in 1960, a great amount of effort has been put forth, both in setting the theory straight and in applying the theory to practice. Within a relatively short period the theory was applied in space guidance and navigation programs such as Ranger, Mariner, and Apollo. In addition, applications have been made of linear filtering theory to aircraft guidance, industrial process control, and other detection problems. In most cases, it is necessary to linearize the system differential equations about the nominal state in order to apply a linear filter.

The explicit problems of deriving the optimal linear filter, showing its stability, and proving asymptotic convergence are solved [KAL1, KAL2, KAL3, BUC1]. The problems remaining in linear filtering theory are related to process modeling and filter mechanization.

The first part of this chapter is concerned with developing the equations for discrete-time filtering of the continuous-time system. Next, some methods that have not appeared in the literature are presented for mechanizing the linear filter. Then an example of an application of linear filtering theory to a steam-turbine generator is presented. Finally, a method for compensation for modeling errors is presented.

3-1. Linear Filtering

The linear set of differential equations that describe the linear or linearized system may be written as

$$\dot{x} = A(t)x + B(t)w \quad (3-1-1)$$

where $A(t)$ is an integrable $n \times n$ matrix, $B(t)$ is an integrable $n \times m$ matrix and w is an m -vector white noise process with covariance $Q(t)\delta(t-\tau)$. Although the white noise process is only a formal representation of (2-1-4), its use is justified since in the solution of (3-1-1), $w(t)$ appears in the integral only and its integral, the Brownian motion process, is a well-defined function. The linear observation process consists of sampled data of the form

$$y(t_k) = M(t_k)x(t_k) + v(t_k) \quad (3-1-2)$$

where $M(t_k)$ is an $r \times n$ linear non-invertible transformation and $v(t_k)$ is an r -vector of a Gaussian white noise sequence with covariance $R(t_k)$. It is assumed that the system is observable.

The linear filtering problem is this: given initial estimates of the state $x(t_0)$ with Gaussian distribution and of $P(t_0)$, the covariance matrix, determine the minimum-variance recursive estimate of the system, (3-1-1) and (3-1-2). Slightly different premises are assumed for the discussion in Section 3-4.

With the modifications given in Section 2-4, the linear discrete-time filter for the linear continuous-time system fits the framework as derived by Kalman [KAL1]. For the linear problem, the process x is Gaussian and only the first and second moments of (2-2-1) and (2-3-1) are needed, namely, the linear version of (2-2-4) and (2-2-6) for estimation extrapolation and (2-3-4) and (2-3-5) for estimation update.

The linear filtering equations are given here without proof since the derivation is commonly found in the literature and only a few steps are required to reduce the above equations. Between measurements, the state estimate and covariance matrix are

$$\hat{x}(t_k|t_{k-1}) = \Phi(t_k, t_{k-1})\hat{x}(t_{k-1}|t_{k-1}) \quad (3-1-3)$$

$$P(t_k|t_{k-1}) = \Phi(t_k, t_{k-1})P(t_{k-1}|t_{k-1})\Phi'(t_k, t_{k-1}) + \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)B(\tau)Q(\tau)B'(\tau)\Phi'(t_k, \tau)d\tau \quad (3-1-4)$$

where Φ is the system transition matrix. At the time of the measurement, the state estimate and covariance matrix are

$$\hat{x}(t_k|t_k) = \hat{x}(t_k|t_{k-1}) + G(t_k)[y(t_k) - M(t_k)\hat{x}(t_k|t_{k-1})] \quad (3-1-5)$$

$$P(t_k|t_k) = P(t_k|t_{k-1}) - G(t_k)M(t_k)P(t_k|t_{k-1}) \quad (3-1-6)$$

where the filter gain is

$$G(t_k) = P(t_k|t_{k-1})M'(t_k)[M(t_k)P(t_k|t_{k-1})M'(t_k) + R(t_k)]^{-1}. \quad (3-1-7)$$

These five equations are the optimal linear filtering equations.

These equations are closely related to the maximum likelihood estimate and the stochastic approximation method as shown by Ho [HO1].

When the system parameters are known exactly the solution of the linear filtering equations is straightforward. However, the real-time computer requirements may be exorbitant and unfeasible. The methods discussed in the next section due to Frame and Needler [FRA2] are useful in decreasing computational time and increasing accuracy. In the case that the system parameters are not known, the methods discussed in Section 3-4 can be used to correct for modeling errors.

3-2. State and Covariance Extrapolation

Minimizing computational time and maximizing computational accuracy are crucial problems in mechanizing an on-line filter for filtering a complex process since these two criteria constrain the filter capability in terms of (1) setting a lower limit on sampling rate, (2) allowing local iterations for linearization of the extended linear filter, and (3) allowing the system model to be sufficiently complex to adequately model the physical system.

Two important aspects in both real-time system estimation and control and in digital simulation of the same are extrapolation of the state of the system and extrapolation of the covariance matrix of the system. In the case that the system may be described by (3-1-1) where A and B are either time invariant or can be approximated as time invariant over the time period of interest, then the computational methods described in the sequel are useful in state and covariance extrapolation between measurements.

The solution of (3-1-1) can be written as

$$x(t) = \Phi(t-t_0)x_0 + \int_0^t \Phi(t-\tau)Bu(\tau)d\tau \quad (3-2-1)$$

where the transition matrix Φ is

$$\Phi(t) = e^{At} = \sum_{k=0}^{\infty} A^k t^k / k! \quad (3-2-2)$$

and t_0 is taken to be zero.

Using the transition matrix to evaluate the transient response of a linear time-invariant unforced system was described by Liou [LI01] and others [BRA1, MAS1]. Ganapathy and Rao [GAN1]* pointed out that

*For corrections to [GAN1], see [RA01].

the A matrix has to be raised to successively higher powers to allow t to be taken over a sufficiently large interval. They in turn show that e^{At} can be expanded in an infinite series of scalars instead of matrices and that A^{n-1} is the largest power of A that is required.

An alternative method requires slightly less computation and is possibly more straightforward than [GAN1]. This new method utilizes the conjoint algorithm to compute the coefficients $d_1, \dots, d_n$ of the characteristic polynomial of A,

$$\lambda^n + d_1 \lambda^{n-1} + \dots + d_{n-1} \lambda + d_n = 0 \quad (3-2-3)$$

and the constant coefficient matrices, $B_0, \dots, B_n$ of the conjoint matrix, defined by

$$B(\lambda) = \text{adjoint } (\lambda I - A) = B_0 \lambda^{n-1} + B_1 \lambda^{n-2} + \dots + B_{n-2} \lambda + B_{n-1} \quad (3-2-4)$$

where $B_0 = I$ and $B_n = B_{-1} = 0$. The conjoint algorithm described by Frame [FRA1] yields the relationships,

$$B_k = A B_{k-1} + d_k I \quad (3-2-5)$$

$$d_k = -(1/k) \text{tr } A B_{k-1} \quad (3-2-6)$$

Define the transition matrix as

$$\Phi(t) = \sum_{j=0}^{n-1} B_j z_j(t) \quad (3-2-7)$$

where

$$z_j(t) = \sum_{k=0}^{\infty} c_k \frac{t^{j+k}}{(j+k)!} \quad (3-2-8)$$

and the c_k are determined so that

$$\sum_{j=0}^m d_j c_{m-j} = 0 \text{ for } m > 0; c_0 = d_0 = 1. \quad (3-2-9)$$

Then since $B_n = B_{-1} = 0$, and $\dot{z}_j = z_{j-1}$ for $j > 0$ we have

$$\dot{\Phi}(t) = \sum_{j=0}^n B_j \dot{z}_j(t) \quad (3-2-10)$$

$$= \sum_{j=1}^n A B_{j-1} \dot{z}_j(t) + \sum_{j=0}^n d_j I \dot{z}_j(t)$$

$$= A \sum_{j=1}^n B_{j-1} z_{j-1}(t) + I \frac{d}{dt} \sum_{j=0}^n \sum_{m=j}^{\infty} d_j c_{m-j} t^m / m! \quad (3-2-11)$$

or

$$\dot{\Phi}(t) = A \Phi(t) + I \cdot 0 = A \Phi(t). \quad (3-2-12)$$

Also it should be noted that

$$\Phi(0) = \sum_{j=1}^{n-1} B_j z_j(0) = B_0 z_0(0) = B_0 c_0 = I \quad (3-2-13)$$

Hence, we have the desired result:

$$\Phi(t) = e^{At} \quad (3-2-14)$$

From equation (3-2-9) we solve for each c_k using at most n multiplications:

$$c_m = - \sum_{j=1}^m d_j c_{m-j}, \quad m \leq n \quad (3-2-15)$$

$$c_m = - \sum_{j=1}^n d_j c_{m-j}, \quad m \geq n$$

By substituting these values of c_k into (3-2-8) and then solving (3-2-7), $\Phi(t)$ can be determined easily

The computational savings are considerable when compared with using (3-2-2) truncated at some value $k = N$ for N large compared to n . On the other hand, when compared to the method utilizing the Hamilton-Cayley theory [GAN1], the savings in computation are less. In particular, the exponential series requires $(N-1) n^3 + N$ multiplications, the Hamilton-Cayley method requires $n^4 - n^3 - \frac{1}{2}n^2 + \frac{1}{2}n + 2nN + N$ multiplications, and the above method requires $n^4 - n^3 - 2n^2 + n + 2nN + N$ multiplications. The number of divisions is negligible and additions are proportional to multiplications. Computer results confirmed the method and the previously mentioned computational savings; for the example discussed later in this chapter, the computation was about 30% as great as the computation of (3-2-2).

Covariance Extrapolation

In the case that u in (3-2-1) is a vector-valued, zero-mean, white-noise process, then to determine the covariance matrix of the error in $x(t)$, it is necessary to evaluate*

$$V(t) = \int_0^t \phi(t-\tau)Q\phi'(t-\tau)d\tau \quad (3-2-16)$$

where

$$Q = E\{B u(t)u'(t)B'\}.$$

*This same quadratic form also appears in optimal control problems with a quadratic performance index over a finite time interval. See [LEE1] and [LEV1].

As pointed out by Levis [LEV1], errors are present from two sources: truncation of the series expansion of $\Phi(t)$ in (3-2-2) and approximate numerical integration methods. In the following, the integral is solved exactly and then a recursive relationship is derived to solve for V to the desired accuracy.

Theorem The solution of (3-2-16) where Q is a symmetric matrix is given by

$$V(t) = \sum_{k=0}^{\infty} G_k \frac{t^{k+1}}{(k+1)!} \quad (3-2-17)$$

where

$$G_{k+1} = AG_k + G_k' A' \quad (3-2-18)$$

and the value of G_0 is

$$G_0 = Q. \quad (3-2-19)$$

Proof: By substitution of variable, (3-2-16) can be rewritten as

$$V(t) = \int_0^t \Phi(\tau) Q \Phi'(\tau) d\tau \quad (3-2-20)$$

After substituting (3-2-2) into (3-2-20) and integrating, the integral becomes

$$V(t) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} A^j Q A^{k,j} \frac{t^{j+k+1}}{j! k! (j+k+1)!} \quad (3-2-21)$$

$$= \sum_{k=0}^{\infty} \sum_{j=0}^k \binom{k}{j} A^j Q A^{k-j,j} \frac{t^{k+1}}{(k+1)!} \quad (3-2-22)$$

If G_k is defined as

$$G_k = \sum_{j=0}^k \binom{k}{j} A^j Q A^{k-j} \quad (3-2-23)$$

then

$$G_k A' = \sum_{j=0}^k \binom{k}{j} A^j Q A^{k+1-j} \quad (3-2-24)$$

Replacing j by $j-1$ in (3-2-23) it follows that

$$G_k = \sum_{j=1}^{k+1} \binom{k}{j-1} A^{j-1} Q A^{k+1-j} \quad (3-2-25)$$

and multiplying by A gives

$$A G_k = \sum_{j=1}^{k+1} \binom{k}{j-1} A^j Q A^{k+1-j} \quad (3-2-26)$$

Thus, the sum of (3-2-24) and (3-2-26) is

$$G_{k+1} = \sum_{j=0}^{k+1} \binom{k+1}{j} A^j Q A^{k+1-j} \quad (3-2-27)$$

and by the symmetry property of G_k ,

$$G_{k+1} = A G_k + (A G_k)'. \quad (3-2-28)$$

Thus, $V(t)$ can be given exactly by

$$V(t) = \sum_{k=0}^{\infty} G_k \frac{t^{k+1}}{(k+1)!} \quad (3-2-29)$$

where G_k is given by (3-2-25) and $G_0 = Q$.

Although the benefits from utilizing this method cannot be evaluated analytically, the increase in accuracy and the savings in computation as compared to numerical integration are considerable in the case of large

t where it is necessary to have a large number of terms in the expansion, as was realized in the following problem. *

3-3. An Example of State Estimation of a Steam Turbo-Generator System

A steam turbo-generator system is composed of a control system, a steam boiler plant, and a turbo-generator. Disturbances occur to the system in terms of a changing load caused by both changing consumer power demands and changing power supplied by alternate sources. A model can be constructed of the steam turbo-generator system by assuming that the coupling to other sources is relatively loose. See for example, Kirchmayer [KIR1], Park [PAR1, PAR2], and Stanton [STA1]. Then the system can be linearized around its operating point to obtain a simulation model. The problem at hand is to estimate the state of the system using this model. The model can be written in the form of (3-1-1) and (3-1-2):

$$\dot{x} = Ax + Bu$$

$$y = Mx + v$$

where in (3-1-1)

$$A = \begin{bmatrix} -1/T_1 & 0 & 0 & 0 & K_1/T_1 \\ 1/T_3 & -1/T_3 & 0 & 0 & 0 \\ 0 & 1/T_5 & -1/T_5 & 0 & 0 \\ 0 & K_2/T_4 & (1-K_2)/T_4 & -1/T_4 & 0 \\ 0 & 0 & 0 & -D/M & 1/M \end{bmatrix},$$

$$B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

*In the following example, less exact numerical integration required approximately twice as much computation time as the above method.

and u is a random disturbance. The reference input is neglected for purposes of this study. An explanation of the physical constants appearing in the A matrix and M matrix as well as their numerical values are given in Appendix C.

From a study of the random disturbances caused by changing loads, it was determined that a random-walk process would be an acceptable model of this phenomenon. Since it can be shown [PAP1] that the random-walk process is a limiting form of the Brownian motion process, the Brownian motion process observed at discrete-time intervals will exhibit the stochastic characteristics of the random-walk process. Thus, in turn, the noise source can be pre-whitened by integrating a white noise source to yield the Brownian motion process. The addition of one state variable is necessary to account for the disturbance. The state can be augmented with the differential equation,

$$\dot{x}_6 = w$$

where w is a white noise source. The augmented A matrix and B matrix now become

$$A = \begin{bmatrix} -1/T_1 & 0 & 0 & 0 & K_1/T_1 & 0 \\ 1/T_3 & -1/T_3 & 0 & 0 & 0 & 0 \\ 0 & 1/T_5 & -1/T_5 & 0 & 0 & 0 \\ 0 & K_2/T_4 & (1-K_2)/T_4 & -1/T_4 & 0 & 0 \\ 0 & 0 & 0 & -D/M & 1/M & 1/M \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

In Fig. 3-1, a block diagram is shown of the overall system where each component is represented by its Laplace Transform system function.

An acceptable model of the observation process depends on the assumptions concerning the measurement devices. One choice is to observe frequency deviation, x_5 , and frequency rate deviation, $\dot{x}_5$. Thus, the observation matrix is

$$M = \begin{bmatrix} 0 & 0 & 0 & 1/M & D/M & 1/M \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} .$$

The noise source $v(t_k)$ can be modeled as a two-vector white-noise sequence since the measurements are refined to the point that bias and other gross errors are eliminated.

The digital simulation was performed using parameter data from [PAR1]. In addition, representative values were taken for the measurement noise and system disturbance statistics. These values are given in Appendix C.

The digital simulation of the steam generating plant including the random effects, and the discrete-time filter demonstrated the effectiveness of filter applications in problems of this type. By using the filter as mechanized in Sections (3-1) and (3-2) it was possible to observe the estimation error time-history for various noise sample functions. An improvement on simulating sample runs is to study the estimation error dispersion by using the filter covariance matrix of estimation error, starting with a large initial variance. In this way, the need for performing Monte Carlo studies is obviated.

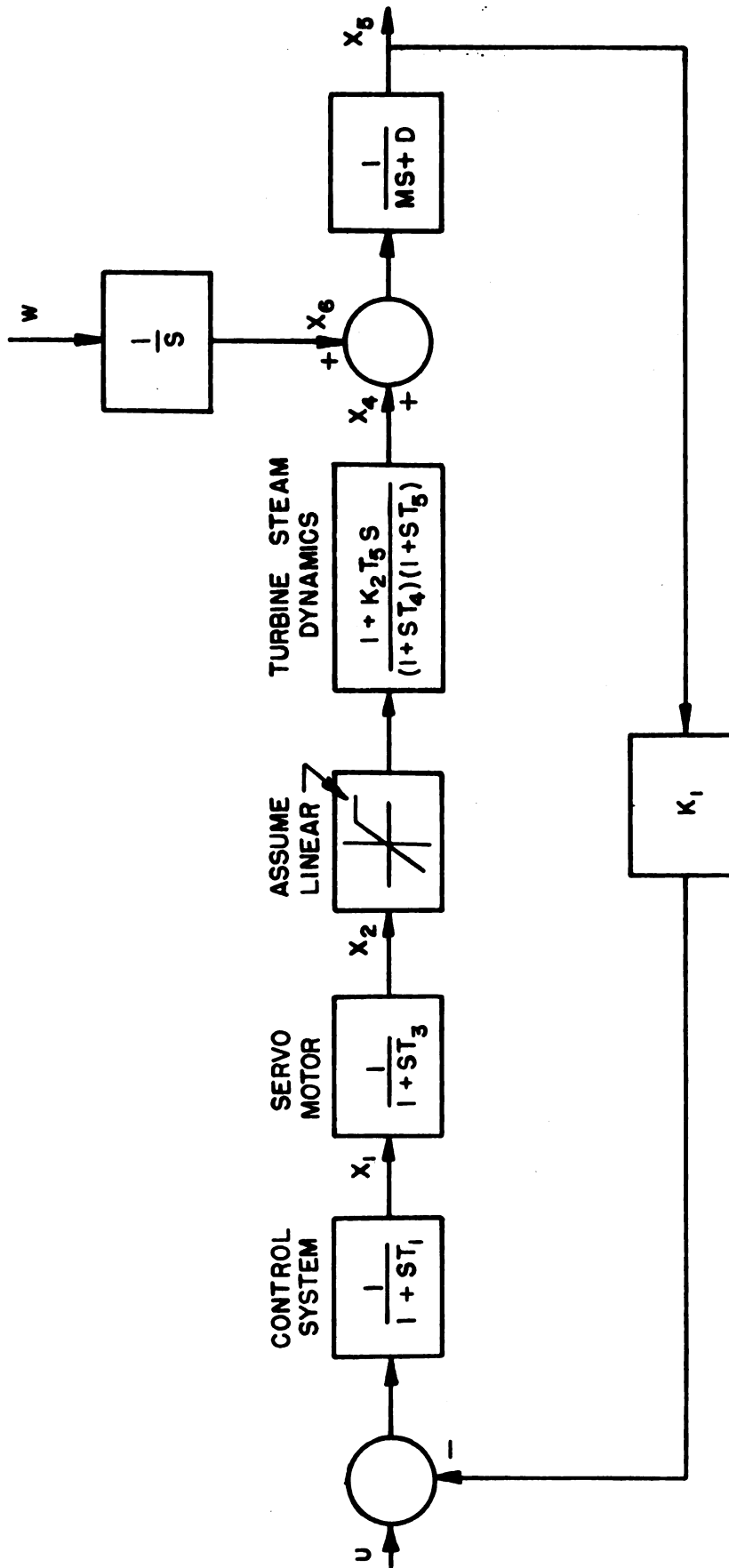


Fig. 3-1. The Block Diagram of a Steam Turbo-Generator System. The variables are the controlled input reference u , the random disturbance w caused by random load changes, and the state variables of the system where, in particular, x_5 is the frequency deviation from the nominal value. The steam value and steam boiler may be approximated as linear if the deviation from the nominal value is small. The final block accounts for the moment of inertia M and friction D of the generator and turbine.

The tradeoff between improved filter performance and the following factors can be evaluated:

1. The required sampling frequency which limits modeling complexity, determines the maximum computer cycle time, and determines measurement transducer required sampling rates.
2. The quality and quantity of measurements that are required.
3. The effect of system disturbances.

Many simulation runs were performed while varying model parameters, such as sampling rate, the measurement accuracy and the magnitude of the disturbance variance. A representative error-study time-history is shown in Fig. 3-2.

In summary, the application of a Kalman filter to a steam-turbo generator plant was simulated and the results were judged to be successful. The methods proposed in Section (3-2) were used. The study should be extended by evaluating the effect of (1) more refined model development, (2) more exact noise and disturbance statistics, and (3) application of a closed-loop optimal controller for minimizing output frequency error. An alternative approach to improved process modeling is presented in the next section.

3-4. The Problem of Filter Divergence

Early in the study of filtering applications, it was found that although in theory the filter estimates converged asymptotically to the state variables, in practice the filter estimates would diverge from the state variables, especially where modeling errors were large relative

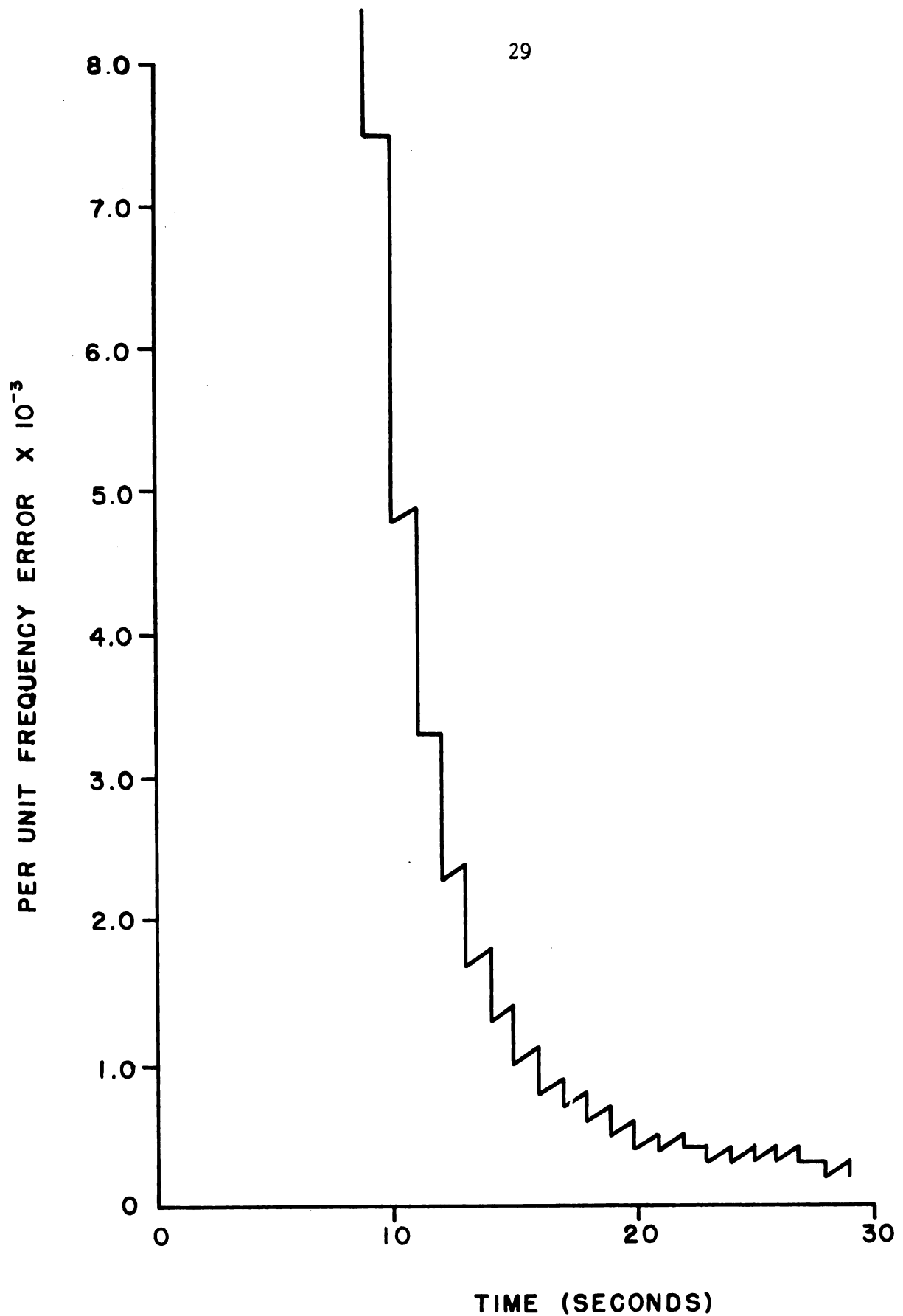


Fig. 3-2. The RMS Estimation Error of Frequency Deviation for the Steam Turbo-Generator. Starting with a very large error, the RMS value converges relatively quickly to a steady-state error.

to system disturbance statistics. This outcome may be a consequence of modeling error in any of the parameter values or covariance specifications [SCH1]. Since the filter has no measure of the consistency of the estimation error covariance and the actual estimation error, it is possible for the covariance matrix to indicate a much more accurate estimate than is actually being performed. The filter divergence is a result of later observations being very lightly weighted since the covariance matrix of the estimation error has converged to a smaller quantity than warranted.

Several approaches have been tried on the problem of filter divergence. One approach is to augment the state vector with the model parameters that are only approximately known [SM11]. By estimating these uncertain parameters, this cause of filter divergence is reduced. The accompanying disadvantage is that a larger state vector leads to greater computational requirements and the possibility of establishing an unobservable system. Another approach is to artificially include a sufficiently large system disturbance covariance such that the filter will not converge improperly [FIT1]. Naturally, a good deal of modeling and simulation needs to be performed to insert proper disturbance statistics. One more approach suggested in [BER1] is to prevent the determinant of the covariance matrix from decreasing below a given reference value. A parameter is chosen such that after the determinant has converged for this length of time, the covariance is scaled so that the determinant remains constant and the covariance matrix does not go to zero.

With these methods, adjustments must be made strictly on the basis of a priori knowledge obtained from measurements and simulation studies.

For this reason, methods are desired that are adaptive to the actual estimation environment during the filtering process.

As pointed out in [JAZ2], the filter divergence problem can be judged by comparing the statistics of the actual measurement errors with the covariance matrix of measurement errors. The only measure of divergence available in the filtering problem is to compare the actual measurement, $y(t_k)$, with the measurement estimate $\hat{y}(t_k)$ based on the previous state estimate; namely, the measurement estimate error $e(t_k)$ is

$$e(t_k) = y(t_k) - E\{y(t_k) | Y(t_{k-1})\} \quad (3-4-1)$$

or

$$e(t_k) = y(t_k) - M(t_k)x(t_k | t_{k-1}). \quad (3-4-2)$$

The measurement estimate error covariance matrix $Z(t_k | t_{k-1})$ is defined as

$$Z(t_k | t_{k-1}) = E\{e(t_k)e'(t_k)\} \quad (3-4-3)$$

$$= M(t_k)P(t_k | t_{k-1})M'(t_k) + R(t_k) \quad (3-4-4)$$

Since the mean measurement error estimate $E\{e(t_k)\}$ is zero, it is possible to judge the consistency of the filter covariance with the actual errors in the measurement estimate by comparing (3-4-4) with the statistics of $e(t_k)e'(t_k)$.

It is further shown in [JAZ2] that in the case of no system disturbances and scalar measurements, a disturbance covariance can be artificially created to adapt the identical noise inputs to the statistics of the error in the measurement estimate. By maximizing the probability density function $p[e(t_k)]$ with respect to a scalar q

where qI represents the system disturbance, the most probable disturbance is determined at each estimate.

Another approach suggested by Anderson, et al., [AND2] is to obtain estimates of the state x , the transition matrix Φ , the input covariance matrix V , and the noise covariance matrix R . The assumptions are that the system is discrete, the parameters Φ , V , and R are constant, the random sources are zero-mean, independent, and identically distributed, but probably most importantly, the observation matrix M is invertible, that is, each state variable is directly observable through noise. Although this latter restriction may be severe in many cases, it should be noted that in a class of system applications this technique may be valuable in taking adequate experimental measurements to ascertain the true values of Φ , V , and R and then using these values for real-time estimation problems--assuming the other requirements are satisfied.

A different approach is used in this thesis to handle the problem of adapting the estimation procedure to the measurement data. This proposed approach has applications in filtering linear systems with parameter uncertainties; however, it has been used more extensively in the non-linear filtering problem. Thus, the discussion of non-linear adaptive filtering is deferred until the next chapter.

IV. Approximate Filtering of Nonlinear Systems

The subject of this chapter is discrete-time filtering of nonlinear continuous-time systems. The case of system disturbances and noisy measurements is covered. A second-order nonlinear filter is developed, that is similar to the one suggested by Athans, et al., [ATH2] but is extended to the case of system disturbances. Then a new technique is developed by the author to adapt the filter to the measurement statistics such that those statistics and the filter covariance estimates are consistent.

4-1. Nonlinear Filtering

The original linear filtering theory was adapted to nonlinear systems by linearizing the system about the nominal value and using the "first-order" or "quasilinear filter" as suggested in [FRI1, MOW1, SCH1] during the early 60's.

By the middle 60's, the problem of nonlinear filtering was being formulated [KUS1, KUS2, BUC2]. The exact equations for the propagation of the state conditional probability density function conditioned on the observation sequence are given by (2-2-1) and (2-3-1). However, an infinite number of moments is required to realize the exact solution of these equations. Thus, considerable effort has been extended to obtain an approximate filtering solution to (2-2-1) and (2-3-1). These efforts

include [KJS3, SOR1, ATH2, JAZ1].

In this chapter, the state model is assumed to be of the form,

$$\dot{x} = f(x,t) + G(x,t)w \quad (4-1-1)$$

where the system dynamics function f is an n -vector which is twice-differentiable with respect to x and once differentiable with respect to t , the input matrix $G(x,t)$ is a continuous matrix function of x,t and as before, w is the zero-mean white-noise disturbance with covariance $Q(t)\delta(t-\tau)$. The observation equation is of the discrete form,

$$y(t_k) = h(x(t_k)) + v(t_k) \quad (4-1-2)$$

where the r vector h is also twice differentiable with respect to x and once differentiable with respect to t , and as before, $v(t_k)$ is a white-noise sequence with covariance $R(t_k)\delta_{k\ell}$. And, finally, $x(t_0)$ is assumed to be a random vector with known Gaussian distribution.

Some definitions are needed to further expound the nonlinear filter. The Jacobian matrix of $f(x)$ is defined as $A(x)$ where the ij th entry of A is

$$A_{ij} = \frac{\partial f_i}{\partial x_j}, \quad i,j = 1, \dots, n. \quad (4-1-3)$$

Define the second partials of f as the Hessian matrix, $F_\ell(x)$ for $\ell = 1, \dots, n$, where the ij th entry of F_ℓ is

$$F_\ell(x)_{ij} = \frac{\partial^2 f_\ell}{\partial x_i \partial x_j}, \quad i,j = 1, \dots, n. \quad (4-1-4)$$

For the output function $h(x,t)$, define the output Jacobian matrix as $M(x)$ where the ij th entry is

$$M_{ij} = \frac{\partial f_i}{\partial x_j}, \quad i = 1, \dots, m, \quad j = 1, \dots, n \quad (4-1-5)$$

Define the second partials of h as the Hessian matrix, $H_\ell(x)$ for $\ell = 1, \dots, m$, where the ij th entry of H_ℓ is

$$H_\ell(x)_{ij} = \frac{\partial^2 h_\ell}{\partial x_i \partial x_j} \quad (4-1-6)$$

The function f may now be expanded in a Taylor series about the conditional mean $\hat{x}$,

$$f(x) \doteq f(\hat{x}) + A(\hat{x})(x-\hat{x}) + \frac{1}{2} \sum_{i=1}^n \phi_i(x-\hat{x})' F_i(x-\hat{x}) \quad (4-1-7)$$

where ϕ_i is the basis vector, $\phi_i = [0 \dots 1 \dots 0]'$ with the one in the i th row. Likewise, the function h may be expanded to give

$$h(x) \doteq h(\hat{x}) + M(\hat{x})(x-\hat{x}) + \frac{1}{2} \sum_{i=1}^m \phi_i(x-\hat{x})' H_i(x-\hat{x}). \quad (4-1-8)$$

Similarly, $G(x,t)Q(t)G'(x,t)$ can be expanded about $\hat{x}$ to give an approximation for the term $E\{GQG'\}$.

Terms that are higher than second order are neglected in these expansions, although they could be included, at least in theory. In practice, the error distributions are assumed to be Gaussian with zero mean; thus all odd-numbered moments vanish [WOZ1], and the fourth-order moments may be decomposed into second-order moments [AND1]. However, it should be noted that the addition of higher-order terms does not greatly affect the filtering equations for small estimation errors since they converge to zero faster than the lower-order terms as the state estimate converges to the state [SOR1]. For larger errors, another approach is suggested in Section 4-3.

For the nonlinear filter, the continuous-time dynamic system can not be transformed to an equivalent discrete-time model as was the case for the linear system by using (2-4-2) through (2-4-4). The differential equations for the extrapolation of the system state estimate must be solved numerically.

The extrapolation equations are found by using the expansions for f in (4-17) and for $E\{QG'\}$ in the moment equations (2-2-4) and (2-2-6):

$$\dot{\hat{x}} = f(\hat{x}, t) + \frac{1}{2} \sum_{i=1}^n \phi_i \text{tr}[F_i(\hat{x})P(t|t_{k-1})], \quad t_{k-1} \leq t < t_k \quad (4-1-9)$$

$$\begin{aligned} \dot{P}(t|t_{k-1}) = & A(\hat{x})P(t|t_{k-1}) + P(t|t_{k-1})A'(\hat{x}) \\ & + E\{G(\hat{x})Q(t)G'(\hat{x})\}, \quad t_{k-1} \leq t < t_k \end{aligned} \quad (4-1-10)$$

The updated equations for the nonlinear filter are found by using the expansion for h in (4-1-8) in the moment equations (2-3-4) and (2-3-5):

$$\begin{aligned} \hat{x}(t_k|t_k) = & \hat{x}(t_k|t_{k-1}) + G(t_k)[y(t_k) - h(\hat{x}(t_k))] - \\ & \frac{1}{2} \sum_{i=1}^m \phi_i \text{tr}[H_i(\hat{x}(t_k))P(t_k|t_k)] \end{aligned} \quad (4-1-11)$$

$$P(t_k|t_k) = P(t_k|t_{k-1}) - G(t_k)M(\hat{x}(t_k))P(t_k|t_{k-1}) \quad (4-1-12)$$

where

$$\begin{aligned} G(t_k) = & P(t_k|t_{k-1})M'(\hat{x}(t_k))[M(\hat{x}(t_k))P(t_k|t_{k-1})M'(\hat{x}(t_k)) \\ & + R(t_k) + O(t_k)]^{-1} \end{aligned} \quad (4-1-13)$$

The term $0(t_k)$ can be approximated as

$$0(t_k) = -\frac{1}{4} \sum_{i=1}^m \phi_i \operatorname{tr}[H_i(\hat{x}(t_k))P(t_k|t_{k-1})] \cdot \sum_{i=1}^m \phi_i' \operatorname{tr}[H_i(\hat{x}(t_k))P(t_k|t_{k-1})] \quad (4-1-14)$$

when only first- and second-order terms are considered, or it may be approximated as

$$0_{ij}(t_k) = \frac{1}{2} \operatorname{tr}[H_i(\hat{x}(t_k))P(t_k|t_{k-1})H_j(\hat{x}(t_k))P(t_k|t_{k-1})] \quad (4-1-15)$$

(0_{ij} is the ij th entry of 0) when the fourth-order central moments are considered. These updated equations are derived in [ATH2, SOR1, and JAZ1], only in slightly different form. The relative performance of linear and nonlinear filters can only be judged by experimental results. In the following section, an estimation problem is chosen that has nonlinearities that are widely varied by making changes in parameter values. In this way, the value of the nonlinear filtering equations are tested.

4-2. An Example of a Nonlinear Filtering Problem

In [WAG1, GRU1, BER1, WIS1, ATH2], the problem of reentry trajectory estimation is discussed. Typically, both the reentry system dynamics and the observation function are severely nonlinear and accurate estimates are difficult to achieve. In this section this problem is discussed and the results are given for a particular example using first-order and second-order filtering.

The Reentry Problem

It is assumed that the force of gravity is negligible when compared to the aerodynamic effects. Thus, the system dynamics for vertical motion of the reentry body can be written as [ALL1, ATH2]

$$\dot{x}_1 = -x_2 \quad (4-2-1)$$

$$\dot{x}_2 = -\frac{C_D A \rho}{2m} x_2^2$$

where x_1 is altitude, x_2 is velocity, C_D is the drag coefficient, A is forward area, ρ is atmospheric density, and m is mass. The atmospheric density is assumed to be an exponential function of altitude,

$$\rho = \rho_0 e^{-\gamma x_1} \quad (4-2-2)$$

where $\gamma = 5 \times 10^{-5} \text{ ft}^{-1}$. Since the ballistic parameter $C_D A \rho / 2m$ is not known it can be estimated by setting

$$x_3 = C_D A \rho_0 / 2m$$

and augmenting the state equations such that

$$\begin{aligned} \dot{x}_1 &= -x_2 \\ \dot{x}_2 &= -x_2^2 x_3 e^{-\gamma x_1} \end{aligned} \quad (4-2-4)$$

$$\dot{x}_3 = 0.$$

It should be noted that no system disturbances are present in this formulation.

The observation equation consists of a scalar radar range measurement taken from an altitude H and a lateral displacement M . The discrete measurements are contaminated with additive white noise so that the expression for the observations is

$$y(t_k) = [M^2 + (x_1(t_k) - H)^2]^{1/2} + v(t_k) \quad (4-2-5)$$

where $v(t_k)$ is assumed to be zero mean with covariance

$$E\{v(t_k)v(t_\ell)\} = R\delta_{k\ell}.$$

For simulation purposes, numerical values were given to the system parameters. Although a great number of cases were tried, a typical case with rather severe nonlinear effects was to offset the radar such that the lateral range is $M = 100,000$ feet and the radar altitude is $H = 100,000$ feet. This evaluation produces the maximum measurement nonlinearity simultaneously to the maximum deceleration nonlinearity. The measurement uncertainty was chosen as $R = 10,000 \text{ feet}^2$. Measurements were assumed to be taken at the rate of one per second.

The initial state vector was chosen as

$$\begin{aligned} x_1(0) &= 300,000 \text{ feet} \\ x_2(0) &= 20,000 \text{ feet/second} \\ x_3(0) &= 10^{-3} \text{ feet}^{-1} \end{aligned} \quad (4-2-6)$$

Initial values for the state estimates were

$$\begin{aligned} \hat{x}_1(0) &= 300,000 \text{ feet} \\ \hat{x}_2(0) &= 20,000 \text{ feet/second} \\ \hat{x}_3(0) &= 3 \times 10^{-5} \text{ feet}^{-1} \end{aligned} \quad (4-2-7)$$

In addition, the initial estimates were assumed to be uncorrelated; thus the initial state covariance matrix is diagonal with elements

$$\begin{aligned} p_{11}(0) &= 10^6 \text{ feet}^2 \\ p_{22}(0) &= 4 \times 10^6 \text{ feet}^2/\text{second}^2 \\ p_{33}(0) &= 10^{-4} \text{ feet}^{-2} \end{aligned} \quad (4-2-8)$$

These particular values are identical to those used by Athans et al. [ATH2] and make a rather close comparison possible for computer simulation studies. In addition, by varying the values of H and M as well as the above initial values, a wide range of nonlinearities is effected.

Reentry Simulation Results

As stated in [ATH2], "The validity and relative advantages and disadvantages of --- (ad hoc assumptions for nonlinear filtering) --- can only be evaluated and analyzed based on the evidence of real or simulation results for particular systems." With this viewpoint, a large number of simulations were performed for the nonlinear problem of reentry estimation.

Since the estimation error depends on the particular random noise sample function, it was necessary to perform a Monte Carlo study. As opposed to the linear filtering case, the RMS error can not be assumed to be identical to the square root of the appropriate covariance matrix variance term since the nonlinear filter produces only approximate minimum-variance estimates and covariance estimates. For this reason several different statistics were found from the Monte Carlo study, including the average of the estimation error, the average of the absolute value of the estimation error, and the RMS estimation error, defined as

$$\text{RMS}(t_k) = \left[\frac{1}{N} \sum_{i=1}^N (x(t_k) - \hat{x}(t_k))^2 \right]^{1/2} \quad (4-2-9)$$

where $\text{RMS}(t_k)$ is a vector. The value of N, the number of trials, was chosen to be twenty.

The results of the Monte Carlo study using the previously given initial conditions are depicted for the linear filter in Fig. 4-1 and for the second-order filter in Fig. 4-2. For this particular simulation, the absolute value of the average error and the RMS error can be compared to results obtained by Athans et al. [ATH2]. Overall the results are similar, but it is interesting to note that whereas both the average and RMS final values of errors are within 10% of the respective values obtained in [ATH2] the average values have a much smoother shape, indicating that more cancellation is taking place in the averaging process than in [ATH2].

Several aspects of the Monte Carlo study of the linear and nonlinear filtering problem should be noted.

1. The second-order filter was more accurate than the first-order filter, which tended to diverge as shown in the RMS curve of Fig. 4-1.
2. The effect of the second-order terms in the state estimate update equation (4-1-11) and the covariance update equation (4-1-12) were nearly negligible whereas the state estimate extrapolation nonlinear terms (4-1-9) produced a significant improvement in the estimation accuracy.
3. Both the linear and nonlinear filter accuracy appeared to be virtually independent of the difference between computing with regular or extended precision.
4. Although results are not shown for velocity x_2 , and ballistic parameter x_3 , the contrast between linear and nonlinear filtering parallels that for altitude in that the second-order filter shows a marked improvement in estimation accuracy.

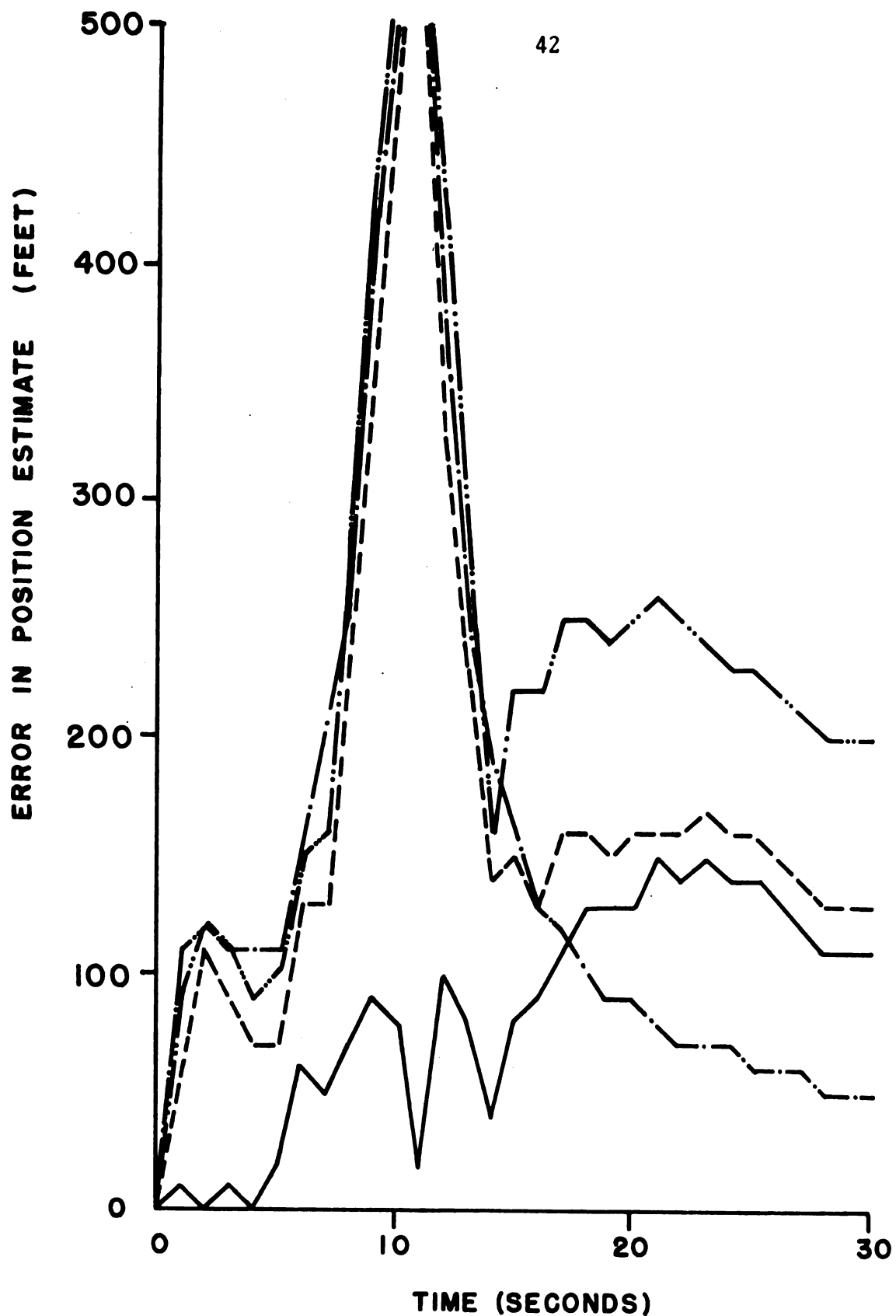


Fig. 4-1. Altitude Estimation Error for the First-Order Filter of the Reentry Estimation Problem.

— Absolute Value of Average Error
 - - - Average of the Absolute Value of Error
 - . . . Square Root of the Filter Altitude Variance
 RMS Error

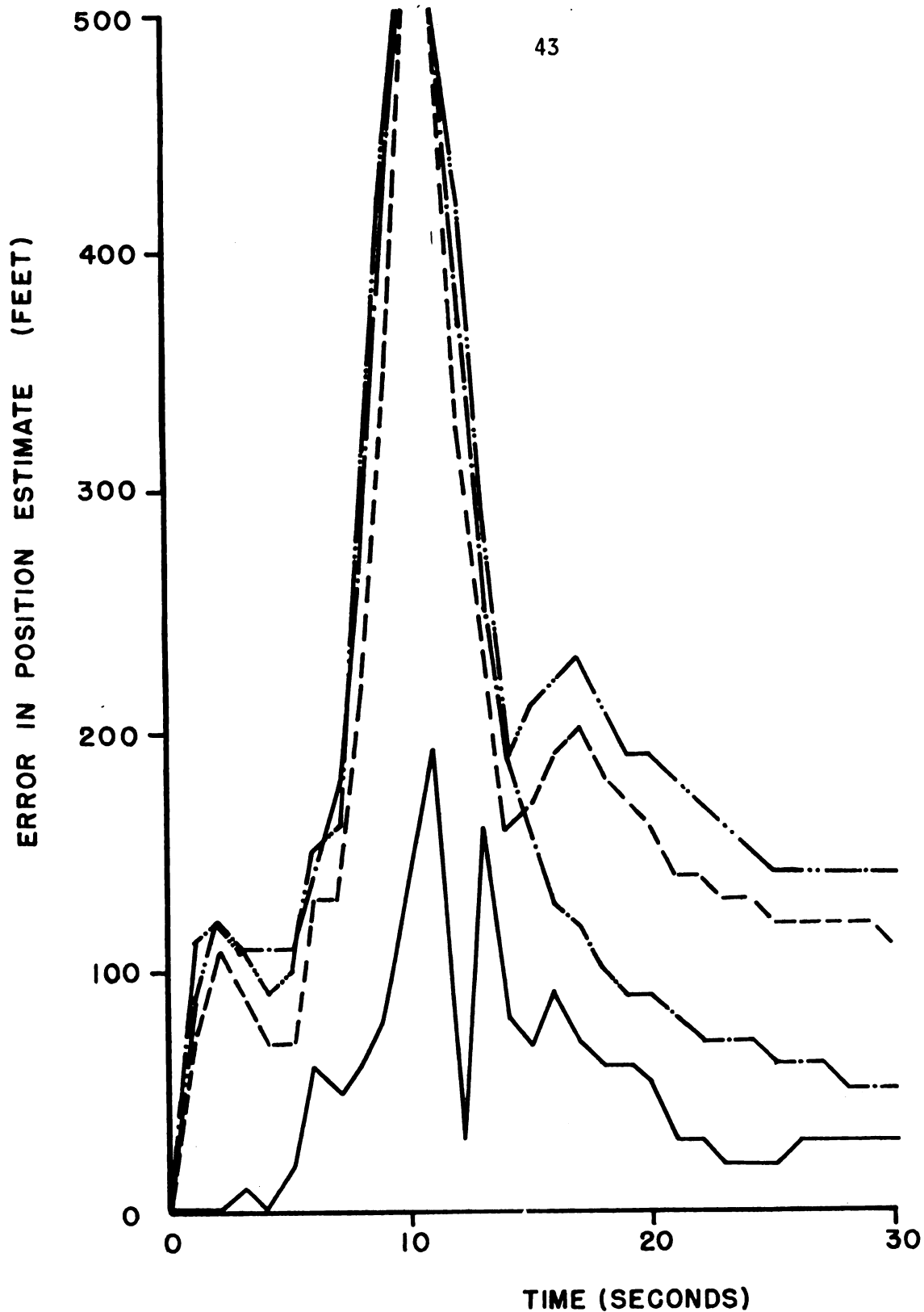


Fig. 4-2. Altitude Estimation Error for the Second-Order Filter of the Reentry Estimation Problem.

- Absolute Value of Average Error
- - - Average of the Absolute Value of Error
- · - · - Square Root of the Filter Altitude Variance
- · · · · RMS Error

5. Neither filter exhibited an RMS error that converged as rapidly as the corresponding square root of the diagonal term of the covariance matrix, although the nonlinear filter reduced the discrepancy as compared to the linear filter.

Other comparisons may be made but they are reserved until the next sections where other factors of filtering theory and filter divergence are discussed.

4-3. Adaptive Filtering for Nonlinear Systems

In the customary development of linear and nonlinear filtering, it is assumed that the system parameters--state coefficients, initial state estimation covariance, measurement coefficients, system disturbance covariance, and measurement noise covariance--are known precisely for the entire filtering time interval. In addition, it is assumed that computation routines are exact, and finally, in the case of nonlinear filtering, that the approximate filtering equations are accurate enough that the Gaussian assumption for the estimation error distribution is satisfactory.

In a large class of filtering applications these assumptions are not warranted. The initial conditions may not be available or the measurement noise and system disturbance covariances may change unexpectedly. And of course, the inaccuracies of numerical solutions and using only approximate filtering must be considered.

It was shown in Section 3-4 that most approaches to the problem of filter divergence are not on-line solutions--they depend on prior information to the filtering process. Although the method suggested by

Jazwinski [JAZ2] is on-line, it applies to linear systems with scalar measurements and without system disturbances. In addition, all of these filter modifications have the effect of reducing the rate of convergence since the state covariance matrix is a priori prevented from converging.

As reported in Denham and Pines [DEN1] and Wishner [WIS1], another approach to filtering a nonlinear system is to use the first-order filter and perform iterations at the point of measurement, called the iterated extended Kalman filter. Simulation results [WIS1] have shown that iterations are as effective as second-order filtering in reducing divergence. However, it should be pointed out that iterations increase the computational load and limit the measurement sampling rate. For this reason, in many systems, the better solution is to use the non-iterative first- or second-order filter and to use a faster measurement sampling rate. Thus, a requirement that is placed on divergence correction techniques is that the computation load must not be significantly increased.

As is the case in the linear filtering problem, Section 3-4, the only measure of filter convergence performance is to compare the statistics of the extrapolated measurement estimation error with the covariance of the measurement estimation error. In the nonlinear case, an approximation must be used to obtain the extrapolated measurement estimation error,

$$\begin{aligned} e(t_k) &= y(t_k) - \hat{y}(t_k) \\ e(t_k) &= y(t_k) - M(\hat{x}(t_k))\hat{x}(t_k|t_{k-1}). \end{aligned} \tag{4-3-1}$$

The measurement estimate error covariance matrix $Z(t_k|t_{k-1})$ is defined as

$$\begin{aligned} Z(t_k|t_{k-1}) &= E\{e(t_k)e'(t_k)\} \\ &= M(\hat{x}(t_k)P(t_k|t_{k-1})M'(\hat{x}(t_k)) + R(t_k) \end{aligned} \quad (4-3-2)$$

The statistics to be used for the extrapolated measurement estimation error are defined to be of the form

$$S(t_k) = \sum_{i=0}^N \alpha_i e(t_{k-i})e'(t_{k-i}) \quad (4-3-3)$$

where N is a constant and α_i is a weighting sequence on the outer product of the past extrapolated measurement estimation errors such that

$$\sum_{i=0}^N \alpha_i = 1.$$

In the case of ideal filtering, $\text{tr } S(t_k)$ should be comparable to $\text{tr } Z(t_k|t_{k-1})$, that is, for Gaussian distribution of $S(t_k)$, $\text{tr } S(t_k)$ should be less than $\text{tr } Z(t_k|t_{k-1})$ on an average of 68.2% of the measurements since the diagonal elements of Z are variances. In the case of filter divergence, $S(t_k)$ will be on the average large compared to $Z(t_k|t_{k-1})$ and the above will not hold.

Now that this manifestation of filter divergence is realized, it must be decided how this quantity is to be utilized to adapt the filter to the real-time measurements.

The author's approach is to observe that the divergence of the filter represents the fact that $Z(t_k|t_{k-1})$ has converged to a smaller value than warranted by actual data, and this in turn means that the

state covariance matrix $P(t_k | t_{k-1})$ has converged to a smaller value than warranted, which has the effect of diminishing the filter gain matrix $G(t_k)$. The discrepancy in the diverging filter can thus be alleviated by using a gain correction matrix $G_c(t_k)$ to compute the adaptive filter gain $G_a(t_k)$,

$$G_a(t_k) = G_c(t_k)G(t_k) \quad (4-3-4)$$

Then the updated adaptive state estimate is

$$\hat{x}_c(t_k | t_k) = \hat{x}(t_k | t_{k-1}) + G_a(t_k)e(t_k) \quad (4-3-5)$$

The gain correction matrix is selected such that the corrected error

$$e_c(t_k) = y(t_k) - M \hat{x}_c(t_k | t_k) \quad (4-3-6)$$

is driven toward zero.

Consider the case of a scalar system. If the state variance has converged incorrectly to a value $P(t_k)/C(t_k)$ where $C(t_k)$ is the divergence factor, then the gain $G(t_k)$ is

$$G(t_k) = P(t_k)M / (M^2P(t_k) + C(t_k)R(t_k)) \quad (4-3-7)$$

and the gain correction should be

$$G_c(t_k) = \frac{M^2P(t_k) + C(t_k)R(t_k)}{M^2P(t_k) + R(t_k)} \quad (4-3-8)$$

A good measure of the divergence C is

$$C(t_k) = \frac{S(t_k)}{Z(t_k | t_{k-1})} \quad (4-3-9)$$

Thus, for the scalar system, the gain correction can be computed exactly based on the measurement error statistic $S(t_k)$. However, it should be observed that the value of $S(t_k)$ is dependent upon the measurement noise sample function and that it is only an approximation to the true measurement estimation error variance. Therefore, a reasonable approximation to (4-3-8) is

$$G_c(t_k) = \left[\frac{S(t_k)}{Z(t_k|t_{k-1})} \right] \frac{1}{2} \quad (4-3-10)$$

The general vector filtering problem may be handled identically to the above except the case of multivariable measurements requires some modification. One modification is to reduce the measurements to scalar measurements processed sequentially in time. Another approach that may be considered is to use cross-correlation to identify those components of the state vector that are correlated with the diverging measurements estimates and apply the gain correction in accordance with the strength of the cross-correlation.

An important property of the diverging filter is that the divergence implies that the gain matrix is too small due to the unwarranted convergence of the state matrix. Thus, by simply testing a criterion on divergence and increasing the gain by an appropriate factor, such as (4-3-10), the filter adapts to the measurement errors without the need to increase the state covariance.

In order to assure filter stability, the gain correction is tested to assure that oscillations do not occur. The test is performed by examining the corrected measurement estimation error,

$$e_c(t_k) = y(t_k) - Mx_c(t_k). \quad (4-3-11)$$

If this error is of the same sign as $e(t_k)$, then the adaptive estimate is unaltered. However, if the corrected error is of the opposite sign as $e(t_k)$, then the adaptive estimate has possibly been over-corrected and to assure the filter will be stable, the adaptive gain is reduced such that $e_c(t_k) = 0$. Therefore, the adaptive filter gain is

$$G_a(t_k) = G_c(t_k)G(t_k), \quad e(t_k)e_c(t_k) \geq 0 \quad (4-3-12)$$

$$G_a(t_k) = \frac{|e(t_k)|}{|e(t_k)| + |e_c(t_k)|} \cdot G_c(t_k)G(t_k), \quad e(t_k)e_c(t_k) < 0$$

Using this method of adapting the filter to the measurement estimation error has the advantages that (1) it is not necessary to increase the state covariance, which reduces or obviates filter convergence, (2) it can be used even in the case that system disturbances are present, and (3) the increased computational load is negligible with respect to the normal amount of filtering computation.

1-4. A Discussion and an Example of Adaptive Nonlinear Filtering

A criterion to establish the performance of a filter is to compare the RMS error of the estimates obtained by a Monte Carlo study with the square root of the variance obtained from the filter state covariance matrix. But first, it is necessary to establish the validity of the approximate filter covariance terms.

Although the accuracy of the covariance matrix cannot be explicitly demonstrated, on the other hand it can be seen that the differential equations governing its extrapolation are smoother than those of the state estimate. This point has been supported by experimental evidence in several ways.

1. The covariance matrix does not vary significantly over the runs used in the Monte Carlo study.
2. The covariance matrix does not vary regardless of whether the first- or second-order filter is being used.
3. The covariance matrix is independent of the integration accuracy for widely varying integration intervals.

Any of these factors would greatly affect the state estimation error, but the constancy of the covariance matrix supports the hypothesis that the approximate filter covariance is nearly exact.

In the case of the second-order filter it can be seen from Figs. 4-1 and 4-2 for the reentry problem that the RMS error is much closer to the covariance matrix values than is the case for the first-order filter. Nonetheless the remaining discrepancy is still large, and the purpose of using adaptive filtering in the reentry estimation problem is to reduce this difference.

Fig. 4-3 shows the altitude estimation error for the reentry problem and is equivalent to Fig. 4-1 except the linear filter is replaced by the linear adaptive filter. The results are significantly improved over the linear filter; in fact, the RMS error is even less than the RMS error for the second-order filter is shown in Fig. 4-2. Thus, in this particular problem, the benefit of using this adaptive filter is

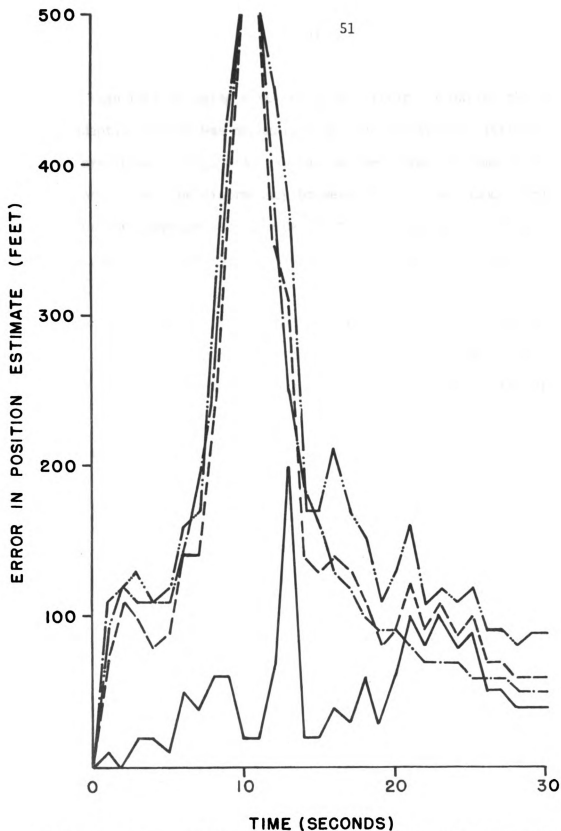


Fig. 4-3. Altitude Estimation Error for the First-Order Adaptive Filter of the Reentry Estimation Problem.

— Absolute Value of Average Error
 - - - Average of the Absolute Value of Error
 - · - Square Root of the Filter Altitude Variance
 · · · RMS Error

greater than that of using a second-order filter. Finally, the second-order adaptive filter was mechanized and the equivalent altitude errors are shown in Fig. 4-4. As can be seen from the time-history of the error curves, the discrepancy between RMS and covariance-originated terms for the adaptive second-order filter is reduced by a factor of four as compared to the second-order filter in the latter part of the run.

The four cases are summarized in the curves shown in Fig. 4-5. The RMS altitude estimation error is shown for the linear, nonlinear, adaptive, and non-adaptive filters as obtained from the Monte Carlo studies.

Several points should be noted with regard to these simulation results.

1. Although only x_1 , the altitude estimation error, is shown in Figs. 4-1 through 4-5, similar results were found for x_2 , the velocity estimation error, and x_3 , the ballistic parameter, with regard to relative filter performance.
2. The value of $S(t_k)$ for these particular results was based on a single sample of measurement error (i.e., $N = 1$ in (4-3-3)), although similar results have been obtained for other trials, such as an exponential weighting function α_i in (4-3-3). The difference in results was slight.
3. The effect of using extended precision and of including a greater number of runs in the Monte Carlo study was negligible.

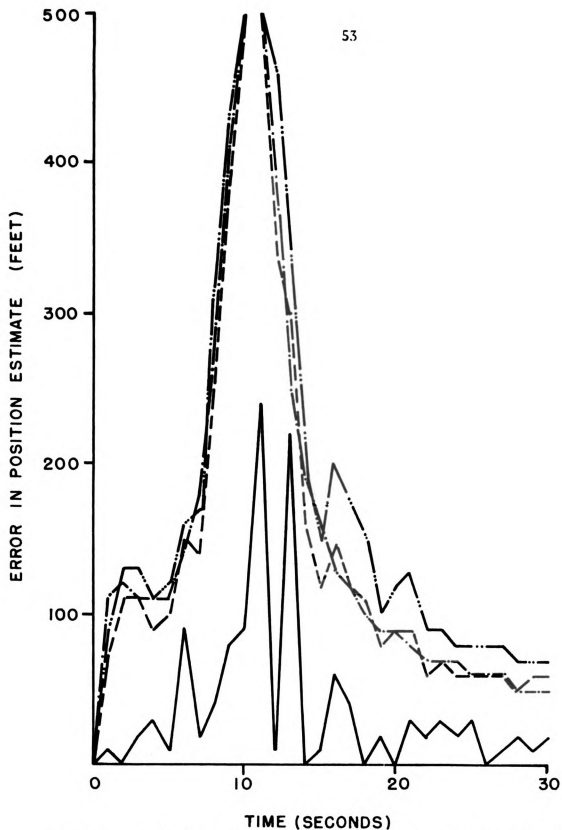


Fig. 4-4. Altitude Estimation Error for the Second-Order Adaptive Filter of the Reentry Estimation Problem.

- Absolute Value of Average Error
- - - Average of the Absolute Value of Error
- Square Root of the Filter Altitude Variance
- RMS Error

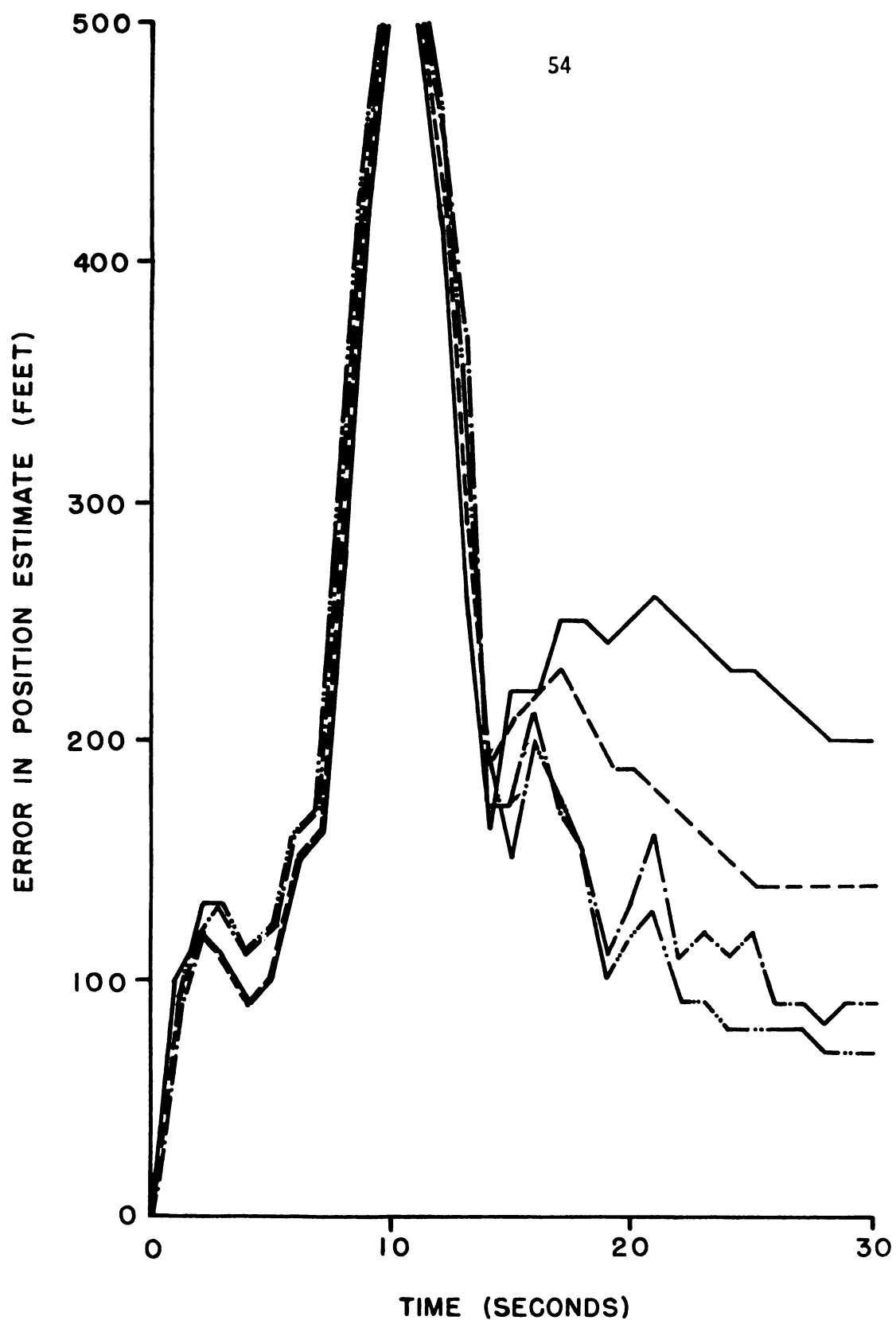


Fig. 4-5. RMS Altitude Estimation Error for Four Similar Filters.

- First-order Non-adaptive Filter
- Second-order Non-adaptive Filter
- First-order Adaptive Filter
- Second-order Adaptive Filter

4. The simulation results did not depend on the noise generator or on the word length; the results obtained on the Michigan State University IBM 1800 computer were substantiated with further runs on the Purdue University CDC 6500 computer.

In the above study, it should be noted that the assumed values of initial state and initial state variance were consistent; that is, the initial estimate of each state component was within one standard deviation of the true state, also, the sample noise function was consistent with the assumed noise variance, and finally, the integration routine was virtually exact. Thus the only source for causing divergence was the approximate filtering equations since the only inexactly known parameter x_3 was included in the augmented state. The other possible uses of adaptive filtering become apparent when other system inconsistencies are considered. For example, it was found that a much coarser value of the integration interval could be used for the adaptive filter without eliciting the problem of divergence that was evident even in the second-order non-adaptive filter. Another example is that even for the second-order filter it was found that the initial state error could only deviate as much as three standard deviations of the initial dispersion estimate for convergence to be assured and this was for a relatively small initial state covariance matrix (4-2-8) whereas for the adaptive filter a large initial error and covariance matrix could be assumed without filter divergence.

Several aspects of adaptive filtering are worthy of further examination. Using the statistic $S(t_k)$ (4-5-3) leads to a very difficult question about what is the optimum weighting sequence

α_i , $i = 1, \dots, N$ and what is the optimum value of N . It appears that N should be small compared to the number of measurements that are required for the system to transcend a sizeable nonlinear range; in addition, for heavily weighted past data, the effect of previous gain correction on the state estimate should be considered. The proper function for the adaptive filter gain (4-3-12) is not analytically determined and apparently must be determined empirically.

In summary, the relative performance of the nonlinear filter and the adaptive filter as they are developed in this chapter is shown in Fig. 4-5 for state estimation of the atmospheric reentry problem and the increase in accuracy for either filter is demonstrated while the best performance is obtained by both nonlinear and adaptive filtering.

V. Conclusions

The purpose of this thesis is the investigation of problems involved in state estimation of a class of dynamic systems.

5-1. Thesis Review

The initial problems include determining a suitable mathematical model for the system and stochastic model for the external disturbances. Due to the universality of the Brownian motion process for modeling real-world phenomena, as well as the availability of useful analytic results, the stochastic disturbance model is assumed to be a Brownian motion process (Appendix A and Section 2-1). Due to the large class of dynamic systems that can be modeled by a set of ordinary differential equations, the differential model (2-1-1) is chosen for the dynamic system model. The output of the system is assumed to be sampled periodically as is the case in digital computer monitors and optimal controllers; therefore, the discrete-time observation model (2-1-10) is used. And finally, it is assumed that real-time estimates are necessary as in the case of an on-line controller; thus the system state estimator is assumed to be recursive. The results and conclusions in this thesis are based on these assumptions.

Using the above assumptions, the filtering equations are developed, based on Kolmogorov's forward equation (2-2-1), Ito's Theorem A-3 and Bayes' rule (2-3-1). The worth of the filter mechanization is determined

by digital computer simulation; for this reason, the discrete-time modeling of continuous-time stochastic processes (Section 2-4) is derived.

In the case of linear filtering problems, two methods are proposed for reducing computation time and increasing accuracy. An example of state estimation of a linearized model of a steam turbo-generator is given. And finally, methods are discussed for handling the problem of filter divergence due to inexactly known models.

In the case of nonlinear filtering, the nonlinear filter is developed and used to estimate the state of a reentry body. The adaptive filter is suggested as a method to reduce the error in the state estimate.

5-2. Results and Conclusions

The major results and conclusions of this thesis are based on over 200 simulations of the steam turbo-generator and the reentry body as well as the results that have been referred to in the literature.*

1. The computational methods proposed in Section 3-2 for state and state covariance matrix extrapolation can significantly reduce computation time and increase accuracy in real-time state estimation. The actual improvement depends on the size of the state and is more significant with a larger dimensional system with more than two or three state variables.

*In addition, some preliminary results on adaptive filtering of the Van der Pol oscillator show favorable convergence properties.

2. The results of the digital computer simulation (Fig. 3-2) suggest that real-time state estimation of a steam turbo-generator is feasible either for optimal control of a single machine or for area-wide control.
3. The second-order filter developed in Section 4-1 successfully estimated the state of a reentry body as determined by the digital computer simulation. In all cases where the nonlinearities are significant, the second-order filter has significantly less error than the first-order filter. Examples of this are shown in Figs. 4-1 and 4-2 for Monte Carlo studies of reentry estimation.
4. The nonlinear filter is superior to the linear filter in the reentry estimation simulation due to the second-order state estimate extrapolation term; the second-order terms in state update and covariance update have virtually no effect on filter performance.
5. The adaptive filter suggested in Section 4-3 successfully reduces the estimation error significantly as compared to the second-order filter for both linear and nonlinear adaptive filtering (Fig. 4-3 and 4-4). The fact that the adaptive second-order filter has an RMS error that is comparable to the filter square-rooted variance indicates that within the given structure, any further improvements in filter accuracy will be slight. These results are summarized in Fig. 4-5.
6. The adaptive filter as developed in this thesis is an adaptive gain filter; i.e., only the filter gain is adapted to the divergence which is manifested in the extrapolated measurement estimation

error. Other methods of divergence compensation such as adaptive noise covariance do not appear to be nearly as satisfactory as adaptive gain filtering according to the simulation efforts of this research effort.

5-3. Suggestions for Further Research

In the area of linear filtering, most of the research effort lies in the area of applications and in filtering of inexactly modeled processes. Naturally, the problem of state estimation and concurrent system identification are not resolved and require further definition. (See Monahan [MON1] for a thorough discussion of system identification.) An example is the problem of estimation and identification where the observation transformation is singular and the system description is only partially known.

In the applications area, more data are required for further refinement of the steam turbo-generator state estimator presented in Section 3-3. An obvious extension is to simulate cascading an optimal controller with the optimal filter and to test the resulting improvement in the quadratic performance index as compared to presently used feedback regulation. Furthermore, the methods of Foshe and Elgerd [FOS1] for inter-area optimal control could be mechanized by coupling the optimal controller with the optimal filter.

The second-order filter has been reviewed and its value has been reviewed and its value has been demonstrated in this thesis. A natural question arises as to the utility of mechanizing even higher-order

filters. Although this topic is certainly an area for future research, the answer appears to be in the negative for several reasons. Higher-order filtering requires higher-order derivatives, which in practice may be difficult to compute either analytically or numerically. Also, higher-order filtering requires a great deal more computation (See [SOR1], for example), especially for a larger system. And finally, higher-order terms appear to have a limited effect in filtering convergence.

The use of adaptive filtering appears to be more promising for improving filter performance of severely nonlinear systems, but this conclusion should be tested over a wide range of examples and conditions. In addition, the exact form of the optimal adaptive filter is an open question and is the subject of further research.

LIST OF REFERENCES

LIST OF REFERENCES

- [ALL1] H.J. Allen and A.J. Eggers, Jr., "A Study of the Motion and Aerodynamic Heating of Missiles Entering the Earth's Atmosphere at High Supersonic Speeds," NACA, Washington, D.C., Tech. Note 4047, October, 1957.
- [AND1] T.W. Anderson, "An Introduction to Multivariate Statistical Analysis," New York, Wiley, 1958.
- [AND2] W.N. Anderson, Jr., G.B. Kleindorfer, P.R. Kleindorfer and M.B. Woodroffe, "Consistent Estimates of the Parameters of a Linear System," The Annals of Mathematical Statistics, Vol. 40, No. 6 December, 1969, pp. 2064-2075.
- [ATH1] M. Athans and E. Tse, "A Direct Derivation of the Optimal Linear Filter Using the Maximum Principle," IEEE Trans. Automatic Control, Vol. AC-12, pp. 690-698, December, 1967.
- [ATH2] M. Athans, R.P. Wishner, and A. Bertolini, "Suboptimal State Estimation for Continuous-time Nonlinear Systems from Discrete Noisy Measurements, 1968 Joint Automatic Control Conf., Ann Arbor, Michigan, pp. 364-382, June, 1968.
- [BAR1] M.S. Bartlett, Stochastic Processes, London, Cambridge University Press, 1961.
- [BEL1] R.E. Bellman, Adaptive Control Processes, pp. 129-131, 152-158, Princeton University Press, Princeton, New Jersey, 1961.
- [BER1] A. Bertolini, "A Trajectory Analysis Program," Technical Note 1967-48, Lincoln Laboratory, MIT, Lexington, Mass., September, 1967.
- [BIE1] G.J. Bierman, "A Remark Concerning Discrete Approximation to White Noise and Covariance Calculation," IEEE Trans. Automatic Control 14, 432-433, 1969.
- [BIR1] M.W. Bird, "Asymptotic Convergence of Nonlinear, Continuous-time Filters," Ph.D. Thesis, Michigan State University, 1969.

- [BRA1] F.H. Branin, Jr., "Computer Methods of Network Analysis," Proc. IEEE, Vol. 55, No. 11, pp. 1787-1801, November, 1967.
- [BRY1] A.E. Bryson, Jr. and Y.C. Ho, Optimization, Estimation and Control, Blaisdell Publishing Co., 1969.
- [BRY2] A.E. Bryson and D.E. Johansen, Linear Filtering for Time-Varying Systems Using Measurements Containing Colored Noise, IEEE Trans. Automatic Control 10, 4-10 (1965).
- [BUC1] R.S. Bucy, "Optimum finite-time filters for a special non-stationary class of inputs," Johns Hopkins University, Appl. Phys. Lab., Baltimore, Md., Internal Memo. BBD-600, 1959.
- [BUC2] R.S. Bucy, "Nonlinear filtering theory," IEEE Trans. Automatic Control, (Correspondence), vol. AC-10, p. 198, April, 1965.
- [CHA1] S.Y. Chan and K. Chuang, "A Study of Effects on Optimal Control Systems Due to Variations in Plant Parameters," Int. J. Control, 1968, Vol. 7, No. 3, pp. 281-298.
- [COC1] W.G. Cochran, and G.M. Cox, Experimental Designs, New York: John Wiley and Sons, 1957.
- [COX1] H. Cox, "On the Estimation of State Variables and Parameters for Noisy Dynamic Systems," IEEE Trans. Automatic Control 9, 5-12, 1964.
- [DEN1] W.F. Denham and S. Pines, "Sequential Estimation when Measurement Function Nonlinearity is Comparable to Measurement Error," AIAA J. 4, 1071-1076 (1966).
- [DEN2] A.R. Dennis, "Functional Updating and Adaptive Noise Variance Determination in Recursive-type Trajectory Estimators," Special Projects Branch Astrodynamics Conf. NASA-GSFC, Greenbelt, Maryland, May, 1967.
- [DOO1] J.L. Doob, Stochastic Processes. New York: Wiley, 1953.
- [DYN1] E.B. Dynkin, Markov Processes, Springer, Berlin, 1965.
- [FEL1] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. II, Wiley, New York, 1966.
- [FER1] J.D. Ferguson and Z.V. Rekasius, "Optimal Linear Control Systems with Incomplete State Measurements," Proc. 6th Ann. Allerton Conf. on Circuit and System Theory (Urbana, Ill., 1968), p. 135, April, 1969.

- [FIT1] R.J. Fitzgerald, Error Divergence in Optimal Filtering Problems, Second IFAC Symp. Automatic Control Space, Vienna, Austria, 1967.
- [FOS1] C.E. Fosha, Jr. and O.I. Elgerd, "The Megawatt-Frequency Control Problem--A New Approach Via Optimal Control Theory," Dept. of Elec. Engr., University of Florida, Gainesville, Florida.
- [FRA1] J.S. Frame, "Matrix Functions and Applications," IEEE Spectrum, Vol. 1, No. 6, pp. 123-131, June, 1964.
- [FRA2] J.S. Frame and M.A. Needler, "State and Covariance Matrix Extrapolation," Proc. IEEE (to be published).
- [FRI1] B. Friedland and I. Bernstein, "Estimation of the State of a nonlinear process in the presence of nongaussian noise and disturbances," J. Franklin Inst., vol. 281, pp. 455-480, June, 1966.
- [GAN1] S. Ganapathy and A. Suba Rao, "Transient Response Evaluation from the State Transition Matrix," Proc. IEEE, Vol. 57, No. 3, pp. 347-349, March, 1969.
- [GAU1] F.K. Gauss, Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections. New York: Dover Publications, Inc., 1963 (reprint).
- [GRU1] M. Gruber, "An Approach to Target Tracking," M.I.T. Lincoln Lab., Lexington, Mass. Tech. Note 1967-8, February, 1967.
- [HO1] Y.C. Ho, "On the Stochastic Approximation Method and Optimal Filtering Theory," Journal of Mathematical Analysis and Applications 6, pp. 152-154, 1962.
- [HO2] Y.C. Ho and R.C.K. Lee, "A Bayesian approach to problems in stochastic estimation and control," IEEE Trans. Automatic Control, vol. AC-9, pp. 333-339, October, 1964.
- [ITO1] K. Ito, "Lectures on Stochastic Processes," Tata Institute for Fundamental Research, Bombay, 1961.
- [ITO2] K. Ito and H.P. McKean, Jr., Diffusion Processes and Their Sample Paths, Academic Press, New York, 1964.
- [JAZ1] A.H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.
- [JAZ2] A.H. Jazwinski, Adaptive Filtering, Proc. IFAC Symp. Multi-variable Control Systems, Dusseldorf, Germany, October, 1968, 2, 1-15.

- [JOS1] R.D. Joseph and J.T. Tou, "On Linear Control Theory," AIEE Trans. Applications and Industry, 80 (1961), pp. 193-196.
- [KAI1] T. Kailath and P. Frost, "Mathematical Modeling of Stochastic Processes," Stochastic Problems in Control, 1968 Joint Automatic Control Conference, University of Michigan, Ann Arbor, Mich., ASME, pp. 1-38, June, 1968.
- [KAL1] R.E. Kalman, "A new approach to linear filtering and prediction problems," Trans. ASME, J. Basic Engrg., vol. 82, pp. 34-45, March, 1960.
- [KAL2] R.E. Kalman, "New methods in Wiener filtering theory," Proc. 1st Symp. on Engrg. Applications of Random Function Theory and Probability, J.L. Bogdanoff and F. Kozin, Eds. New York: Wiley, 1963.
- [KAL3] R.E. Kalman and R.S. Bucy, "New results in linear filtering and prediction theory," Trans. ASME, J. Basic Engrg., ser. D vol. 83, pp. 95-107, December, 1961.
- [KIR1] L.K. Kirchmayer, Economic Control of Interconnected Systems, Wiley, New York, 1959.
- [KOL1] A.N. Kolmogorov, "Interpolation and extrapolation of stationary random sequences," Bull. Acad. Sci. USSR, Math. Ser. vol. 5, 1941. A translation has been published by the RAND Corp., Santa Monica, Calif., as Memo. RM-3090-PR.
- [KUS1] H.J. Kushner, "On the differential equations satisfied by conditional probability densities of Markov processes, with applications," J. SIAM Control, vol. 2, pp. 106-119, 1964.
- [KUS2] H.J. Kushner, "Dynamical equations for optimum nonlinear filtering," J. Differential Equations, vol. 3, pp. 179-190, 1967.
- [KUS3] H.J. Kushner, "Nonlinear filtering: the exact dynamical equations satisfied by the conditional mode," IEEE Trans. Automatic Control, vol. AC-12, pp. 262-267, June, 1967.
- [LEE1] E.B. Lee and L. Markus, Foundations of Optimal Control Theory, Wiley, New York, 1967.
- [LEE2] R.C.K. Lee, Optimal Estimation, Identification, and Control, Research Monograph No. 28, M.I.T. Press, 1964.
- [LEV1] A.H. Levis, "Some Computational Aspects of the Matrix Exponential," IEEE Trans. of Automatic Control, Vol. AC-14, No. 4, pp. 410-411, August, 1969.

- [LIO1] M.L. Liou, "A Novel Method of Evaluating Transient Response," Proc. IEEE, Vol. 54, No. 1, pp. 20-23, January, 1966.
- [MAS1] E.J. Mastascusa, "A Relation Between Liou's Method and the Fourth-order Runge-Kutta Method for Evaluating Transient Response," Proc. IEEE, Vol. 57, No. 5, pp. 803-804, May, 1969.
- [MCK1] H.P. McKean, Jr., Stochastic Integrals, Academic Press, New York, 1969.
- [MON1] D.D. Monahan, "System Identification by Stochastic Approximation," Ph.D. Thesis, Michigan State University, 1970.
- [MOW1] V.O. Mowery, "Least squares recursive differential correction estimation in nonlinear problems," IEEE Trans. Automatic Control, vol. AC-9, pp. 399-407, October, 1965.
- [NAY1] T.H. Naylor, J.L. Balintty, D.S. Burdick, K. Chu, Computer Simulation Techniques, Wiley, New York, 1966.
- [PAP1] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.
- [PAR1] G.L. Park, J. Preminger, A.P. Wurmlinger, and W. Panyan, "A Deterministic and Probabilistic Analysis of Response of Power Systems to Load Changes," Prog. Rt. 2, Div. Eng. Res., Michigan State University, East Lansing, Michigan, October, 1967.
- [PAR2] G.L. Park, et al, "A Deterministic and Probabilistic Analysis of Response of Power Systems to Load Changes," Final Report, Div. of Engr. Research, Michigan State University, East Lansing, Michigan, October, 1969.
- [PEA1] J.B. Pearson and C.Y. Ding, "Compensator Design for Multi-variable Linear Systems," IEEE Trans. Automatic Control, Vol. AC-11, p. 130, April, 1969.
- [PIN1] S. Pines, H. Wolf, A. Bailie, and J. Mohan, "Modifications of the Goddard Minimum Variance Program for the Processing of Real Data," Analytical Mechanics Associates, Inc., Westbury, New York, Rept. Contract NAS 5-2535, October, 1964.
- [POT1] J.E. Potter, "A Guidance-Navigation Separation Theorem," Rep. Re-11, Experimental Astronomy Laboratory, Mass. Institute of Technology, Cambridge, 1964.
- [PUR1] P.E. Purser, M.A. Faget, N.F. Smith, Eds., Manned Spacecraft: Engineering Design and Operation, Fairchild Publications, Inc., New York, 1965, pp. 226-231.

- [RAO1] V.P.C. Rao and S. Ganapathy, 'Comments on Transient Evaluation from the State Transition Matrix,' Proc. IEEE, Vol. 58, No. 5, p. 814, May, 1970.
- [SAR1] V.V.S. Sarma and B.L. Deekshatula, 'Optimal Control When Some of the State Variables Are Not Measurable,' Int. J. Control, 1968, Vol. 7, No. 3, pp. 251-256.
- [SCH1] F.H. Schlee, C.J. Standish and N.F. Toda, Divergence in the Kalman Filter, AIAA J. , 5, 1114-1120, 1967.
- [SCH2] F.C. Schweppe, 'Algorithms for estimating a reentry body's position, velocity, and ballistic coefficient in real time or for post flight analysis,' M.I.T. Lincoln Lab., Lexington, Mass., Group Rept. 1964-64, DDC609524, December, 1964.
- [SMI1] G.L. Smith, S.F. Schmidt and L.A. McGee, 'Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle,' NASA TR R-135, 1962.
- [SOR1] H.W. Sorenson and A.R. Stubberud, 'Recursive Filtering for Systems with Small but Non-negligible Non-linearities,' Int. J. Control, Vol. 7, No. 3, 271-280, 1968.
- [STA1] K.N. Stanton, 'Estimation of Turbo-Alternator Transfer Functions using Normal Operating Data,' Proc. IEE, Vol. 112, p. 173, 1965.
- [STR1] R.L. Stratonovich, 'Application of the theory of Markov process for optimum filtration of signals,' Radiotekhn. i. Electron., vol. 5, no. 11, pp. 1751-1763, 1960. (English transl. in Radio Engrg. and Electronics, vol. 1, pp. 1-19.)
- [STR2] R.L. Stratonovich, 'A New Form of Representing Stochastic Integrals and Equations, J. SIAM Control 4, 362-371, 1966.
- [WAG1] W.E. Wagner, 'Re-entry Filtering, Prediction and Smoothing,' J. Spacecraft Rockets 3, 1321-1327, 1966.
- [WIE1] N. Wiener, The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications. New York: Wiley, 1949. Originally issued as a classified report by M.I.T. Radiation Lab., Cambridge, Mass., February, 1942.
- [WIS1] R.P. Wishner, J.A. Tabaczynski and M. Athans, 'On the Estimation of the State of Noisy Nonlinear Multivariable Systems,' Proc. IFAC Symp. on Multivariable Control Systems, Dusseldorf, Germany, October, 1968.

- [WON1] W.M. Wonham, "On the Separation Theorem of Stochastic Control,"
SIAM Journal on Control, May, 1968, Vol. 6, No. 2.
- [WON2] E. Wong, and M. Zakai, "On the Relation Between Ordinary and
Stochastic Differential Equations, Internat. J. Engrg. Sci.
3, 213-229, 1965.
- [WOZ1] J.M. Wozencraft and Irwin M. Jacobs, Principles of Communication
Engineering, Wiley, New York, pp. 205-206, 1965.

APPENDICES

APPENDIX A

Stochastic Processes and Filtering Theory

In random processes, a Gaussian white noise process serves as a model for purely random effects that appear to be pure noise with mutually independent components. However, care must be taken since, by definition of its correlation function, a white-noise process is almost nowhere continuous, and thus, not Riemann integrable. For this reason, some properties of the Ito stochastic integral are needed [IT01].

Definition A-1. Let $\beta(t, \omega)$ be a Brownian motion process defined on the interval $t \in [t_0, t_N]$ and let $f(t, \omega)$ be measurable with respect to $B(t)$, the smallest σ field induced by (t, ω) , $t \in [t_0, t_N]$. Then the Ito stochastic integral is defined as

$$\int_{t_0}^{t_N} f(t, \omega) d\beta(t) = \lim_{\tau \rightarrow 0} \sum_{i=0}^{N-1} f(t_i, \omega) [\beta(t_{i+1}, \omega) - \beta(t_i, \omega)] \quad (A-1)$$

where

$$f(t, \omega) = f(t_i, \omega), \quad t_i \leq t < t_{i+1}, \quad t_0 \leq t < t_N \quad (A-2)$$

and $\tau = \max \{t_{i+1} - t_i\}$.

The need for careful attention to solving stochastic integral and differential equations is elaborated on in Appendix A of [BIR1]. See also [KAI1].

The following two important properties of the Ito stochastic integral shall be stated without proofs (see [IT01] or [D001]).

Theorem A-2. Let the random vector $f(t, \omega)$ satisfy the previous definition, then

$$E\left\{\int_T f(t, \omega) d\beta(t)\right\} = 0 \quad (\text{A-3})$$

and

$$E\left\{\int_T f(t, \omega) d\beta(t) \left[\int_T f(t, \omega) d\beta(t)\right]'\right\} = P_\beta \int_T E\{f(t, \omega) \cdot f'(t, \omega)\} dt \quad (\text{A-4})$$

where $P_\beta = \text{cov}\{\beta(t, \omega)\}$.

These relations are needed to deal with stochastic differential equations of the form

$$\dot{x} = f(x, t) + G(x, t)\dot{\beta} \quad (\text{A-5})$$

which must be interpreted as a formal representation of the vector Ito stochastic differential equation

$$dx = f(x, t)dt + G(x, t)d\beta \quad (\text{A-6})$$

The solution is of the form

$$x = \int_T f(x, t)dt + \int_T G(x, t)d\beta \quad (\text{A-7})$$

where the first integral can be interpreted as an ordinary Riemann integral but the second integral has meaning only if properly interpreted since $\beta(t)$ is of unbounded variation.

An alternative to the earlier defined Ito stochastic integral is the Stratonovich integral [STR2]:

$$\int G(x,t) d\beta = \lim_{\tau \rightarrow 0} \sum_{i=1}^{N-1} G\left(\frac{x(t_i) + x(t_{i+1})}{2}, t_i\right) [\beta(t_{i+1}) - \beta(t_i)] \quad (A-3)$$

where l.i.m. is the mean-square limit. It has been shown by Wong and Zakai [WON2] that if (A-6) is defined in terms of the Stratonovich integral, then to be interpreted in the sense of Ito, it will become

$$dx = f(x,t)dt + \frac{1}{2} \frac{\partial G(x,t)}{\partial x} G(x,t)dt + G(x,t)d\beta \quad (A-9)$$

where

$$\frac{\partial G(x,t)}{\partial x} G(x,t) = \sum_{j=1}^n \sum_{k=1}^m \frac{\partial G_{jk}(x,t)}{\partial x_j} G_{jk}, \quad i=1, \dots, n \quad (A-10)$$

and the components of $d\beta$ are independent. The difference between (A-6) and (A-9) shows that the interpretation of the stochastic integral of the stochastic differential equation used to model the physical system governs the solution properties. The Ito stochastic integral will be used exclusively in this work since it is more general; however, by using the results of Wong and Zakai, transformation can be made from the Ito to the Stratonovich differential equation.

The main result from Ito [IT01] is the following theorem.

Theorem A-3. (Ito) Given (A-6), then for $\phi(x,t)$ a scalar-valued function of x and t ,

$$d\phi = \phi_t dt + \phi'_x dx + \frac{1}{2} \text{tr } GQG' \phi_{xx} dt \quad (A-11)$$

where the subscripts denote the partials, and the necessary partials do exist.

APPENDIX B

The Multivariable Normal Distribution

If a Gaussian-distributed random vector x has expected value

$$\bar{x} = E\{x\} \quad (B-1)$$

and covariance

$$V = \text{cov} \{x\} = E\{(x-\bar{x})(x-\bar{x})'\} \quad (B-2)$$

then a simple method for generating random variates with the probability density function

$$p(x) = \frac{\exp[-\frac{1}{2}(x-\bar{x})'V^{-1}(x-\bar{x})]}{|2\pi V|^{1/2}} \quad (B-3)$$

is given in the sequel. (See also [NAY1] or [BAR1].)

Using the transformation,

$$x = Tu + \bar{x} \quad (B-4)$$

where $\bar{x}$ is an n -vector given by (B-1), T is a lower triangular matrix, and u is an n -vector with unit-variance, zero-mean, independent, Gaussian-distributed random components, then the variance of x can be expressed as

$$V = TT'. \quad (B-5)$$

Due to the symmetric form of V and the fact that T is triangular, T can be determined from the relationships

$$\begin{aligned}
 t_{i1} &= v_{i1}/v_{11}^{1/2} \quad 1 \leq i \leq n \\
 t_{ij} &= (v_{ij} - \sum_{k=1}^{j-1} t_{ik}t_{jk})/t_{jj} \quad 1 < j < i \leq n \\
 t_{ii} &= (v_{ii} - \sum_{k=1}^{i-1} t_{ik}^2)^{1/2} \quad 1 < i \leq n
 \end{aligned} \tag{B-5}$$

Thus, the generation of the multivariable random vector x with covariance V requires the solution of (B-4) where T is determined in (B-6). The components of u are generated independently by a Gaussian random variate generator as in [COC1], for example.

APPENDIX C

Parameters of the Steam Turbo-Generator Filtering Problem

The parameters of the steam plant model are listed below.

- T_1 - governor time constant
- T_3 - steam valve motor time constant
- T_4 - steam system time constant
- T_5 - reheat steam system time constant
- K_1 - reciprocal of regulation coefficient
- K_2 - reheat coefficient
- M - effective rotary inertia of machine
- D - damping torque coefficient
- x_1 - valve control signal
- x_2 - valve position deviation
- x_4 - torque output deviation
- x_5 - frequency deviation
- x_6 - load deviation

The values of the above constants are chosen for a typical machine:

- | | |
|----------------|----------------|
| $T_1 = 94.25$ | $K_1 = 62.24$ |
| $T_3 = 3.8$ | $K_2 = 0.324$ |
| $T_4 = 113.1$ | $M = 18,221.6$ |
| $T_5 = 4524.0$ | $D = 5.94$ |

All of these quantities are on a per unit basis and must be transformed for real values.

In addition to the above parameters, values for the stochastic model are chosen. The frequency and frequency rate measurement errors are statistically independent with zero mean and variance

$$r_{11} = 0.0001$$

$$r_{22} = 0.0004$$

The system disturbance is a scalar with zero mean and variance

$$q = 0.0008$$

The initial value for the state covariance matrix is chosen to be very large with uncorrelated components as follows:

$$p_{11}(0) = 1.0$$

$$p_{22}(0) = 1.0$$

$$p_{33}(0) = 1.0$$

$$p_{44}(0) = 1.0$$

$$p_{55}(0) = 1.0$$

$$p_{66}(0) = 1.0$$

The initial state estimate is chosen to be zero, $\hat{x}(0) = 0$, whereas the initial true state is given the value $x(0) = 0.1$; that is, each component is set at 0.1. The sampling rate was varied widely in many different simulation runs, but a value of 2 seconds was commonly used since this value allows a reasonable computation interval and yet is not too large with respect to the system dynamics.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03145 7561